

Université de Montréal

**Amélioration des stratégies diagnostiques pour détecter la bronchopneumonie infectieuse
chez les veaux de race laitière**

Par

Julie Berman

Département de sciences cliniques

Faculté de médecine vétérinaire

Thèse présentée à la Faculté de médecine vétérinaire
en vue de l'obtention du grade de *Philosophiae doctor* (Ph. D.)
en sciences vétérinaires

Juin 2022

© Julie Berman, 2022

Université de Montréal

Département de sciences cliniques, Faculté de médecine vétérinaire

Cette thèse intitulée

**Amélioration des stratégies diagnostiques pour détecter la bronchopneumonie infectieuse
chez les veaux de race laitière**

Présentée par

Julie Berman

A été évaluée par un jury composé des personnes suivantes

Jean-Philippe Roy
Président-rapporteur

David Francoz
Directeur de recherche

Sébastien Buczinski
Codirecteur

Simon Dufour
Codirecteur

Hélène Carabin
Membre du jury

Bart Pardon
Examineur externe

Résumé

Objectif : La bronchopneumonie infectieuse (**BPI**) est une affection des voies respiratoires inférieures due à l'interaction entre des agents microbiens, des facteurs environnementaux et des facteurs propres à l'individu. Malgré des années de recherche, la BPI reste prévalente en élevage de génisses de remplacement et de veaux lourds engendrant des pertes économiques majeures et une forte consommation d'antibiotiques dans ces élevages. Le score clinique respiratoire diagnostique (**SCRD**), l'auscultation thoracique, l'échographie thoracique et la radiographie thoracique sont des tests couramment utilisés pour détecter la BPI. Cependant, leurs performances diagnostiques actuelles à savoir : leurs facultés à détecter les veaux malades (sensibilité; **Se**) et leurs facultés à détecter les veaux non malades (spécificité; **Sp**) sont sous-optimales pour détecter adéquatement les veaux de race laitière nécessitant d'être traités avec des antibiotiques et/ou des anti-inflammatoires.

L'objectif général de cette thèse était donc d'améliorer les stratégies de détection du statut actif de la BPI, statut nécessitant un traitement (d'antibiotiques et/ou d'anti-inflammatoires), afin, d'une part, de mieux diagnostiquer les veaux malades, diminuer les impacts économiques de la maladie et améliorer le bien-être des veaux ; et d'autre part, favoriser l'emploi judicieux des antibiotiques en traitant uniquement les veaux qui en ont besoin. Pour répondre à cet objectif général, 3 sous-objectifs spécifiques ont été définis : (1) développer un SCRd, en identifiant et validant les signes cliniques avec une bonne répétabilité interopérateurs pour diagnostiquer le statut actif de la BPI à l'échelle individuelle et populationnelle; (2) Améliorer les stratégies diagnostiques des tests d'imagerie médicale en comparant l'échographie et la radiographie thoracique pour détecter les lésions pulmonaires et le statut actif de la BPI; (3) Améliorer l'échographie thoracique en standardisant la technique (ajout ou non du lobe crânial droit) et l'interprétation (seuil de profondeur) afin d'en optimiser ces performances pour détecter le statut actif de la BPI.

Méthodes : (1) 800 veaux lourds de 80 lots ont été filmés à l'aide d'une caméra portative lors de leur évaluation clinique. Différentes séquences vidéo ont été aléatoirement sélectionnées et soumises à un panel d'évaluateur constitué de producteurs de veaux lourds, techniciens et médecins vétérinaires. La variabilité interopérateurs de chaque signe clinique a été évaluée. Les signes cliniques ayant la meilleure répétabilité ont été retenus pour le développement et la validation d'un

SCRD en utilisant un modèle bayésien d'analyse de classe latente. Enfin, des stratégies diagnostiques des BPI au niveau populationnel ont été développées; (2) 50 veaux hospitalisés ont reçu une échographie et une radiographie thoracique. La tomodensitométrie thoracique (CT scan) a été utilisé comme *gold standard* pour vérifier les individus positifs à l'un des deux tests. Les performances des deux tests ont été comparées pour détecter d'abord les lésions pulmonaires en comparant avec les résultats du CT scan, puis le statut actif de la BPI défini par trois experts; (3) la Se et la Sp pour diagnostiquer le statut actif de BPI de différents sites d'échographie thoracique et différents seuils de profondeur de consolidation pulmonaire ont été évalués à l'aide de modèles utilisant une analyse bayésienne à classe latente dans une population de génisse de remplacement et une population de veaux lourds.

Résultats : (1) La position des oreilles et la toux induite étaient les signes cliniques de BPI les plus répétables. Le SCRD développé et validé consiste à évaluer ces deux signes cliniques ainsi que la température rectale (anormale $T \geq 39.5^{\circ}\text{C}$) sur 10 veaux d'un lot, deux semaines après leur arrivée en parc d'engraissement. La présence de 3 veaux avec deux de ces prédicteurs implique que le lot a 94 % de chance d'avoir une prévalence du statut actif de BPI ≥ 0.10 . Avec moins de 3 veaux, le lot à 95 % de chance de ne pas avoir une prévalence ≥ 0.10 ; (2) Pour détecter les lésions pulmonaires, les Se et Sp de l'échographie thoracique étaient de 0,81 (intervalle de crédit bayésien à 95 % (**ICB95%**): 0,65; 0,92) et 0,90 (ICB95 %: 0,81; 0,96), respectivement. Les Se et Sp de la radiographie thoracique étaient de 0,86 (ICB95%: 0,62; 0,99) et 0,89 (ICB95%: 0,67; 0,99), respectivement. Pour détecter le statut actif de la BPI, les Se et Sp de l'échographie thoracique étaient de 0,84 (intervalle de confiance à 95% (**IC95%**): 0,60; 0,97) et 0,74 (IC95 %: 0,57; 0,86), respectivement. La Se et Sp de la radiographie thoracique étaient de 0,89 (IC95%: 0,67; 0,99) et 0,58 (IC95%: 0,39; 0,75), respectivement. Aucune différence n'était présente entre les deux tests pour détecter les lésions pulmonaires ou le statut actif de BPI; (3) La détection échographique de lésions de consolidation pulmonaire des sites caudaux au cœur avec une profondeur ≥ 3 cm conclut à la présence du statut actif de BPI avec une Se de 0,89 (ICB95%: 0,55; 1,00) et une Sp de 0,95 (ICB95%: 0,92; 0,98).

Conclusion : Dans cette thèse, nous avons amélioré les stratégies diagnostiques pour détecter les veaux à traiter de BPI. L'utilisation de ces résultats pour élaborer des algorithmes décisionnels

devrait ultimement permettre de diminuer les pertes économiques et de raffiner l'utilisation d'antibiotiques.

Mots-clés : bovin; complexe respiratoire bovin; tests diagnostiques; poumons; consolidation; modèle de classe latente; kappa; méthode de panel diagnostique.

Abstract

Objectives: Infectious bronchopneumonia (**BPI**) is the infection of the lower respiratory tract implying an interaction between microbial agents, environment, and host. Despite decades of research, BPI remains omnipresent in dairy and veal calves, responsible for major economic losses and antimicrobial consumption. Clinical respiratory scoring system (**CRSC**), lung auscultation, thoracic ultrasonography, and thoracic radiography are the most popular tests used to detect BPI in calves. However, the sensitivity (**Se**) (i.e., faculty to detect sick calves) and the specificity (**Sp**) (i.e., faculty to detect healthy calves) are suboptimal to accurately detect dairy and veal calves to treat.

Our main objective was to improve the diagnostic strategies to detect active BPI (BPI status needed a treatment) in order to: firstly, better diagnose sick calves, reduce BPI economic losses and improve calves' welfare; secondly, reduce antimicrobial consumption by treating only calves that need it. For this purpose, three sub objectives have been defined: (1) develop and validate a CRSC including inter operator reliable respiratory clinical signs to detect active BPI at calf-level and group-level; (2) compare thoracic ultrasonography and thoracic radiography to detect lung lesions and active BPI; (3) standardize thoracic ultrasonography technique (sites to ultrasound) and interpretation (lung consolidation depth threshold) to promote thoracic ultrasonography Se and Sp to detect active BPI.

Methods: (1) 800 veal calves from 80 batches were filmed with a portative camera during their clinical exam. Videos were randomly selected and assessed by producers, technicians, and veterinarians. Reliability of each respiratory clinical sign was assessed. The most reliable respiratory clinical signs were kept in order to develop and validate a CRSC using latent class Bayesian analysis. Batch-level diagnostic strategies were developed; (2) Fifty hospitalized calves underwent thoracic ultrasonography and thoracic radiography. Thoracic tomography (CT scan) were used as a gold standard to check positive calves on one of both tests. Se and Sp of both tests were compared to detect lung lesions and active BPI defined by three experts; (3) Se and Sp of different ultrasound sites and different lung consolidation depth thresholds were estimated to detect active BPI using latent class Bayesian analysis in both dairy calves and veal calves' populations.

Results: (1) Ear droop/head tilt and induced cough were the most reliable respiratory clinical signs. The CRSC developed and validated implies assessing those clinical signs with rectal temperature (abnormal $T \geq 39.5^{\circ}\text{C}$) in 10 calves from a batch at two weeks after arriving at the fattening unit. Having two abnormal characteristics of those predictors in three calves implies that a batch has 94% of having an active BPI prevalence ≥ 0.10 . A batch with <3 positive calves on 10 has 95% of not having an active BPI prevalence ≥ 0.10 ; (2) For detecting lung lesions, the Se and Sp of thoracic ultrasonography were 0.81 (95% Bayesian credible interval (**95%BCI**): 0.65; 0.92) and 0.90 (95%BCI: 0.81; 0.96), respectively. The Se and Sp of thoracic radiography were 0.86 (95%BCI: 0.62; 0.99) and 0.89 (95%BCI: 0.67; 0.99), respectively. For detecting active BPI, the Se and Sp of thoracic ultrasonography were 0.84 (95% confidence interval (**95%CI**): 0.60; 0.97) and 0.74 (95%CI: 0.57; 0.86), respectively. The Se and Sp of thoracic radiography were 0.89 (95%CI: 0.67; 0.99) and 0.58 (95%CI: 0.39; 0.75), respectively. There was no difference between both tests to detect both lung lesions and active BPI; (3) Thoracic ultrasonography of the sites caudal of the heart with a depth lung consolidation ≥ 3 cm has a Se of 0.89 (95%BCI: 0.55; 1.00) and a Sp of 0.95 (95%BCI: 0.92; 0.98).

Conclusion: We improved diagnostic strategies to detect active BPI in this thesis. Using our results to elaborate decisional algorithms would reduce economic losses and antimicrobial consumption of BPI in dairy and veal calves.

Keywords: cattle; bovine respiratory disease; diagnostic tests; lung; consolidation; latent class model; kappa; panel diagnostic method.

Table des matières

Résumé.....	3
Abstract	6
Table des matières	8
Liste des tableaux	21
Liste des figures	27
Liste des sigles et abréviations	29
Remerciements.....	32
INTRODUCTION	35
Chapitre 1 – Revue de littérature	36
PARTIE I : Définition de la bronchopneumonie infectieuse chez les veaux de race laitière	36
I- La pathophysiologie de la bronchopneumonie infectieuse	36
I-1- Le poumon sain.....	36
I-1-1- L’anatomie du poumon.....	36
I-1-2- Les mécanismes de défense du poumon.....	38
I-2- Le poumon infecté	38
I-2-1- Les voies d’infection du poumon	38
I-2-2- Les facteurs favorisant la bronchopneumonie infectieuse.....	39
i- Les facteurs environnementaux	39
ii- Les facteurs de stress	39
iii- Les facteurs intrinsèques au veau	40
iv- Les agents microbiens.....	40

I-2-3- Les lésions de bronchopneumonie infectieuse	41
II- Les manifestations cliniques de la bronchopneumonie infectieuse	43
III- La définition diagnostique de la bronchopneumonie infectieuse.....	46
III-1- La nécessité d'améliorer la définition diagnostique de la bronchopneumonie infectieuse.....	46
III-2- Les définitions diagnostiques actuelles de la bronchopneumonie infectieuse.....	47
III-2-1- Distinction bronchopneumonie clinique <i>versus</i> non clinique.....	47
III-2-2- Distinction infection des voies respiratoires supérieures <i>versus</i> voies respiratoires inférieures	48
III-2-3- Distinction bronchopneumonie non-infectieuse/infectieuse virale <i>versus</i> bronchopneumonie infectieuse bactérienne	49
III-3- Vers une nouvelle définition diagnostique de bronchopneumonie infectieuse chez les veaux de race laitière: distinction bronchopneumonie infectieuse active <i>versus</i> non-active	51
PARTIE II : Importance d'améliorer la détection de la bronchopneumonie infectieuse chez les veaux de race laitière	53
I- La prévention et le contrôle actuels de la bronchopneumonie infectieuse chez les veaux de race laitière	53
I-1- Chez les génisses de remplacement	53
I-1-1- Le contexte d'élevage	53
I-1-2- Les mesures de prévention et de contrôle de la bronchopneumonie infectieuse	53
I-2- Chez les veaux lourds	55
I-2-1- Le contexte d'élevage	55
I-2-2- Les mesures de prévention et de contrôle de la bronchopneumonie infectieuse	56
I-3- Les limites d'efficacité de ces mesures de prévention et de contrôle chez les veaux de race laitière	58
II- Les conséquences actuelles de la bronchopneumonie infectieuse	59

II-1- Les impacts économiques de la bronchopneumonie infectieuse	59
II-1-1- Les impacts liés au problème de bronchopneumonie infectieuse	59
i- Les pertes visibles	59
ii- Les pertes invisibles	61
II-1-2- Les impacts liés aux dépenses et réaction aux problèmes de bronchopneumonie infectieuse.....	61
II-2- Les impacts sur la santé de la bronchopneumonie infectieuse	62
II-3- Les impacts sur la consommation d'antibiotiques de la bronchopneumonie infectieuse	63
II-3-1- L'utilisation d'antibiotiques	63
II-3-2- La contribution au développement de résistance	63
III- L'intérêt d'améliorer la détection de la bronchopneumonie infectieuse.....	65
III-1- Réduction des impacts économique et sur la santé de la bronchopneumonie infectieuse.....	65
III-2- Réduction de la consommation d'antibiotique pour la bronchopneumonie infectieuse	65
PARTIE III : Les tests diagnostiques actuels pour détecter la bronchopneumonie infectieuse chez les veaux de race laitière.....	67
I- Les scores cliniques respiratoires diagnostiques.....	67
I-1- Les différents scores cliniques respiratoires diagnostiques rapportés chez les veaux de race laitière	68
I-2- L'interprétation des scores cliniques respiratoires diagnostiques.....	68
I-3- Les scores cliniques respiratoires diagnostiques <i>versus</i> l'observation et l'interprétation des signes cliniques	70
I-3-1- Variation d'expression et de détection des signes cliniques de la bronchopneumonie infectieuse.....	70

I-3-2- L'objectivité des scores cliniques respiratoires diagnostiques par rapport à l'observation et l'interprétation des signes cliniques	71
I-4- L'utilisation des scores cliniques respiratoires diagnostiques chez les veaux de race laitière.....	72
II- L'auscultation pulmonaire.....	73
II-1- La technique d'auscultation pulmonaire	73
II-2- L'interprétation de l'auscultation pulmonaire	73
III- Les tests d'imagerie médicale	75
III-1- L'échographie thoracique	76
III-1-1- La technique d'échographie thoracique.....	76
i- Première description de la technique	76
ii- Simplification de la technique	76
iii- Les ajouts à la technique initiale.....	77
III-1-2- L'interprétation de l'échographie thoracique	78
i- La lésion la plus évocatrice de bronchopneumonie infectieuse.....	78
ii- La mesure objective des lésions de consolidation pulmonaire pour détecter la bronchopneumonie infectieuse.....	80
III-2 La radiographie thoracique.....	82
III-2-1 La technique de radiographie thoracique	82
i- La première description de la technique	82
ii- Simplification de la technique	83
iii- Les ajouts à la technique initiale.....	83
i- La lésions la plus évocatrice de bronchopneumonie infectieuse	84
ii- La mesure objective des lésions alvéolaires pour détecter la bronchopneumonie infectieuse.....	86

PARTIE IV : Les limites de performance des tests actuels pour détecter la BPI chez les veaux de race laitière.....87

I- Les limites de développement des tests pour détecter la bronchopneumonie infectieuse.....87

I-1- Les étapes générales nécessaires pour développer un test87

I-1-1- Le développement d'un test.....87

I-1-2- Le développement d'un score clinique diagnostique.....88

I-2- Le développement des tests pour détecter la bronchopneumonie infectieuse89

I-2-1- L'inexactitude de la définition de la bronchopneumonie infectieuse89

I-2-2- Les limites de développement90

I-2-3- Les limites de standardisation.....90

II- Les limites de détermination des performances des tests diagnostiques pour détecter la bronchopneumonie infectieuse.....92

II-1- Procédure générale de détermination des performances d'un test diagnostique.....92

II-1-1- Test *gold standard* versus test standard de référence.....92

II-1-2- Les paramètres de performances individuelles93

II-1-3- Les paramètres de performances de groupe95

II-1-4- Les paramètres de performances de plusieurs tests.....97

II-1-5- Les facteurs affectant la détermination des performances d'un test97

i- Les facteurs affectant la précision.....97

ii- Les facteurs affectant l'exactitude98

II-2- Les performances actuelles des tests diagnostiques pour détecter la bronchopneumonie infectieuse chez les veaux de race laitière.....99

II-2-1- Les performances individuelles.....99

i- Les scores cliniques respiratoires diagnostiques.....99

ii- L'auscultation thoracique.....101

iii-	Les tests d'imagerie médicale.....	102
II-2-2-	Les performances de groupe	104
II-2-3-	Les performances de plusieurs tests	104
III-	Les limites de répétabilité et reproductibilité des tests pour détecter la bronchopneumonie infectieuse chez les veaux de race laitière.....	105
III-1-	Procédure générale d'évaluation de la répétabilité d'un test diagnostique.....	105
III-1-1-	Les indicateurs de répétabilité d'un test	105
III-1-2-	L'interprétation des indicateurs de répétabilité	108
III-1-3-	Les facteurs affectant l'estimation des indicateurs de répétabilité :	108
i-	Les facteurs affectant la précision.....	108
ii-	Les facteurs affectant l'exactitude	109
iii-	Les facteurs affectant l'interprétation des indicateurs de répétabilité.....	109
III-2-	La répétabilité et la reproductibilité des tests diagnostiques pour détecter la BPI chez les veaux de race laitière	110
III-2-1-	Les scores cliniques respiratoires diagnostiques	110
III-2-2-	L'auscultation thoracique	110
III-2-3-	Les tests d'imagerie médicale.....	110
PARTIE V : Conclusion et objectif général.....		114
Chapitre 2 – Développement d'un score clinique respiratoire diagnostique pour détecter la bronchopneumonie infectieuse chez les veaux lourds.....		116
ARTICLE 1.....		116
ABSTRACT		116
I- INTRODUCTION.....		118
II- MATERIALS AND METHODS		119
II-1-	Collection of the videos.....	119
II-1-1	Selection of calves.....	119

II-1-2- Recording of videos	120
II-1-3- Selection of videos	120
II-2- Assessment of the videos by the raters	122
II-2-1- Selection and training of the raters.....	123
II-2-2- Assessment of the videos	124
II-3- Statistical analyses.....	124
II-3-1- Power sample size justification of number of videos and rater	124
II-3-2- Assessment of different scoring combinations	125
II-3-3- Assessment of the inter-rater agreement	125
II-3-4- Interpretation of the inter-rater agreement	126
II-3-5- Differences between veterinarians, technicians, and producers.....	127
III- RESULTS.....	127
III-1- Participants, videos, and raters.....	127
III-2- Overall inter-rater agreement for the different scoring combinations	127
III-3- Inter-rater agreement between veterinarians, technicians and producers	132
IV- DISCUSSION	134
V- CONCLUSION	138
CONCLUSION ARTICLE 1	139
ARTICLE 2	140
ABSTRACT	140
I- INTRODUCTION.....	143
II- MATERIALS AND METHODS	145
II-1- Step 1: Developing a new clinical score (VcCRS): attributing weights to each clinical sign using a quantitative method and choosing a cut-off to detect active BRD (800 veal calves).....	146
II-1-1- Study population	146

II-1-2- Collection of data	146
II-1-3- VcCRS development.....	148
II-1-4- Sample size calculation.....	152
II-1-5- From the logistic analysis to the VcCRS	152
II-1-6- Assessment of model sensitivity to priors.....	152
II-1-7- Software	152
II-1-8- Determination of the optimal cut-off	152
II-2- Step 2: External validation of the developed score and estimation of its performance at the individual level using different populations (209, 313, and 722 veal calves).....	154
II-3- Step 3: Simulation: Comparing and choosing interpretation for batch-level treatment decision (800 veal calves from 80 batches)	155
III- RESULTS.....	157
III-1- Step1: Development of the VcCRS	157
III-1-1- Study population.....	157
III-1-2- Descriptive statistics	159
Clinical signs (predictors):	159
Reference standard tests:.....	159
III-1-3- Final VcCRS.....	161
III-1-4- Optimal cut-off	162
III-2- Step2: External validation of the VcCRS	164
III-3- Step 3: Application at batch level	164
IV- DISCUSSION	166
V- CONCLUSION	170
SUPPLEMENTARY METHODS	171
CONCLUSION ARTICLE 2	179

Chapitre 3 – Comparaison entre l'échographie thoracique et la radiographie thoracique chez les génisses de remplacement180

ARTICLE 3.....	180
ABSTRACT	181
I- INTRODUCTION.....	182
II- MATERIALS AND METHODS	183
II-1- Population.....	183
II-2- Data collection.....	184
II-2-2- Thoracic ultrasonography.....	184
II-2-1-General examination and arterial blood sampling	185
II-2-3- Thoracic radiography	185
II-2-4- Thoracic CT-Scan	186
II-3- Definitions	187
II-4- Statistical analyses.....	187
II-4-1- Bayesian modeling for stage one	188
II-4-2- Bayesian modeling for stage two	189
II-4-3- Prior distributions.....	189
II-4-4- The sequential use of TR and TUS	190
II-4-5- Comparison between TR and TUS	190
II-4-6- Assessment of model sensitivity to assumptions	191
II-4-7- Assessment of model sensitivity to priors.....	191
II-4-8- Software	191
II-4-9- Models assessment and fit.....	192
III- RESULTS.....	192
III-1- Population	192

III-2- Thoracic lesions	194
III-3- Comparison between TUS and TR	194
III-4- Sequential use of TR and TUS.....	196
III-5- Assessment of the Bayesian model.....	196
IV- DISCUSSION	198
V- CONCLUSION	202
SUPPLEMENTARY METHODS	204
CONCLUSION ARTICLE 3	209
ARTICLE 4	210
ABSTRACT	210
I- INTRODUCTION.....	211
II- MATERIALS AND METHODS	214
II-1- Study sample	214
II-2- Clinical definition of active BP	215
II-3- Index tests.....	217
II-3-1- Thoracic ultrasonography.....	217
II-3-2- Thoracic radiography	218
II-4- Statistical analysis	219
II-4-1- Assessment of the inter-expert agreement	219
II-4-2- Estimation of performances of TUS and TR.....	219
II-4-3- Comparison of TUS and TR	220
II-4-4- Sequential use of TUS and TR.....	220
III- RESULTS.....	221
III-1- Study sample.....	221
III-2- Clinical diagnosis of active BP	222

III-3- Performances of TUS and TR.....	222
III-4- Comparison between TUS and TR	225
III-5- Sequential use of TUS and TR.....	225
IV- DISCUSSION	225
V- CONCLUSION	230
SUPPLEMENTARY METHODS	230
SUPPLEMENTARY RESULTS	233
CONCLUSION ARTICLE 4	235
Chapitre 4 – Amélioration de la technique de l’échographie thoracique pour détecter la bronchopneumonie infectieuse chez les veaux de race laitière	236
ARTICLE 5	236
ABSTRACT	237
I- INTRODUCTION.....	238
II- MATERIALS AND METHODS	241
II-1- Study design	241
II-2- Participants	241
II-2-1- Population of veal calves	241
II-2-2- Population of dairy calves.....	242
II-3- Test methods.....	242
II-4- Statistical analysis	244
II-4-1- Main model (initial LCM).....	244
II-4-2- Priors	245
II-4-3- Sensitivity analysis - Assumption of conditional independence.....	245
II-4-4- Sensitivity analysis - Assumption of constant Se and Sp across populations..	246
II-4-5- Evaluation of the test’s practical application	246

II-4-6- Assessment of convergence and statistical software.....	246
III- RESULTS.....	247
III-1- Descriptive Statistics.....	247
III-2- Comparison between TUS case definition.....	249
III-3- Sensitivity analysis.....	251
III-4- Illustration of PPV and NPV of TUS.....	253
IV- DISCUSSION	254
V- CONCLUSION	257
CONCLUSION ARTICLE 5	235
Chapitre 5 – Discussion générale	258
I- Résumé des résultats et réponse à nos objectifs	258
I-1- Amélioration des tests pour détecter les veaux de race laitière à traiter et apporter une définition diagnostique du statut actif de bronchopneumonie infectieuse	258
I-1-1- De la définition conceptuelle à la définition opérationnelle de la bronchopneumonie infectieuse.....	258
I-1-2- Les tests sélectionnés	259
I-1-3- Amélioration du score clinique respiratoire diagnostique	260
i- Amélioration du développement du score clinique respiratoire diagnostique	260
ii- Amélioration de la simplification et de l'utilisation du score clinique respiratoire diagnostique	262
I-1-4- Amélioration de la technique d'échographie thoracique	264
I-2 Estimation de l'exactitude des tests diagnostiques pour détecter le statut actif de la bronchopneumonie infectieuse.....	264
I-2-1- Les performances du score clinique respiratoire diagnostique pour détecter la prévalence de bronchopneumonie infectieuse active dans un lot	264
I-2-2- Détermination des performances de l'échographie thoracique	265

I-2-3- Détermination des performances de la radiographie thoracique	267
II- Validité des résultats	267
II-1- L'exactitude des performances des tests sélectionnés	267
II-2- La répétabilité des tests sélectionnés	269
II-3- La précision des tests sélectionnés	272
II-4- La validation externe de nos résultats	273
III- Des résultats de notre thèse vers l'élaboration d'algorithmes décisionnels	275
III-1- À l'échelle individuelle.....	275
III-1-1- Les utilisateurs	276
III-1-2- Les tests de dépistage de la bronchopneumonie infectieuse.....	277
III-1-3- Le traitement à donner	278
III-1-4- Suivi dans le temps	279
III-2- À l'échelle d'un groupe de veaux	279
IV- Impacts anticipés de nos résultats	280
IV-1- Les impacts anticipés sur l'économie de l'industrie des veaux de race laitière.	280
IV-1-1- le gain économique anticipé	280
IV-1-2- Les études futures à envisager	281
IV-2- Les impacts anticipés sur l'utilisation des antibiotiques chez les veaux de race laitière.....	282
IV-2-1- À l'échelle individuelle :.....	282
IV-2-2- À l'échelle d'un groupe de veaux	282
CONCLUSION	284
Références bibliographiques	285
Annexes	314

Liste des tableaux

Tableau 1. – Principaux agents microbiens isolés et considérés comme pathogènes lors de la bronchopneumonie infectieuse chez les veaux de race laitière, d’après Pardon et Buczinski (2020).....	41
Tableau 2. – Définition diagnostique de la bronchopneumonie infectieuse selon les résultats du score de Wisconsin et du score d’échographie thoracique de Ollivett et Buczinski (2016).	48
Tableau 3. – Principaux tests disponibles pour mettre en évidence un agent microbien de bronchopneumonie infectieuse chez un veau de race laitière.	50
Tableau 4. – Impacts de la bronchopneumonie infectieuse au début de vie sur la reproduction chez les génisses de remplacement.	61
Tableau 5. – Principaux tests diagnostiques disponibles pour mettre en évidence les symptômes, les lésions et les conséquences de la bronchopneumonie infectieuse chez les veaux de race laitière.....	67
Tableau 6. – Principaux scores cliniques respiratoires diagnostiques rapportés pour diagnostiquer la bronchopneumonie infectieuse chez les veaux de race laitière.	69
Tableau 7. – Sons audibles lors de l’auscultation thoracique chez les veaux de race laitière, d’après Curtis et al. (1986).....	74
Tableau 8. – Images échographiques normales et anormales détectables lors de bronchopneumonie infectieuse chez les veaux de race laitière.....	79
Tableau 9. – Différents seuils et sites échographiques rapportés dans la littérature pour détecter la bronchopneumonie infectieuse chez les veaux de race laitière.	81
Tableau 10. – Limites des techniques, des interprétations et de l’accessibilité des tests diagnostiques couramment utilisés pour diagnostiquer la bronchopneumonie infectieuse chez les veaux de race laitière.....	86
Tableau 11. – Étapes de développement des scores cliniques respiratoires diagnostiques rapportés pour détecter la bronchopneumonie infectieuse chez les veaux de race laitière.	91
Tableau 12. – Avantages et inconvénients des méthodes d’élaboration d’un test standard de référence.....	93
Tableau 13. – Paramètres couramment utilisés pour décrire les performances individuelles et de groupe d’un test diagnostique dichotomique.....	96

Tableau 14. – Principaux biais affectant l’exactitude des performances d’un test diagnostique.....	99
Tableau 15. – Sensibilités et spécificités des scores cliniques respiratoires diagnostiques (test index) pour détecter la bronchopneumonie infectieuse (BPI) chez les veaux de race laitière.....	100
Tableau 16. – Sensibilités et spécificités de l’auscultation thoracique (test index) pour détecter la bronchopneumonie infectieuse chez les veaux de race laitière.....	101
Tableau 17. – Sensibilités et spécificités de l’échographie thoracique (test index) pour détecter la bronchopneumonie infectieuse chez les veaux de race laitière.....	103
Tableau 18. – Principaux indicateurs utilisés pour quantifier la répétabilité inter et intra opérateur d’un test diagnostique catégorique.....	107
Tableau 19. – Échelles d’interprétation des indicateurs de répétabilité intra et inter opérateurs.....	108
Tableau 20. – Répétabilité des tests diagnostiques couramment utilisés pour détecter la bronchopneumonie infectieuse (BPI) chez les veaux de race laitière.....	113
Tableau 21. – Répartition des articles scientifiques constituant cette thèse ainsi que les différents auteurs et leurs contributions respectives.....	115
Tableau 22. – Scoring system used by the 18 raters (6 veterinarians, 6 technicians, and 6 producers) to assess the severity of clinical signs of bovine respiratory disease (BRD) of 524 filmed veal calves.....	122
Tableau 23. – Experience years (1 to <5; 5 to <10; +10) of work in calves or veal calves’ health of the 6 veterinarians (V1 to V6), the 6 technicians (T1 to T6) and the 6 producers (P1 to P6) assessing videos of clinical signs associated with bovine respiratory disease of 524 veal calves.....	123
Tableau 24. – Number of videos of the 254 filmed veal calves for each score assigned to each Bovine Respiratory Disease clinical sign by 18 raters: 6 veterinarians (V1 to V6); 6 technicians (T1 to T6), and 6 producers (P1 to P6).....	128
Tableau 25. – Overall inter-rater agreements (and 95% CI) of the different scoring combination of each bovine respiratory disease (BRD) clinical sign from 254 filmed veal calves; inter-rater agreement was estimated with the median percentage agreement (Pa), the median weighted (scoring categories > 2) or unweighted (scoring categories < 2) kappa (κ), and the Gwet’s	

agreement coefficient type 1 (AC1). The scoring combination with the higher inter-rater agreement are highlighted in grey for each BRD clinical sign. 129

Tableau 26. – Probabilistic statistical approach to interpret the magnitude of the inter-rater bovine respiratory disease (BRD) clinical signs agreements from 254 filmed veal calves, for the 2-levels scoring combination: 0,1/2,3 (absence-mild/moderate-severe) for nasal discharge, ocular discharge, and ear droop/head tilt; or 0/1,2 (absence/presence) for abnormal respiration and induced cough. For any given agreement coefficient (median percentage agreement (Pa), median weighted (scoring categories > 2) or unweighted (scoring categories < 2) kappa (κ), and Gwet’s agreement coefficient type 1 (AC1)), the probability that the extent of agreement falls into each of the benchmark intervals defining by Landis and Koch (1977) was quantified (A). The cumulative probability for each interval was then computed (B). The final benchmark level was determined by the interval associated with the smallest cumulative probability that exceeds 0.95 (highlighted in grey)..... 131

Tableau 27. – Inter-veterinarian, -technician, and -producer agreement and their 95% CI of each bovine respiratory disease (BRD) clinical sign from 254 filmed veal calves for the 2-levels scoring combination: 0,1/2,3 (absence-mild/moderate-severe) for nasal discharge, ocular discharge, and ear droop or head tilt; or 0/1,2 (absence/presence) for abnormal respiration and induced cough.¹ 133

Tableau 28. – Demographic and clinical features of the validation populations. Predictors (cough, increased rectal temperature, and ear droop/head tilt) were measured identically as in the score development population but sampled previously by other investigators (temporal validation). The target condition (active bovine respiratory disease) was estimated by the same reference tests (thoracic ultrasonography and serum haptoglobin concentration) as in the developing population in population 1, but only by thoracic ultrasonography in population 2. 154

Tableau 29. – Descriptive data (milk/grain fed batches, season, breed, sex, treatment before sampling, co-morbidity and bacteria isolated from a deep naso-pharyngeal sample) of the random 800 veal calves included in the study. Data collection was performed 2 weeks after arrival at the fattening unit. Calves were around 3 to 4 weeks old. 158

Tableau 30. – Posterior median distributions and 95% Bayesian credible intervals (BCI) regression coefficients and their weights for increased rectal temperature, ear droop/head tilt, and

cough using the sensitivity (Se) and specificity (Sp) of thoracic ultrasonography (TUS) (Main Model) or the Se and Sp of haptoglobin concentration (Hap) (Sensitivity analysis scenario). ...161

Tableau 31. – Posterior median and 95% credible intervals based on 4 Bayesian latent class models (LCM) for the prevalence of active bovine respiratory disease, sensitivity (Se) and specificity (Sp), positive (PPV) and negative (NPV) predictive values, positive (LR+) and negative (LR-) likelihood ratios of the clinical respiratory score (VcCRS), and Se and Sp of thoracic ultrasound (TUS) and concentration of haptoglobin (Hap) used to diagnose active bovine respiratory disease.164

Tableau 32. – Descriptive data of the 47 calves included in the study (≥ 7 days old; $pCO_2 \leq 53$ mmHg; able to stand; ≤ 100 kg).193

Tableau 33. – Descriptive data of the thoracic lesions identified on thoracic ultrasonography (TUS), thoracic radiography (TR), and thoracic computed tomography (CT) in the 47 dairy calves included in the study.194

Tableau 34. – Descriptive data summarizing the comparison between thoracic ultrasonography (TUS) and thoracic radiography (TR) in the 47 calves included in the study. The first step corresponds to the direct comparison between TUS and TR. The second step highlights the prior densities and median estimates (95% Bayesian credibility interval (BCI)) from the comparison of TUS and TR combined based on the presence or absence of computed tomography (CT) thoracic lesions.195

Tableau 35. – Posterior densities and median estimates, including 95% Bayesian credibility interval (BCI) of test sensitivity, test specificity, predictive positive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+), and negative likelihood ratio (LR-), for thoracic ultrasonography (TUS) and thoracic radiography (TR) using a two-stage Bayesian method.195

Tableau 36. – Posterior median and 95% credible intervals based on four plausible Bayesian latent class models for the prevalence of bronchopneumonia, the sensitivity (Se) and specificity (Sp) of thoracic radiography (TR), thoracic ultrasonography (TUS), and computed tomography (CT) used to diagnose BP in 47 dairy calves.197

Tableau 37. – Comparison of the two-stage model the two-stage model used in this study with a one-stage model considering two populations: one population including calves where thoracic

ultrasonography (TUS), thoracic radiography (TR), and thoracic computed tomography(CT) were performed, and another population including only calves where TUS and TR were performed.203

Tableau 38. – Relevant information collected by one single ACVIM-FA board certified veterinarian on the 50 calves included in the study and used for the panel diagnosis method. ...216

Tableau 39. – Descriptive data of the 50 calves recruited in the study (≥ 7 days-old; $pCO_2 \leq 53$ mmHg; ≤ 100 kg; able to stand).....222

Tableau 40. – 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.223

Tableau 41. – 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.....224

Tableau 42. – 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.....224

Tableau 43. – Number of calves with a given tests results profile using three tests (thoracic ultrasonography [TUS], concentration of haptoglobin [Hap], calf respiratory score [CRSC]) on the two populations (veal and dairy calves) as function of TUS case definition.248

Tableau 44. – Estimates of sensitivity (Se), specificity (Sp), positive (LR+) and negative (LR-) likelihood ratio of systematic thoracic ultrasonography (TUS) using Bayesian latent-class models of the main model for the different TUS case definitions in dairy and veal calves during the pre-weaning period.....250

Tableau 45. – Posterior median and 95% credible intervals based on four plausible Bayesian latent class models (LCM) for the prevalence of active pneumonia in two different populations, and sensitivity (Se) and specificity (Sp) of the clinical respiratory score chart (CRSC), thoracic ultrasound (TUS) (with the TUS case definition of caudal ≥ 3 cm) and dosage of haptoglobin (Hap) used to diagnose active pneumonia in dairy and veal calves during the pre-weaning period.....252

Tableau 46. – Comparaison des caractéristiques de développement des scores cliniques respiratoires diagnostiques précédemment décrits chez les veaux de race laitière avec celui développé et validé dans cette thèse.....261

Tableau 47. – Comparaison des caractéristiques des scores cliniques respiratoires diagnostique précédemment décrits chez les veaux de race laitière avec le score clinique respiratoire diagnostique développé et validé dans cette thèse.	263
Tableau 48. – Données actualisées des sensibilités et spécificités de l'échographie thoracique pour détecter la bronchopneumonie infectieuse des veaux de race laitière à la suite de nos travaux de thèse.	266
Tableau 49. – Sensibilités et spécificités de la radiographie thoracique pour détecter la bronchopneumonie infectieuse des veaux de race laitière.	267
Tableau 50. – Principaux tests disponibles pour définir le statut actif de la bronchopneumonie infectieuse.	268
Tableau 51. – Moyens utilisés pour limiter les biais pouvant affecter l'exactitude des résultats des travaux de notre thèse.	269
Tableau 52. – Données actuelles de répétabilité des signes cliniques de la bronchopneumonie infectieuse chez les veaux de race laitière à la suite de nos travaux de thèse.	271
Tableau 53. – Performances des principaux tests de dépistage de la bronchopneumonie infectieuse chez les veaux de race laitière.	277
Tableau 54. – Proposition d'un algorithme décisionnel pour décider d'un traitement d'antibiotiques d'un lot de veaux lourds.	280
Tableau 55. – Changements prédictifs de productivité en réponse à la réduction de l'incidence de bronchopneumonie infectieuse (BPI) et bénéfique de l'éradication dans le secteur de génisse de remplacement et de veaux lourds. Adapté de Delabouglise et al. (2017).	281

Liste des figures

Figure 1. – Schéma anatomique des voies respiratoires supérieures (en jaune) et inférieures (en orange) d'un veau de race laitière (A : côté gauche ; B : côté droit).	37
Figure 2. – Images radiographiques normales et anormales détectables lors de bronchopneumonie infectieuse chez les veaux de race laitière.	85
Figure 3. – Flowchart of the selection (4 steps) of the videos from the 254 filmed veal calves per category of BRD clinical signs: nasal discharge, ocular discharge, ear droop/head tilt, abnormal respiration, and induced cough.....	121
Figure 4. – Boxplot showing the distribution of kappa value calculated between pairs of raters between the 6 veterinarians (blue), the 6 technicians (red), and the 6 producers (green) for the different clinical signs of bovine respiratory disease (BRD). The distribution of the kappa is displayed using seven values: minimum and maximum (whisker); first quartile (bottom of box), median (horizontal bar), and third quartile (upper of box); mean (+); and outliers (dots). Columns that do not share at least one letter are significantly different (Wilcoxon; p -value < 0.05).....	132
Figure 5. – Diagram of the latent class models (LCM): LCM-Regression and LCM-TUS performances (a); LCM-Validation (b); used for determining and validating the clinical scoring system chart (VcCRS) to detect active bovine respiratory disease (BRD) in veal calves. Priors of each model are shown.	151
Figure 6. – Upset plot showing the total set size and overlaps between the presence of the 3 predictors included in the developed score (ear droop/head tilt, cough, and increased rectal temperature), TUS results, and haptoglobin results in 795 veal calves. The number of positive tests is indicated on the y-axis. The shaded circles connected by solid lines in the lower panel show the intersecting positive test datasets.	160
Figure 7. – Misclassification cost-term (MCT) of the cut-offs used to detect active bovine respiratory disease in a population of 800 veal calves.	163
Figure 8. – Batch-level positive (continued line) and negative (dashed line) predictive values for a bovine respiratory disease prevalence threshold of 15%, considering the true proportion of positive batches of 5%, for different batch cut-off values (k) and numbers of sampled calves (n).....	165

Figure 9. – Batch level positive (HPPV) and negative (HNPV) predictive values for a minimal number of positive calves to consider batch as positive (k), true proportion of positive batches (HTP) and active bovine respiratory disease (BRD) prevalence when the clinical respiratory chart score (VcCRSC) is applied on 10 veal calves in a batch.169

Figure 10. – Data collection and time frame for the 50 calves enrolled in the study. Each test was conducted in chronological order from 1 to 4 within 24 hours. Steps 1 and 2 were performed by one experiment operator (JB). Tests in Steps 3 and 4 were analyzed by another experienced operator (IM) who was unaware of the results from the previous tests.184

Figure 11. – During the study period, a total of 50 calves were enrolled in the study. They were then classified according to the results obtained by computed tomography (CT; deemed positive, CT+, if pulmonary lesions were present), thoracic ultrasonography (TUS; deemed negative, TUS -, in the absence of pulmonary lesions or when the lesions ≤ 1 cm), and thoracic radiography (TR; deemed negative, TR-, in the absence of thoracic lesions).193

Figure 12. – Flowchart illustrating the overall flow of calves in the study.215

Figure 13. – Ultrasound images corresponding to the four lesions assessed in dairy calves based on Babkine and Blond (2009). Pleural irregularities, B-lines, and comet-tails were overlooked since their clinical relevance in calves is controversial (Ollivett et al., 2015).218

Figure 14. – Illustration of thoracic ultrasonography (TUS) sites using the TUS-based case definitions (bilateral caudal sites only, vs. right cranial site only vs. a combination of bilateral caudal sites and right cranial site).243

Figure 15. – Positive predictive values (PPV) and negative predictive values (NPV) of thoracic ultrasonography (using a consolidation threshold of ≥ 3 cm caudal to the heart), according to different plausible prevalence values of active pneumonia in dairy and veal calves during the pre-weaning period.253

Figure 16. – Proposition d'un algorithme décisionnel pour mieux détecter et traiter un veau de race laitière atteint de bronchopneumonie infectieuse.276

Liste des sigles et abréviations

α K = α - Kripendorff

BHV-1 = Herpesvirus Bovin type 1

BPI = Bronchopneumonie infectieuse

BRSV = Virus Syncytial Respiratoire Bovin

BVD-1 = Virus de la Diarrhée Virale Bovine type 1

BVD-2 = Virus de la Diarrhée Virale Bovine type 2

Co = Coronavirus type A

CT scan = Tomodensitométrie

EIC = Espace intercostal

FN = Faux négatif

FP = Faux positif

GRAAS = *Guidelines for Reporting Reliability and Agreement Studies*

HPV- = Valeur prédictive négative de groupe

HPV+ = Valeur prédictive positive de groupe

HSe = Se d'un test de groupe

HSp = Spécificité d'un test de groupe

HTP = Probabilité pré-test qu'un groupe soit positif

IC95% = intervalle de confiance obtenu lors d'une analyse statistique fréquentiste à 95%

ICB95% = intervalle de crédit bayésien obtenu lors d'une analyse statistique bayésienne à 95%

κ = kappa de Cohen

LR- = Rapport de vraisemblance négatif

LR+ = Rapport de vraisemblance positif

MCT = *Misclassification Cost Term*

OMSA = Organisation mondiale de la santé animale

Pa = Pourcentage d'accord

PABA κ = Kappa ajusté à la prévalence et aux biais

PI-3 = Virus Parainfluenza type 3

SCDR = Score Clinique Respiratoire Diagnostique

Se = Sensibilité

Se_{parallèle} = Sensibilité de deux tests utilisés en parallèle

Se_{série} = Sensibilité de deux tests utilisés en série

Sp = Spécificité

Sp_{série} = Spécificité de deux tests utilisés en série

Sp_{parallèle} = Spécificité de deux tests utilisés en parallèle

TRIPODS = *Transparence Reporting of multivariable prediction model for Individual Prognosis or Diagnosis*

VN = Vrai négatif

VP = Vrai positif

VPN = Valeur prédictive négative

VPP = Valeur prédictive positive

VRS = Voies respiratoires supérieures

VRI = Voies respiratoires inférieures

*À André Guy, producteur laitier,
qui m'a transmis sa passion des bovins et qui m'a dit un jour
que si je soignais au moins une vache dans ma vie il en serait vraiment fier.*

Remerciements

À Monsieur le Professeur Jean-Philippe ROY,

De l'Université de Montréal,

Pour m'avoir fait l'honneur de présider mon jury de thèse,

Hommages respectueux.

À Monsieur le Professeur Bart PARDON,

De l'Université de Ghent,

Pour m'avoir fait l'honneur d'évaluer mon travail,

Sincères remerciements.

À Madame la Professeur Hélène CARABIN,

De l'Université de Montréal,

Pour m'avoir fait l'honneur d'accepter de constituer mon jury de thèse,

Sincères remerciements.

À Messieurs les Professeurs Simon DUFOUR, Sébastien BUCZINSKI, et David FRANCOZ

De l'Université de Montréal,

Pour m'avoir guidé tout le long de la réalisation de ce travail,

Pour vos disponibilités, vos passions et dévouement pour la recherche et vos précieux conseils,

Veillez recevoir mes remerciements les plus sincères et le témoignage de toute ma reconnaissance.

« Parfois notre lumière s'éteint, puis elle est rallumée par un autre être humain. Chacun de nous doit de sincères remerciements à ceux qui ont ravivé leur flamme. »

Albert Schweitzer

Parce que le doctorat est une longue épreuve munie de découragement, comme lors de la réception des révisions des articles, de frustration comme lors d'échec dans les manipulations (petit clin d'œil aux fameux lavages broncho-alvéolaires), ou encore de manque d'inspiration devant une feuille blanche et des quantités phénoménales d'informations à résumer. Mais face à tous ces sentiments, plusieurs personnes ont su me soutenir, me guider et me faire grandir pour devenir la chercheuse que je suis devenue aujourd'hui. C'est pourquoi je tenais à remercier tout particulièrement :

Vous mes directeurs, David pour ton mentorat, ton soutien inconditionnel et ta bienveillance; Sébastien, pour ta disponibilité, ton investissement, tes conseils toujours pertinents et ta faculté à transmettre ta passion de la recherche; et Simon, pour ta faculté à rendre les statistiques les plus compliquées si simples; merci pour votre encadrement d'exception.

Vous mes collaborateurs et comité-conseil, la professeure Isabelle Masseur et le professeur Gilles Fecteau, pour votre contribution précieuse et votre vision clinique enrichissante; le professeur Jocelyn Dubuc et la professeure Mathilde Leclère pour vos précieux conseils, dont le plus précieux, celui de prendre son temps;

Vous mes partenaires d'échantillonnage, Judith, Eloi, Kim, Monem, techniciens d'imagerie et plusieurs étudiants français, pour vos précieuses aides dans ces heures interminables et incomptables de récolte de données;

Vous mes soutiens de rédaction, Pablo mon partenaire colocataire félin, Station W, le café Centrale, les nombreuses bibliothèques montréalaises, et surtout l'espace Théséz-Vous et leurs retraites qui m'ont initié aux tomates et à la communauté de thésard en tout genre (Stéphanie la mère de famille, Nicholas l'expert du plein air ou encore Séverine la physiothérapeute);

Vous mes ami(e)s les thésards, Jon pour me comprendre tellement sur la difficulté d'allier clinique/PhD, Phil pour ta rigueur exemplaire, tes précieux conseils sur ma revue de littérature et surtout pour être devenu mon idole avec ta thèse exceptionnelle, et Claire pour ta

passion tellement communicatrice pour ton sujet, ta faculté à gagner des *awards* sur la visibilité et surtout ton soutien inconditionnel, précieux, hilarant, et enfin pour avoir relue cette thèse le temps d'un trajet de train ou tes talents de ex-journaliste et pro-travailleuse nomade a su combattre ma dyslexie;

Vous mes ami(e)s non-thésards, Nab, Véro, Anne-So, Elo, Rox et les autres, pour m'avoir changé les idées quand il le fallait et à qui je n'éprouverai aucun remords sur votre pression innocente de savoir quand est-ce que j'allais terminer;

Vous ma famille, loin d'être des adeptes des grandes études, qui ne sait toujours pas trop ce que je fais, mais qui ont toujours su me soutenir et m'inspirer à distance.

INTRODUCTION

La bronchopneumonie infectieuse (**BPI**) est une infection virale ou bactérienne descendante d'origine multifactorielle de l'appareil respiratoire inférieur bovin. Les causes de cette maladie ne se résument pas à un agent pathogène ou un unique facteur de risque, mais à une multitude de facteurs qui varient selon la localisation géographique, le type de production, la ferme ou encore, l'individu. Cette variabilité causale fait de la BPI une maladie difficile à prévenir et à traiter efficacement et universellement.

Les veaux sont fréquemment affectés par cette maladie dans l'industrie laitière, que ce soient les génisses de remplacement gardées pour la relève, ou les veaux lourds envoyés à l'engraissement. L'infection en jeune âge de ces veaux engendre des conséquences négatives sur leur bien-être et leur productivité future. Il est donc essentiel de maintenir ces veaux de race laitière en bonne santé afin d'assurer la pérennité de l'industrie laitière.

Autre problématique majeure de cette maladie : la forte consommation actuelle d'antibiotiques nécessaire pour la contrôler. Cette forte consommation pourrait contribuer à l'émergence de résistance bactérienne problématique pour la santé publique. Des actions visant à réduire et raisonner l'utilisation des antibiotiques pour contrôler la BPI sont donc nécessaires.

Une des solutions pour réduire les pertes économiques engendrées par la maladie, améliorer la santé des veaux et raisonner l'utilisation d'antibiotiques est de perfectionner sa détection afin de traiter la BPI précocement, efficacement et raisonnablement avec des antibiotiques.

L'objectif de cette thèse était donc de perfectionner les stratégies diagnostiques afin (1) de mieux diagnostiquer les veaux infectés et ainsi améliorer leur bien-être et diminuer les impacts économiques de la BPI ; (2) de favoriser l'emploi judicieux des antibiotiques en traitant uniquement les veaux qui en ont besoin.

Chapitre 1 – Revue de littérature

PARTIE I : Définition de la bronchopneumonie infectieuse chez les veaux de race laitière

I- La pathophysiologie de la bronchopneumonie infectieuse

L'appareil respiratoire des veaux se constitue : (1) des voies respiratoires supérieures (**VRS**) incluant : la cavité nasale, les sinus paranasaux, le passage nasal, le nasopharynx, l'oropharynx, les tonsilles et la portion supérieure du larynx ; et (2) des voies respiratoires inférieures (**VRI**) incluant : la portion inférieure du larynx, la trachée, et le poumon constitué de bronches, de bronchioles et d'alvéoles (**Figure 1**) (Barone, 2001). Lorsqu'on parle de BPI, on se réfère à l'infection des VRI.

I-1- Le poumon sain

I-1-1- L'anatomie du poumon

Le poumon du veau de race laitière s'étend du 1^{er} au 11^e espace intercostal (**EIC**) (Barone, 2001). Il est constitué du côté droit : d'un lobe crânial (portion crâniale et caudale), d'un lobe moyen, d'un lobe accessoire et d'un lobe caudal ; et du côté gauche : d'un lobe crânial (portion crâniale et caudale), d'un lobe moyen et d'un lobe caudal (**Figure 1**). Les EIC crâniaux au cœur se situent sous le membre thoracique où logent le lobe crânial (portion crâniale) à droite et le thymus à gauche.

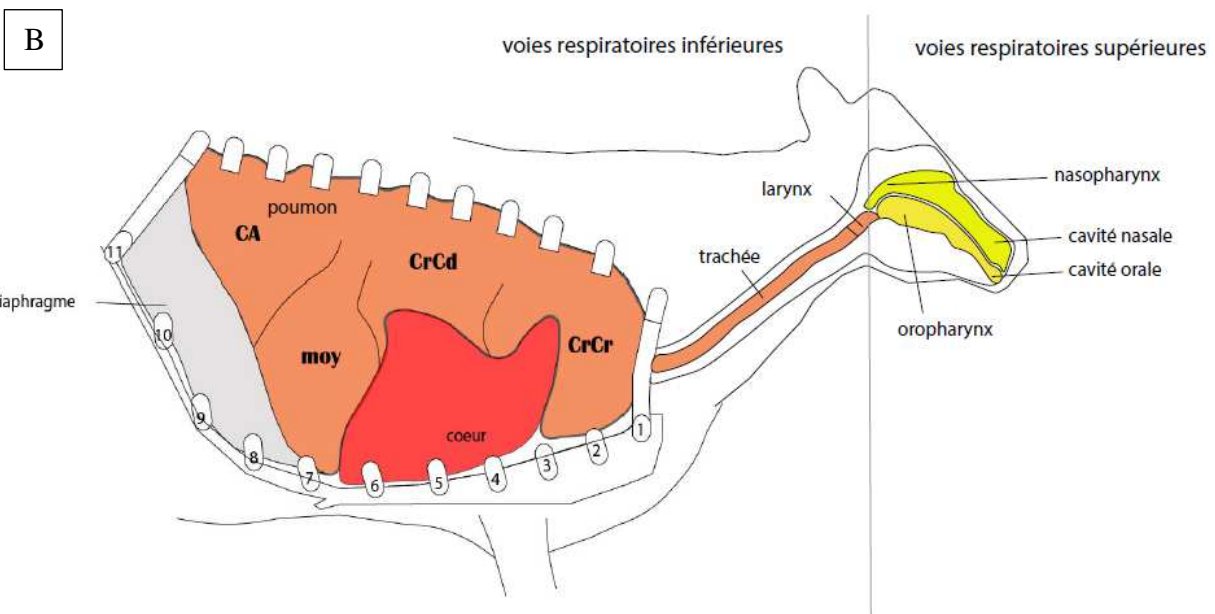
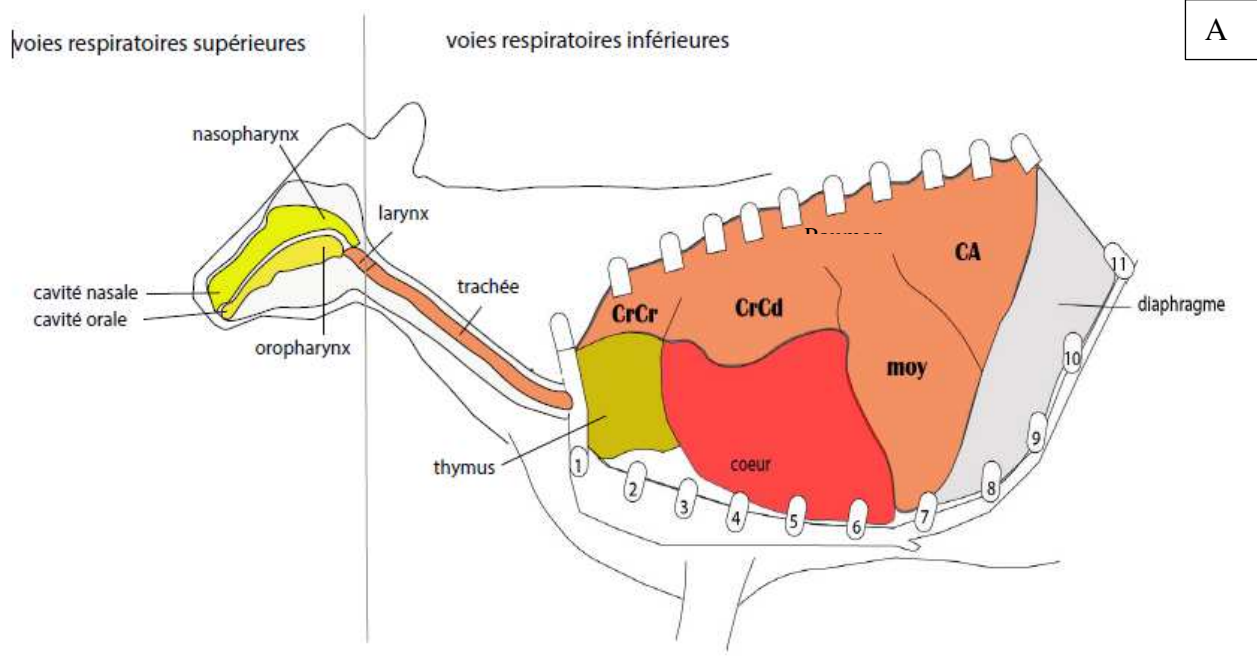


Figure 1. – Schéma anatomique des voies respiratoires supérieures (en jaune) et inférieures (en orange) d'un veau de race laitière (A : côté gauche ; B : côté droit).

CA: lobe caudal; Moy: lobe moyen; CrCd: lobe crânial portion caudale; CrCr: lobe crânial portion crâniale.
 À noter que le lobe accessoire n'est pas visible car il est situé en profondeur du poumon droit.

I-1-2- Les mécanismes de défense du poumon

Le poumon a pour principale fonction d'oxygéner le sang et d'éliminer le CO₂ produit par les cellules. Afin de réaliser cette fonction, l'épithélium des VRI est constitué d'une seule couche de cellules. Cette fine couche rend le poumon plus vulnérable aux gaz inhalés, aux particules étrangères et aux agents microbiens (Caswell, 2014). Pour pallier cette vulnérabilité, des mécanismes de défense sont présents dans les VRI (Ackermann et al., 2010, Osman et al., 2018, Zeineldin et al., 2019, McGill et Sacco, 2020) :

- La production de **mucus** qui emprisonne les particules étrangères ;
- Le **système muco-ciliaire** qui élimine les particules emprisonnées par le mucus ;
- **Les composants de l'immunité innée** (macrophages alvéolaires, neutrophiles recrutés) qui ont un effet bactéricide, anti-inflammatoire, cicatriciel et immunomodulateur ;
- **Les composants de l'immunité acquise** qui contribuent à la reconnaissance et l'élimination des agents microbiens ;
- **Le microbiote pulmonaire** qui synthétise des molécules permettant une communication avec le système immunitaire contribuant à la protection du poumon.

I-2- Le poumon infecté

I-2-1- Les voies d'infection du poumon

Le poumon d'un veau s'infecte de trois manières : (1) par aspiration où le matériel infecté atteint directement le poumon, sans migration ; (2) par voie embolique, où il y a dissémination bactérienne d'un foyer primaire *via* la circulation sanguine ; (3) ou finalement par colonisation, où les agents microbiens colonisent le poumon des VRS vers les VRI, lors de BPI (Pancier et Confer, 2010). À la différence des pneumonies par aspiration ou emboliques, où une quantité importante de bactéries est amenée directement dans le poumon, la BPI nécessite que de nombreuses défenses soient contournées avant qu'un agent microbien ne soit capable de coloniser le poumon (Caswell, 2014). L'infection peut ainsi se restreindre aux VRS si les défenses des VRI ne sont pas submergées, ou

atteindre les VRI et causer la BPI. La voie par colonisation responsable de BPI est la voie d'infection la plus fréquente (Lillie, 1974).

I-2-2- Les facteurs favorisant la bronchopneumonie infectieuse

Pour que les VRI soient infectées et que la BPI apparaisse, l'homéostasie pulmonaire à savoir l'équilibre microbiote/système immunitaire doit être perturbée (Zeineldin et al., 2019). Cette perturbation passe notamment par un affaiblissement des facteurs de défense de l'arbre respiratoire (Caswell, 2014). Plusieurs facteurs fragilisent ces défenses :

i- Les facteurs environnementaux

- **Les conditions climatiques** qui favorisent la colonisation des VRI et perturbent l'élimination des particules étrangères *via* l'appareil mucociliaire (Jericho et al., 1986; Jones, 1987, Diesel et al., 1991). L'exposition a un air froid (-10°C pour Jones [1987], 1°C pour Jericho et al [1986]), la variation d'humidité (passage de 75 % à 84 % [Diesel et al., 1991]) ou encore le changement rapide dans la température (passage de 5°C à 13°C [Diesel et al., 1991]) en sont des exemples fréquents ;
- **La qualité de l'air** où des particules inhalées comme des bactéries, des particules inertes comme la poussière ou des particules chimiques comme l'ammoniac, engendrent de l'inflammation chronique (Lago et al., 2006, Van Leenen, 2020a) et donc une perturbation de l'homéostasie pulmonaire.

ii- Les facteurs de stress

Tout stress en général fragilise les défenses du poumon (Earley et al., 2017, Zeineldin et al., 2019). Plusieurs facteurs de stress peuvent se retrouver chez les veaux de race laitière :

- **Le transport** de la ferme d'origine vers une autre localisation comme l'encan ou un parc d'engraissement (Earley et al., 2017) ;
- **Le sevrage** qui engendre un stress alimentaire (Murray et al., 2018) ;
- **Le regroupement de veaux** d'origines différentes constituant un stress social comme lors de la vente dans les encans ou la constitution de groupes au sevrage (Losinger et Heinrichs, 1996).

iii- Les facteurs intrinsèques au veau

- **La génétique** qui influence l'efficacité du système immunitaire inné (Caswell, 2014) et du microbiote (Zeineldin et al., 2019) faisant ainsi varier la tolérance du veau face à l'infection (Bassel et al., 2020) ;
- **L'immunité du veau** avec l'importance du transfert d'immunité passive et de la vaccination pour se défendre contre la colonisation des agents microbiens (Ackermann et al., 2010) ;
- **Une sous-alimentation** qui diminue l'efficacité de la réponse leucocytaire et donc la résistance aux infections (Ballou et al., 2015, Sharon et al., 2015).

iv- Les agents microbiens

L'arbre respiratoire des veaux de race laitière est naturellement composé de virus et de bactéries constituant le microbiote pulmonaire (Zeineldin et al., 2019, Zhang et al., 2019). Chaque individu possède ainsi un microbiote unique qui varie selon différents facteurs incluant : l'âge, l'environnement, le microbiote maternel, la génétique, la localisation (VRS *versus* VRI), la vaccination, les rassemblements, le transport, les changements alimentaires, le sevrage, ou encore l'exposition à de nouveaux agents microbiens (Zeineldin et al., 2019). Chez un veau, les agents microbiens responsables de BPI peuvent être déjà présents dans le microbiote et proliférer à la suite de son déséquilibre. Ils peuvent être également étrangers d'une famille différente ou d'une famille identique à un agent microbien déjà présent, mais avec des facteurs de virulence différents, et coloniser le poumon après transmission (Stanton et al., 2013, Timsit et al., 2013). C'est pourquoi on peut retrouver des agents microbiens responsables de BPI chez des individus sains comme constituant naturel de leur microbiote.

Les principaux agents microbiens isolés et considérés comme pathogènes de BPI sont présentés dans le **Tableau 1** (Pardon et Buczinski, 2020). À savoir que d'autres agents microbiens peuvent être isolés dans le poumon sans être considérés comme pathogènes (p. ex., Virus de la rhinite Bovine type A et B, *Mycoplasma bovirhinis*, *Mycoplasma dispar*) ou retrouvés lors de pneumonie par aspiration ou embolique (p.ex., *Salmonella* spp., *Escherichia coli*).

Les virus	Les bactéries
Herpesvirus Bovin type-1 (BHV-1)	<i>Mycoplasma bovis</i>
Virus Parainfluenza type-3 (PI-3)	<i>Histophilus somni</i>
Virus Syncytial Respiratoire Bovin (BRSV)	<i>Mannheimia haemolytica</i>
Virus de la Diarrhée Virale Bovine type 1 et 2 (BVD-1, BVD-2)	<i>Pasteurella multocida</i>
Adenovirus type-3	<i>Trueperella pyogenes</i> et d'autres bactéries anaérobies dans les cas chroniques
Le Coronavirus type-A (Co)	<i>Bibersteinia trehalosi</i>
Virus Influenza D	<i>Clamidia psittaci</i>
	<i>Moraxella bovis/ovis</i>

Tableau 1. – Principaux agents microbiens isolés et considérés comme pathogènes lors de la bronchopneumonie infectieuse chez les veaux de race laitière, d'après Pardon et Buczinski (2020).

Chacun de ces agents microbiens considérés pathogènes possède des facteurs de virulence capables d'engendrer des lésions pulmonaires (**Annexe 1**). Si les défenses du poumon sont optimales, la prolifération de ces agents microbiens ne causera pas nécessairement la maladie. Au contraire, si les défenses du poumon sont fragilisées ou submergées par une prolifération abondante de ces agents, la BPI va apparaître (Zeineldin et al., 2019).

En plus de leur contagiosité, ces agents microbiens sont également synergiques (Fulton et al., 2000). En effet, les virus ont la capacité de fragiliser les défenses respiratoires par immunosuppression et par altération de la surface de la muqueuse rendant l'adhésion des bactéries facilitées sur les cellules déjà infectées par le virus. La colonisation des VRI par les bactéries est alors facilitée (Griffin et al., 2010, Panciera et Confer, 2010, Caswell, 2014). Une infection virale favorise donc la surinfection bactérienne rendant l'infection souvent polymicrobienne. Cependant, la proportion de surinfection bactérienne des BPI virale reste inconnue chez les bovins de race laitière. Cela est problématique pour savoir comment traiter ces BPI virales notamment avec des antibiotiques, si le risque de surinfection est important, ou uniquement des anti-inflammatoires et attendre la fin de la période d'incubation (Buczinski et Pardon, 2020).

I-2-3- Les lésions de bronchopneumonie infectieuse

Anatomiquement, étant donné que l'infection se réalise par colonisation (migration des VRS vers les VRI), les lésions se localisent tout d'abord proches de la bifurcation des

premières bronches localisées dans les portions ventrales du poumon. Ces lésions vont ensuite évoluer selon la gravité, dans les portions dorsales du poumon. L'infection se fait indifféremment des deux côtés (Panciera et Confer, 2010, Ollivett et al., 2013).

Pathophysiologiquement, l'infection s'étend d'alvéole en alvéole par progression de la bactérie ou du virus engendrant l'extravasation vasculaire de liquides et de cellules inflammatoires vers l'espace interstitiel et les alvéoles. Une large partie du poumon, un lobe ou encore le poumon entier peuvent ainsi se remplir de liquides et de débris cellulaires remplaçant l'air. On parle alors de consolidation pulmonaire (Guyton, 2011).

Histologiquement, différentes inflammations sont retrouvées lors de BPI. Elles varient selon l'agent microbien impliqué et la réponse immunitaire de l'hôte (Yates, 1982, Ames et al., 1985, Panciera et Confer, 2010). On distingue :

- **Une inflammation broncho-interstitielle** qui correspond généralement à une infection virale non compliquée par une infection bactérienne. L'inflammation se localise à l'espace interstitiel et aux bronches. Le poumon est rarement consolidé (Panciera et Confer, 2010) ;
- **Une inflammation suppurative** qui se caractérise par une infiltration par des polynucléaires neutrophiles des tissus infectés par des bactéries. Une large partie du poumon, un lobe ou encore le poumon entier peuvent devenir consolidés. L'inflammation suppurative est l'inflammation la plus souvent retrouvée lors de BPI (Panciera et Confer, 2010) ;
- **Une inflammation fibrineuse de l'espace pleural (pleurobronchopneumonie)** est causée par l'intervention d'agents microbiens particuliers infectant l'espace pleural comme *Mannheimia haemolytica*, *Mycoplasma bovis* et *Histophilus somni*. Cette inflammation est toujours associée à une inflammation suppurative du poumon (Panciera et Confer, 2010) ;
- **L'abcédation** qui correspond à l'évolution d'une inflammation suppurative en abcès (Dagleish et al., 2010). Les agents microbiens souvent retrouvés sont *Trueperella pyogenes*, *Mycoplasma bovis* et des bactéries anaérobies ;

- **Une inflammation caséo-nécrotique** qui se caractérise macroscopiquement par la présence de nodules contenant un matériel caséux de nécrose (Caswell et Archambault, 2007). Ce type d'inflammation est caractéristique d'une infection à *Mycoplasma bovis* (Pancier et Confer, 2010) mais n'est pas systématiquement présente lors d'infection par cette bactérie (Caswell et Archambault, 2007). Elle est souvent retrouvée avec d'autres inflammations (p. ex., suppurative, fibrineuse).

II- Les manifestations cliniques de la bronchopneumonie infectieuse

Selon leurs étendues, les lésions pulmonaires entraîneront différentes répercussions cliniques chez les veaux infectés. Un veau de race laitière atteint de BPI peut présenter les signes cliniques suivant :

- **Un jetage nasal** qui survient par augmentation des sécrétions glandulaires des VRS pour éliminer l'infection. Ce jetage est séreux, mucoïde ou purulent selon la stimulation neutrophilique de l'infection. Ce signe clinique est le reflet des mécanismes de défense des VRS. Un jetage nasal n'est donc pas spécifique à la BPI, car il peut survenir lors d'infection localisée aux VRS ou en réponse à l'inhalation d'air de mauvaise qualité ou de particules alimentaires (Wilson et al., 2015, Woolums, 2015) ;
- **Un jetage oculaire** qui survient également lors d'augmentation des sécrétions glandulaires des VRS pour éliminer l'infection *via* le canal naso-lacrimal dans la cavité nasale. Tout comme le jetage nasal, ce jetage peut être séreux, mucoïde ou purulent selon la stimulation neutrophilique de l'infection. Sa présence est souvent associée au jetage nasal, mais comme ce dernier peut ne pas être toujours détectable (p. ex., élimination par léchage du veau), un jetage oculaire peut être observé seul. De plus, sa présence peut être accentuée si l'agent microbien responsable de la BPI a un tropisme oculaire comme c'est le cas lors d'infection à BHV-1 ou *Mycoplasma bovis* (Loy et al., 2021). Ce signe clinique est donc le reflet des mécanismes de défense des VRS et/ou oculaire. Ainsi, un jetage oculaire n'est pas spécifique de BPI et peut survenir lors d'infection localisée aux VRS ou d'affection oculaire (p. ex., conjonctivite) (Wilson et al., 2015, Woolums, 2015) ;

- **De la toux** qui correspond à une réponse réflexe physiologique visant à évacuer les particules étrangères emprisonnées dans le mucus *via* le système muco-ciliaire. Ce signe clinique est le reflet des mécanismes de défense des VRI. Une étude récente a montré que la toux, spontanée ou induite, était le seul signe clinique performant, associé statistiquement avec la présence de consolidation pulmonaire à l'échographie dans toutes les productions bovines (Lowie et al., 2022). Malgré cette association, la présence de toux n'est pas 100 % spécifique de BPI et peut également survenir pour se défendre contre l'inhalation d'air de mauvaise qualité ou de particules alimentaires (Wilson et al., 2015, Woolums, 2015) ;
- **De la fièvre** qui est le reflet d'un mécanisme systémique de défense. La BPI étant une maladie infectieuse, l'organisme va combattre les agents microbiens incriminés en augmentant sa température. Cependant, l'augmentation de température n'est pas forcément de la fièvre, mais peut être de l'hyperthermie due à un stress, un effort physique ou secondaire à la température ambiante. De plus, ce signe clinique n'est pas directement observable et doit être mesuré. Selon les instruments utilisés, on peut avoir une erreur de mesure ou une erreur liée à la technique de l'opérateur qui le fait. La valeur de la température est donc à interpréter avec précaution et d'autres signes systémiques sont nécessaires pour conclure à une fièvre comme un foyer d'infection. Une infection autre que la BPI (p.ex., omphalite, arthrite) peut être responsable de fièvre ce qui en fait un signe clinique non spécifique de BPI (Wilson et al., 2015, Woolums, 2015) ;
- **De la tachypnée** qui correspond à une augmentation de la fréquence respiratoire. La fréquence respiratoire normale d'un jeune veau (< 30 jours) est en moyenne de 48 respirations par minute variant entre 30-60 (Terra et Reynolds, 2015). Lorsque le veau grandit, cette fréquence respiratoire diminue atteignant une moyenne de 24 respirations par minute variant entre 12 et 36 à l'âge adulte (Terra et Reynolds, 2015). Lors d'un dysfonctionnement pulmonaire, la tachypnée permet d'améliorer la ventilation et les échanges gazeux. Cependant, une tachypnée peut également être présente lors de stress, de désordres acido-basiques, de fièvre, ou encore de

- douleur. Une tachypnée est donc non spécifique de BPI (Wilson et al., 2015, Woolums, 2015) ;
- **De la dyspnée** qui correspond à une modification de la dynamique respiratoire, soit inspiratoire, quand l'air ne rentre pas adéquatement (p. ex. ; obstruction des VRS), ou est expiratoire quand la compliance du poumon ne permet pas de faire sortir l'air adéquatement (p. ex., BPI). Une dyspnée peut être présente lors d'un dysfonctionnement de l'appareil respiratoire, mais également lors de désordres thoraciques comme des fractures de côtes ou des désordres diaphragmatiques. Une dyspnée n'est donc pas spécifique à la BPI (Wilson et al., 2015, Woolums, 2015) ;
 - **Une position anormale des oreilles** est présente lors de BPI par extension de l'infection des VRS au conduit auditif, aussi appelé trompe d'Eustache, engendrant une otite moyenne. L'infection est responsable de douleur dans la région auriculaire et peut provoquer un dysfonctionnement du nerf facial impliqué dans la position normale des oreilles. Lors d'infection grave unilatérale de l'oreille moyenne, le veau présente une tête penchée secondaire à la douleur et/ou l'atteinte neurologique (Wilson et al., 2015, Woolums, 2015). La position anormale des oreilles ou une tête penchée peut également être le reflet de l'expression d'une douleur autre qu'auriculaire chez le veau. La position des oreilles n'est donc pas spécifique d'infection de la BPI et peut survenir lors d'infection localisée aux VRS, lors d'otite externe, lors d'affection douloureuse (p. ex., boiterie) ou lors de toute autre raison d'abattement (Wilson et al., 2015, Woolums, 2015) ;
 - **De l'abattement et d'une anorexie** qui survient lors de l'expression : (1) d'un état fébrile ; (2) de douleur ; (3) d'une hypoxémie (p. ex., anémie ou dysfonctionnement pulmonaire). Ces signes peuvent donc être présents dans tout état pathologique, ce qui en fait des signes non spécifiques de BPI (Wilson et al., 2015, Woolums, 2015).

En résumé, les signes cliniques de BPI sont variables et peuvent être le reflet de : (1) l'expression des mécanismes de défense de l'appareil respiratoire contre les agents microbiens incluant les réactions mécaniques à l'irritation (toux), la réponse immunitaire locale (jetage nasal, jetage oculaire) et/ou systémique (fièvre) ; (2) l'expression du

dysfonctionnement du poumon (dyspnée, tachypnée) ou d'autres structures concomitantes infectées (position anormale des oreilles lors d'otite); (3) l'expression de la douleur secondaire à l'infection (position anormale des oreilles, abattement ou anorexie). Un veau infecté présentera alors un certain nombre de signes cliniques selon la gravité et l'étiologie de l'infection, sa faculté à exprimer la douleur, ou encore l'expression de son propre système immunitaire. Les manifestations cliniques ne sont donc pas universelles et pathognomoniques de BPI. En d'autres termes, ce n'est pas parce qu'un veau a un ou plusieurs de ces signes cliniques qu'il est atteint nécessairement d'une BPI. À l'inverse, un veau atteint de BPI peut n'avoir aucun signe clinique apparent (Apley, 2014). Cette caractéristique de la BPI rend son diagnostic clinique difficile. D'autres tests sont donc nécessaires pour détecter la BPI et définir avec exactitude et précision un veau atteint *versus* non atteint.

III- La définition diagnostique de la bronchopneumonie infectieuse

On parle de définition diagnostique, la définition de la BPI qui se base sur les résultats de tests diagnostiques pour définir un veau atteint *versus* non atteint de la maladie. Cette définition se veut opérationnelle (résultats de tests diagnostiques) plutôt que conceptuelle (malade, non malade).

III-1- La nécessité d'améliorer la définition diagnostique de la bronchopneumonie infectieuse

Pendant des années, la BPI était incluse dans le terme global de « complexe respiratoire bovin » et se définissait selon l'épidémiologie de l'infection. On parlait, par exemple, de la pneumonie enzootique lorsque des veaux d'un même troupeau étaient infectés, ou encore de fièvre de transport lorsque les veaux étaient affectés après un transport (Lillie, 1974). Bien qu'utilisée depuis des décennies, cette définition résumait davantage la complexité multifactorielle d'apparition de la maladie que ses particularités pathologiques permettant de dire si un veau était malade ou non. La décision de traiter un cas individuel était alors mal définie, limitant une détection efficace de la maladie et une utilisation raisonnée des traitements antibiotiques (Buczinski et Pardon, 2020). Une meilleure définition, basée sur le diagnostic de la BPI, est actuellement primordiale pour définir avec précision et

exactitude un veau malade afin de contrôler efficacement la maladie et d'utiliser les antibiotiques raisonnablement pour la traiter. Pour que la définition diagnostique de BPI soit la plus appropriée aux conditions d'élevage actuelles, **les veaux à traiter avec des antibiotiques** doivent être clairement définis par les résultats d'outils diagnostiques rapides, précis, performants et peu coûteux.

III-2- Les définitions diagnostiques actuelles de la bronchopneumonie infectieuse

III-2-1- Distinction bronchopneumonie clinique *versus* non clinique

Une définition dite « clinique » de la BPI a souvent été utilisée comme définition diagnostique pour détecter les veaux à traiter avec des antibiotiques de BPI. Cette définition repose sur l'observation de signes cliniques chez un veau et la mesure de la température. Cependant, des études récentes ont souligné les limites de cette définition en mettant en évidence une proportion importante de veaux (69 à 87 %) n'ayant jamais été traités pour BPI mais présentant des lésions de consolidations pulmonaires à l'échographie thoracique, signe de BPI contractée dans le passé (Ollivett et al., 2011, Buczinski et al., 2014). On peut donc conclure que, si ces veaux n'ont pas reçu de traitement, c'est qu'ils n'ont pas été observés avec des signes cliniques par les producteurs. Cela souligne les limites de détection des veaux à traiter basées sur l'expression et la détection des signes cliniques. Cette limite se retrouve également à l'échelle d'un groupe de veaux où l'évaluation des signes cliniques a été jugée inadéquate pour prédire les lots de veaux lourds avec une forte prévalence de BPI, et donc nécessitant potentiellement un traitement antibiotique de groupe pour diminuer cette prévalence (Leruste et al., 2012). Ainsi, bien que couramment utilisée, la définition diagnostique de la BPI basée sur l'observation et la détection des signes cliniques n'est **pas adéquate pour diagnostiquer précisément et avec exactitude un veau ou un groupe de veaux à traiter avec des antibiotiques.**

**III-2-2- Distinction infection des voies respiratoires supérieures
versus voies respiratoires inférieures**

Afin d'améliorer la détection des veaux à traiter avec des antibiotiques, une définition diagnostique basée sur la détection de lésions échographiques a été proposée pour améliorer la simple détection clinique (Buczinski et Pardon, 2020).

En médecine humaine, une distinction est réalisée entre les infections des VRS (p. ex., rhinite, pharyngite) *versus* VRI (bronchite, pneumonie) pour faciliter la décision de traiter l'infection aux antibiotiques (Kim et Marikar, 2019). En effet, les infections des VRS nécessitent plus rarement des antibiotiques que les infections des VRI (Kim et Marikar, 2019). Basée sur ces pratiques en médecine humaine, la localisation de l'infection a donc été suggérée comme définition diagnostique pour définir les **veaux à traiter avec des antibiotiques** chez les veaux de race laitière (Ollivett et Buczinski, 2016, Buczinski et Pardon, 2020). Cette définition diagnostique se base sur la combinaison des résultats de deux tests : le score clinique du Wisconsin (**Annexe 2**) (McGuirk, 2008) et un score d'échographie thoracique développé par Ollivett et Buczinski (2016) (**Annexe 3**). Le **Tableau 2** résume les différents statuts de BPI obtenus selon cette définition diagnostique.

Statut de la bronchopneumonie infectieuse	Score clinique ^a	Échographie thoracique ^b	Indication de traiter
Sain	Négatif	Négatif	Non
Atteinte des VRS	Positif	Négatif	Non sauf si autres conditions (otite)
Atteinte des VRI sous cliniques	Négatif	Positif	Peut-être
Atteinte des VRI cliniques	Positif	Positif	Oui

^a score clinique de Wisconsin définit comme positif si ≥ 5 , négatif si < 5

^b score d'échographie thoracique positif si ≥ 3 , négatif si < 3

Tableau 2. – Définition diagnostique de la bronchopneumonie infectieuse selon les résultats du score de Wisconsin et du score d'échographie thoracique de Ollivett et Buczinski (2016).

L'avantage de cette définition est qu'elle est basée sur les résultats de tests rapides, peu coûteux et faciles à réaliser en conditions d'élevage actuelles. Autre avantage, la gravité de la maladie peut être estimée selon l'étendue des lésions échographiques, ce qui peut être intéressant pour estimer la durée et les chances de succès du traitement. De plus, bien

qu'utilisée essentiellement par les médecins vétérinaires actuellement, la simplicité de l'échographie thoracique pourrait rendre son utilisation par les producteurs envisageables dans le futur rendant cette définition diagnostique accessible et pratique pour les producteurs à la ferme (voir Partie III).

Cependant, au-delà de leurs simplicités, ces deux tests diagnostiques utilisés présentent des limites. Premièrement, les performances modérées de détection du score de Wisconsin (**Tableau 15**) et sa répétabilité (**Tableau 20**) peuvent engendrer de mauvaises classifications et des variations de détection selon l'opérateur (voir plus de détails, partie 4). *A contrario*, bien que performant (**Tableau 17**) et répétable (**Tableau 20**) pour détecter des lésions pulmonaires chez les veaux de race laitière, les performances de l'échographie thoracique sont inconnues pour savoir si les lésions détectées sont actives, c.-à-d., infectées et enflammées, ou inactives, c.-à-d., cicatricielles ou en voie de guérison. La décision de traiter un veau avec une lésion échographique est donc encore à définir, c'est pourquoi la décision de traiter des veaux atteints de VRI sans signes cliniques (veaux sous-cliniques) n'est pas claire. Deuxièmement, avec cette définition, les veaux atteints de BPI virale (s'exprimant avec des signes cliniques et des lésions échographiques) recevront des antibiotiques, ce qui peut être discutable en absence de surinfection bactérienne. En conclusion, la combinaison de ces deux tests imparfaits pour détecter les veaux à traiter avec des antibiotiques présente des limites sujettes à l'amélioration.

III-2-3- Distinction bronchopneumonie non-infectieuse/infectieuse virale *versus* bronchopneumonie infectieuse bactérienne

Face aux incertitudes de cette dernière définition diagnostique de traiter précisément et avec exactitudes les veaux avec des antibiotiques, une définition basée sur la mise en évidence de l'agent microbien a été envisagée en compléments de l'observation des signes cliniques et des lésions échographiques (Buczinski et Pardon, 2020). Dans cette définition, la maladie se définit selon l'isolement d'un agent microbien, c.-à-d., comme étant : (1) une BPI non infectieuse et secondaire à une irritation et l'inflammation des voies respiratoires par l'inhalation d'élément polluant (Van Leenen, 2020a) ; (2) une BPI virale et ne nécessitant pas de traitement antibiotique ; et (3) une BPI bactérienne et nécessitant un

traitement antibiotique. Les tests disponibles pour isoler ces agents microbiens sont présentés dans le **Tableau 3**.

Les techniques de prélèvement	Les tests visant à mettre en évidence l'agent microbien
<ul style="list-style-type: none"> ◆ Écouvillon nasal (double ou simple gaine) ◆ Écouvillon nasopharyngé (double ou simple gaine) ◆ Lavage trachéal ◆ Lavage bronchoalvéolaire (endoscopique ou à l'aveugle) 	<ul style="list-style-type: none"> ◆ Culture microbienne ◆ Isolement viral ◆ Spectométrie de masse (MALDI-TOF MS) sur culture bactérienne enrichie ◆ PCR ◆ Cytologie* ◆ Séquençage de l'ADN
Analyse sanguine	Sérologie (ELISA anticorps).

MALDI-TOF MS: *Matrix-Assisted Laser Desorption/Ionization Time of Flight Mass Spectrometry*; ELISA: *enzyme-linked immunosorbent assay*, technique immune-enzymatique sur support solide.

*La cytologie n'est pas réalisée actuellement sur les écouvillons nasaux chez les veaux de race laitière

Tableau 3. – Principaux tests disponibles pour mettre en évidence un agent microbien de bronchopneumonie infectieuse chez un veau de race laitière.

Bien que cette définition semble prometteuse pour détecter **les veaux à traiter avec des antibiotiques** avec plus de précision et d'exactitude, les tests diagnostiques utilisés pour isoler les agents microbiens présentent des limites. Premièrement, hormis la cytologie, l'obtention des résultats des tests n'est pas instantanée, retardant l'instauration d'un traitement. De son côté, bien que plus rapide, l'interprétation de la cytologie seule manque de Se et Sp pour distinguer les BPI infectieuses bactérienne *versus* virale ou non infectieuse (Van Leenen et al., 2020b). Deuxièmement, il n'existe pas de technique de prélèvement 100 % efficace (Doyle et al., 2017, Van Driessche et al., 2017). En effet, selon la technique, l'agent microbien peut ne pas être prélevé ou ne pas être responsable de la maladie et provenir d'une contamination lors du prélèvement, ne pas être prélevé au site de l'infection, ou appartenir au microbiote respiratoire normal (Doyle et al., 2017, Van Driessche et al., 2017; Zeineldin et al., 2019). Finalement, cette définition diagnostique de la BPI est accessible uniquement pour des médecins vétérinaires, limitant son utilisation quotidienne par les producteurs.

III-3- Vers une nouvelle définition diagnostique de bronchopneumonie infectieuse chez les veaux de race laitière: distinction bronchopneumonie infectieuse active *versus* non-active

En médecine humaine, un des moyens pour distinguer les pneumonies à traiter ou non avec des antibiotiques est de doser une protéine de l'inflammation aiguë, la procalcitonine, sur des individus dont les signes cliniques et les examens d'imagerie sont en faveur d'une pneumonie (Pinzone et al., 2014). En effet, le niveau de cette protéine renseigne sur le caractère inflammatoire et infectieux des lésions pulmonaires et permet de justifier l'instauration d'un traitement antibiotique et d'en contrôler sa durée (Christ-Crain et al., 2006). En d'autres termes, lorsque des lésions pulmonaires sont présentes et que le niveau de procalcitonine est élevé, la pneumonie est considérée comme active et est traitée avec des antibiotiques. À l'inverse, lorsque des lésions pulmonaires sont présentes, mais que le niveau de procalcitonine est bas, la pneumonie est non active. Le traitement antibiotique n'est alors pas réalisé ou arrêté (Christ-Crain et al., 2006). Cette distinction basée sur la nature des lésions a permis de définir **un statut actif** de la maladie où les lésions dans le poumon sont secondaires à une infection et une inflammation nécessitant un traitement antibiotique, *versus* **un statut non actif**, où les lésions dans le poumon sont cicatricielles ou susceptibles de disparaître sans traitement. Cette distinction a permis de réduire la consommation d'antibiotique secondaire aux pneumonies chez les humains (Bosso et Drew, 2011). On retrouve cette distinction en médecine vétérinaire et notamment en médecine équine chez les poulains atteints de pneumonie à *Rhodococcus equi* (Arnold-Lehna et al., 2020). Le test utilisé est cette fois-ci l'échographie thoracique. En effet, selon la taille des lésions échographiques, la pneumonie est soit active et le poulain est traité avec des antibiotiques, soit non active, le poulain ne reçoit alors aucun traitement et les lésions disparaissent au fil du temps (Arnold-Lehna et al., 2020 ; Venner et al., 2012).

Une définition diagnostique impliquant des tests capables de savoir si les lésions pulmonaires de BPI nécessitent un traitement antibiotique et/ou anti-inflammatoire n'existe pas actuellement en médecine bovine. Le développement et l'amélioration de stratégies diagnostiques définissant la BPI comme active *versus* non active, permettraient de mieux raisonner traiter les veaux de race laitière atteint de BPI.

Conclusion PARTIE I : La BPI est une maladie multifactorielle due à l'interaction entre des agents microbiens, des facteurs environnementaux et des facteurs propres à l'individu. Par ce caractère multifactoriel, l'expression clinique de la BPI est pléomorphe et difficile à détecter. Actuellement, aucune définition diagnostique appropriée aux conditions d'élevage actuelles ne permet d'identifier précisément et avec exactitude les veaux à traiter ou non avec des antibiotiques. Une amélioration des outils diagnostiques pour distinguer le statut actif *versus* non actif de la BPI permettrait de mieux détecter et traiter la maladie.

PARTIE II : Importance d'améliorer la détection de la bronchopneumonie infectieuse chez les veaux de race laitière

I- La prévention et le contrôle actuels de la bronchopneumonie infectieuse chez les veaux de race laitière

La prévention et le contrôle de la BPI chez les veaux de race laitière varient entre l'élevage de génisses de remplacement et l'élevage de veaux lourds.

I-1- Chez les génisses de remplacement

I-1-1- Le contexte d'élevage

Les génisses de remplacement constituent la relève des vaches en production laitière. La gestion de l'élevage varie selon la localisation et la taille des fermes (Urie et al., 2018a). Dans ce type d'élevage, les veaux sont soit élevés à la ferme de naissance, soit transportés dans une ferme spécialisée dans l'élevage de génisses de remplacement (Stanton et al., 2013, Urie et al., 2018a). Durant la période présevrage, les génisses reçoivent une alimentation lactée et peuvent être logées individuellement, par paire, en groupe ou libre dans une pouponnière. Au sevrage, les génisses reçoivent une alimentation solide de fibre et de concentrés, et peuvent être rassemblées en groupe ou restées séparées en stabulation entravée (Barkema et al., 2015, Urie et al., 2018a).

Deux périodes sont à risque pour la survenue de la BPI : (1) la période de présevrage qui comprend des événements stressants (p. ex. ; écornage, transport, rassemblement) qui sont associés à la baisse naturelle des anticorps maternels. Cette période constitue la période la plus à risque de BPI (Urie et al., 2018b) ; (2) la période de sevrage qui constitue un stress alimentaire et social lors de rassemblement en groupe (Callan et Garry, 2002).

I-1-2- Les mesures de prévention et de contrôle de la bronchopneumonie infectieuse

La période de présevrage constitue une période critique pour l'apparition de la BPI chez les génisses de remplacement avec une incidence maximale de 0.08 veau-semaine entre 5 et 7 semaines d'âge (Svensson et al., 2007, Urie et al., 2018b). C'est pourquoi la majorité des mesures de prévention et de contrôle s'effectuent durant cette période et reposent sur :

- L'administration d'un **colostrum** de bonne qualité, en bonne quantité et au bon moment pour assurer un bon transfert d'immunité pour combattre la BPI et d'autres maladies (Van Donkersgoed et al., 1993, Donovan et al., 1998, Windeyer et al., 2014) ;
- Une **nutrition de qualité et en quantité adéquate** pour apporter de l'énergie nécessaire à la croissance et au bon fonctionnement du système immunitaire (Khan et al., 2011, Shivley et al., 2018, Dubrovsky et al., 2019) ;
- **Un logement favorisant une bonne qualité de l'air**. En effet, l'humidité élevée tend à fragiliser l'arbre respiratoire et favoriser la survie des agents microbiens transmis par les aérosols. Des mesures peuvent être prises pour réduire l'humidité environnante telles que réduire la densité de logement, accroître le taux de renouvellement d'air par la ventilation, garder le logement propre, améliorer le drainage de l'urine et de l'eau, et limiter les activités qui font augmenter l'humidité comme la préparation du lait de remplacement dans le même espace libre que celui des veaux (Lago et al., 2006, Svensson et Liberg, 2006, Windeyer et al., 2014) ;
- Diminuer le **contact avec des bovins plus âgés** afin d'éviter la transmission d'agents microbiens (Callan et Garry, 2002, Lundborg et al., 2005) ;
- La surveillance plus rapprochée des **génisses plus à risque** de développer la BPI comme les génisses prématurées ou qui sont nées après une dystocie (Perez et al., 1990, Lombard et al., 2007) ;
- La **vaccination** afin de renforcer la réponse immunitaire aux agents microbiens (Gorden et Plummer, 2010). Des vaccins contre les virus IBR-1, PI-3, BRSV, BVD-1 et BVD-2 sont disponibles. Une vaccination contre les bactéries *Mannheimia haemolytica*, *Histophilus somni* et *Pasteurella multocida* peut être ajoutée selon l'historique du troupeau. Les protocoles de vaccination sont modulables selon les risques présents dans le troupeau ;
- L'administration d'**antibiotiques avant l'apparition de la BPI** (Stanton et al., 2010, Stanton et al., 2012, Teixeira et al., 2017a). En général, les antibiotiques sont donnés soit : (1) en **prophylaxie**, où l'antibiotique est administré à des veaux sains pour prévenir l'apparition de la BPI, soit, (2) en **métaphylaxie**, où l'antibiotique est administré à des veaux en santé et à des veaux diagnostiqués avec la BPI, tous

présents dans un même groupe ou bâtiment (Aarestrup, 2005). Chez les génisses de remplacement, l'administration d'antibiotique avant l'apparition de la BPI n'est pas une pratique courante. Quelques études rapportent son utilisation en prophylaxie lors des périodes à risques au début de la période de présevrage, après un transport ou encore au sevrage afin de diminuer l'impact du stress engendré et la pression infectieuse importante secondaire au rassemblement des veaux (Stanton et al., 2010, Stanton et al., 2013, Celestino et al., 2020). À noter que cette pratique n'est cependant pas un emploi raisonné et judicieux des antibiotiques et tend à ne plus être réaliser ;

- **Le diagnostic de la BPI, l'isolement quand cela est possible, et le traitement individuel** réalisé à partir d'un protocole élaboré par le médecin vétérinaire et appliqué par les producteurs (McGuirk, 2008).

I-2- Chez les veaux lourds

I-2-1- Le contexte d'élevage

L'élevage de veaux lourds correspond à l'engraissement de veaux provenant de fermes laitières. La majorité des veaux sont des mâles et sont abattus entre 6 et 8 mois d'âge (Sans et Fontguyon, 2009, Pardon et al., 2014). Selon l'alimentation reçue et l'aspect de la viande, on distingue : (1) l'élevage de veaux de lait où les veaux sont nourris essentiellement avec une alimentation lactée afin d'obtenir une viande de couleur pâle (Sargeant et al., 1994b) ; et (2) l'élevage de veaux de grain où les veaux sont tout d'abord nourris avec une alimentation lactée et aux grains (période de démarrage) suivis d'une période de finition où les veaux sont engraisés uniquement aux grains. La viande est de couleur rosée (Sargeant et al., 1994a).

La gestion d'élevage varie selon la localisation et le type d'élevage lait/grain (Pardon et al., 2014). L'élevage peut se réaliser dans la ferme laitière d'origine avec l'engraissement des veaux mâles nés sur la ferme et l'achat éventuel de veaux supplémentaires (Lava et al., 2016a), ou se dérouler dans des parcs d'engraissement où l'alimentation et le personnel technique sont fournis par des intégrateurs (Pardon et al., 2014, Jarrige et al., 2017, Renaud et al., 2018). Dans ce dernier type d'élevage, les veaux âgés d'environ 1 à 2 semaines (variable selon les réglementations spécifiques liées au transport des jeunes veaux) sont soit

directement transportés dans le parc d'engraissement, ou dans un encan où des veaux de différentes origines sont rassemblés avant d'être répartis dans différents parcs d'engraissement (Bahler et al., 2012, Leruste et al., 2012, Winder et al., 2016, Marquou et al., 2019).

En élevage, les veaux sont en général logés individuellement par mesure de biosécurité dans les 4 à 8 premières semaines afin de limiter la transmission de maladie (p. ex., diarrhée, BPI) (Miller et al., 1980, Callan et Garry, 2002, Pardon et al., 2014). Par la suite, les veaux sont rassemblés en groupe de tailles variables (Stull et McDonough, 1994, Bokkers et Koene, 2001) pour compléter leur période d'engraissement qui varie de 16 à 26 semaines (Sans et Fontguyon, 2009, Brscic et al., 2012, Leruste et al., 2012). La période de finition des veaux de grain peut se réaliser sur un site différent ou sur le même site (Fertner et al., 2016, Earley et al., 2019).

Le jeune âge des veaux, l'origine multiple, les transports et les rassemblements favorisent la survenue de la BPI chez les veaux lourds. Dans les parcs d'engraissement, l'installation de la maladie est lente et débute dès la première semaine après l'arrivée pour progressivement atteindre son incidence maximale entre 3.0 et 3.5% de veaux à risque (Pardon et al., 2012b), 3 à 5 semaines après l'arrivée (Rérat et al., 2012, Santman-Berends et al., 2018).

I-2-2- Les mesures de prévention et de contrôle de la bronchopneumonie infectieuse

Tout comme pour les génisses de remplacement, la période de présevrage constitue la période critique pour la survenue de la BPI en élevage de veaux lourds (Pardon et al., 2012b, Renaud et al., 2018). C'est pourquoi la majorité des mesures de prévention et de contrôle s'effectuent durant cette période, soit dans la ferme d'origine avant le transport, soit en parc d'engraissement à l'arrivée et durant les premières semaines. Ces mesures reposent sur :

- L'administration de **colostrum à la ferme d'origine** qui doit être de bonne qualité et en bonne quantité pour assurer un bon transfert d'immunité pour combattre la

- BPI et d'autres maladies (Stull et McDonough, 1994, Wilson et al., 2000, Pardon et al., 2015) ;
- La réception de veaux **en bonne santé à l'arrivée** ayant reçu une alimentation et des soins de qualité à la ferme d'origine. L'état du nombril (désiré sec et cicatrisé), l'absence de faiblesse lors du passage de la position couchée à debout, un bon état d'éveil et leurs capacités à se tenir solidement sur leurs pattes constituent des critères d'évaluation de l'état de santé des veaux à l'arrivée en parc d'engraissement (Brscic et al., 2012, Pardon et al., 2012b, Winder et al., 2016) ;
 - Des conditions de **transport** limitant le stress, à savoir de durée limitée et avec des veaux d'âge minimum respecté avant le départ de la ferme d'origine (Pardon et al., 2014, Renaud et al., 2018, Marcato et al., 2020). Cet âge minimum est en moyenne de deux semaines, mais reste variable selon les réglementations spécifiques des pays ;
 - **Un logement favorisant une bonne qualité de l'air** à l'intérieur du parc d'engraissement grâce à une ventilation efficace, au retrait fréquent du fumier et une densité d'animaux appropriée (Jones et Webster, 1984, Santman-Berends et al., 2018) ;
 - **Des contacts limités** entre les veaux les premières semaines de vie et entre les veaux de différents groupes d'âge afin de limiter la transmission d'agents microbiens et de mieux les surveiller (Miller et al., 1980, Callan et Garry, 2002) ;
 - **La vaccination** afin de renforcer la réponse immunitaire aux agents microbiens (Gorden et Plummer, 2010). Cette vaccination se réalise sur les veaux la première semaine après l'arrivée en parc d'engraissement, par voie intranasale contre IBR-1, PI-3, BRSV, +/- *Mannheimia haemolytica* et *Pasteurella multocida*. Plus tard, entre 35 et 45 jours de la période d'engraissement, les veaux peuvent recevoir un vaccin systémique contre IBR-1, PI-3, BRSV, BVD-1, BVD-2 +/- *Mannheimia haemolytica* et *Pasteurella multocida* (Woolums, 2015) ;
 - L'administration d'un traitement **antibiotique avant l'apparition de la BPI** (Pardon et al., 2012a, Rérat et al., 2012, Pardon et al., 2015). Ce traitement est généralement donné la première semaine après l'arrivée en parc d'engraissement,

- en métaphylaxie et vise principalement à contrôler les diarrhées présentes en démarrage (Bahler et al., 2012, Pardon et al., 2012a, Lava et al., 2016b) ;
- **Le diagnostic et le traitement individuel** réalisé à partir d'un protocole élaboré par les médecins vétérinaires et les techniciens et appliqué par les producteurs. Ce traitement vise indirectement à limiter la propagation des agents microbiens notamment bactériens (Bahler et al., 2012, Rérat et al., 2012).

I-3- Les limites d'efficacité de ces mesures de prévention et de contrôle chez les veaux de race laitière

Bien que présentes dans la majorité des élevages de génisses de remplacement et de veaux lourds *via*, par exemple, l'adhésion à des codes de pratiques, ces mesures peinent à réduire la morbidité et la mortalité secondaire au BPI qui restent stables depuis plusieurs années (Windeyer et al., 2017). Plusieurs arguments expliquent le manque d'efficacité de ces mesures de prévention et de contrôle de la BPI chez les veaux de race laitière :

- **L'origine multifactorielle de la BPI**, qui fait que si on veut prévenir et contrôler la BPI on ne peut pas se restreindre au contrôle strict d'un seul facteur (Windeyer et al., 2017). À titre d'exemple, un élevage de veaux lourds peut être constitué de veaux vaccinés et ayant reçu un bon transfert d'immunité, mais, si la ventilation dans le parc d'engraissement n'est pas adéquate, ces mesures seront insuffisantes pour prévenir la survenue de la BPI ;
- **Les variabilités dans la gestion des fermes** selon les producteurs, les pays, ou encore les intégrateurs en élevage de veaux lourds (Brscic et al., 2012, Windeyer et al., 2017). Ces variabilités font que certaines mesures seront mieux appliquées dans certaines fermes que d'autres, contribuant à la variabilité d'efficacité globale de ces mesures ;
- **Les variabilités de collaboration interprofessionnelle** entre les médecins vétérinaires/techniciens/producteurs pour la mise en place de ces mesures et les protocoles de détection de la BPI (Virtala et al., 1999, Olson et al., 2019, Sicho et al., 2019), et d'autre part, **intersectorielle** entre les élevages de génisses de remplacement fournissant les veaux des élevages de veaux lourds (Renaud et al., 2017, Santman-Berends et al., 2018, Marquou et al., 2019) ;

- **L'inefficacité de détection des veaux malades** avec des études récentes recensant une proportion importante de génisses de remplacement avec des lésions échographiques pulmonaires n'ayant jamais reçu de traitement dans le passé, synonyme d'un manque de détection (Buczinski et al., 2014, Ollivett et Buczinski, 2016, Cramer et Ollivett, 2019, Binversie et al., 2020). Cette détection suboptimale se retrouve également chez les veaux lourds lorsque l'on compare la forte proportion de lots avec des lésions pulmonaires à l'abattage (> 50 %) *versus* le faible nombre de veaux détectés cliniquement dans ce même lot (5 %) (Leruste et al., 2012).

II- Les conséquences actuelles de la bronchopneumonie infectieuse

Ces limites d'efficacité de prévention et contrôle de la BPI font que cette maladie reste prévalente avec, pour les élevages de génisses de remplacement, une prévalence qui varie de 0.12 à 0.23 selon les études (Staton et al., 2013 ; Urie et al., 2018 ; Dubrovsky et al., 2019). Pour les veaux lourds, la prévalence varie entre 0.14 à 0.61 selon les études (Brscic et al., 2012 ; Bahler et al., 2012 ; Pardon et al., 2015). Ces chiffres font de la BPI l'une des maladies les plus prévalentes avec la diarrhée néonatale chez les veaux de race laitière (Pardon et al. 2012 ; Windeyer et al., 2014). En plus d'être prévalente, la BPI engendre des conséquences importantes chez les veaux de race laitière sur : (1) l'économie de l'industrie, (2) la santé des veaux et (3) la santé publique. Revoyons brièvement ces impacts dans les prochains paragraphes.

II-1- Les impacts économiques de la bronchopneumonie infectieuse

II-1-1- Les impacts liés au problème de bronchopneumonie infectieuse

i- Les pertes visibles

Sortie du cycle de production (mortalité, euthanasie, réforme) : La BPI constitue une des premières causes de mortalité chez les génisses de remplacement (Sivula et al., 1996, Svensson et al., 2003, Urie et al., 2018b) et les veaux lourds (Sargeant et al., 1994b, a, Wilson et al., 2000, Lava et al., 2016a, Lava et al., 2016b). À titre d'exemple, le dernier rapport de santé du département de l'agriculture des États-Unis rapporte que 24 % des

mortalités des génisses de remplacements en période de présevrage sont secondaires à la BPI (APHIS, 2018). Dans les veaux lourds, la maladie est responsable de 16 à 50 % des mortalités durant la période d'engraissement (Lava et al., 2016a, Winder et al., 2016). Cependant, les risques de mortalité ne surviennent pas uniquement au moment de l'épisode de la maladie. En effet, une récente méta-analyse rapporte qu'une génisse atteinte de BPI durant sa première année de vie aurait presque 3 fois plus de chance de mourir avant son premier vêlage (OR = 2.85 ; IC95% : 1.22 ; 6.69) et environ 2 fois plus de chance de sortir du troupeau (OR = 2.30 ; IC95% : 1.75 ; 3.03) (Buczinski et al., 2021a). Dans le secteur de production de veaux lourds, une étude belge rapporte qu'un veau présentant un épisode de BPI a 5,5 fois plus de chance de mourir pendant le cycle d'engraissement (HR = 5.5 ; IC95% : 3.0 ; 10.2) (Pardon et al., 2013). Au-delà de la perte directe du veau lors de cette sortie du cycle de production, des paramètres rapportés en élevage de boucherie comme une perte liée à l'investissement dépensé dans la gestation de la mère, une perte liée à l'achat, et une perte liée à l'élevage du veau sont également à considérer chez les veaux lourds et les génisses de remplacement (Smith et al., 2020).

Perte de productivité (croissance, gain, production laitière) : Cette perte de productivité engendre les principales pertes économiques de l'industrie (Delabouglise et al., 2017). Ainsi, Buczinski et al. (2021a) rapportent dans leur méta-analyse qu'une génisse de remplacement atteinte de BPI la première année de vie va avoir un gain de poids moyen quotidien réduit en moyenne de 0,067 kg/j (IC95% : -0.099 ; -0.034) et une production laitière diminuée en moyenne de 121,2 kg (IC95% : -184.9 ; -57.5) durant la première lactation. Il est également important de mentionner qu'une génisse avec une lésion de consolidation pulmonaire de profondeur ≥ 3 cm à l'échographie thoracique durant la période de présevrage, avec ou sans symptômes de BPI, aura une production laitière réduite en moyenne de 525 kg (IC95% : -992.81 ; -60.25) à la première lactation (Dunn et al., 2018). Chez les veaux lourds, cette perte de productivité va se traduire par une réduction du poids de la carcasse de 4,3 kg à 9,7 kg (Van der Mei et Van den Ingh, 1987, Pardon et al., 2013) et une diminution de la qualité de la viande par une altération de la classification des carcasses (Pardon et al., 2013).

ii- Les pertes invisibles

Parmi les pertes invisibles, on répertorie dans le **Tableau 4** tous les impacts de la BPI sur **la reproduction** des génisses de remplacement.

Paramètre affecté	Définition de la bronchopneumonie infectieuse	Période du diagnostic	Résultats	Référence
Un premier vêlage retardé	Critère clinique (toux, jetage nasal ou oculaire, dyspnée).	≤ 90 jours de vie	BPI0 = 50% des génisses 32 mois vs BPI+ = 50% des génisses à 38 mois HR = 0.5 (IC95%: 0.3; 0.8) pour les BPO de ne pas avoir un premier vêlage retardé.	(Correa et al., 1988)
	Critère clinique (toux, jetage nasal ou oculaire, dyspnée).	≤ 90 jours de vie	BPI0 = 29 mois (médiane) vs BPI+ = 32 mois (médiane) HR = 0.5 (IC95%: 0.4; 0.8) pour les BPO de ne pas avoir un premier vêlage retardé.	(Warnick et al., 1994)
	Veau défini avec la maladie quand traité par le producteur.	Jusqu'à 650 jours d'âge	BPI0 = 724 jours (moyenne) vs BPI4 ou plus = 732 jours (moyenne) <i>Aucune donnée pour BPI 1,2 ou 3.</i>	(Bach et al., 2011)
	Veau défini avec la maladie quand traité par le producteur.	< 1 an de vie	BPI0 = 24.69 mois (SE : 0.07) vs BPI1 = 24.89 mois (SE : 0.07) vs BPI2 = 25.22 mois (SE : 0.11).	(Rossini, 2004)
	Critère clinique (abattement, diminution de l'appétit, manque d'énergie, température rectale > 39.5°C, jetage nasal ou dyspnée).	Suivi de 56 jours à 120 jours	BPI0 = 83% vèlées ≤ 25 mois vs BPI+ = 73% vèlées ≤ 25 mois HR de vêler avant 25 mois de 0.6 (IC95%: 0.4; 0.8) pour les BPO BPI0 = 702 jours (médiane) vs BPI+ = 714 jours (médiane).	(Stanton et al., 2012)
	Consolidation pulmonaire (≥ 1 cm) avec ou sans signe clinique de BPI.	À 60 jours au moment du sevrage	Absence de consolidation = 679.8 jours (SE : 1.4) vs présence de consolidation = 687.4 jours (SE : 2.0).	(Teixeira et al., 2017b)
Un succès d'être gestante diminué	Consolidation pulmonaire (≥ 1 cm) avec ou sans signe clinique de BPI.	À 60 jours au moment du sevrage	HR de devenir gestante de 0.7 (CI95%: 0.6; 0.8) pour les BPI0.	(Teixeira et al., 2017b)
Un risque de dystocie au premier vêlage	Critère clinique (abattement, diminution de l'appétit, manque d'énergie, température rectale > 39.5°C, jetage nasal ou dyspnée).	Suivi pendant 2 mois de 56 jours à 120 jours	BPI0 = 46% avec une dystocie BPI+ = 50% avec une dystocie HR = 1.5 (IC95%: 1.1; 2.2) pour BPI+ de présenter une dystocie.	(Stanton et al., 2012)
	Critère clinique (toux, jetage nasal ou oculaire, dyspnée).	≤ 90 jours de vie	HR de présenter une dystocie au premier vêlage de 2.4 (IC95%: 1.0; 5.7) pour BPI+.	(Warnick et al., 1994)

BPI0 = Absence de bronchopneumonie infectieuse; BPI+ = Présence de bronchopneumonie infectieuse; BPI1 = Présence de 1 épisode de bronchopneumonie infectieuse; BPI2 = Présence de 2 épisodes de bronchopneumonie infectieuse; BPI4 = Présence de 4 épisodes de bronchopneumonie infectieuse; IC95% = intervalle de confiance à 95%; HR = Hazard ratio qui correspond à la mesure de l'effet d'avoir une BPI sur un résultat au fils du temps; SE= Erreur standard.

Tableau 4. – Impacts de la bronchopneumonie infectieuse au début de vie sur la reproduction chez les génisses de remplacement.

II-1-2- Les impacts liés aux dépenses et réaction aux problèmes de bronchopneumonie infectieuse

Coût du traitement (antibiotiques, anti-inflammatoires) : Lorsqu'il est détecté, le veau va être traité. Pour avoir un ordre d'idée, Kaneene et Hurd (1990) rapportent qu'à court

terme, une génisse de remplacement traitée pour BPI coûte en plus par année approximativement 14.71\$ USD (soit environ 20\$ CAD), variant de 0-119\$ USD, qu'une génisse sans BPI. Considérant que la prévalence de la BPI est d'environ 0.20 chez les génisses de remplacement en période de présevrage (Urie et al., 2018b, Dubrovsky et al., 2019), les coûts engendrés peuvent très vite devenir importants. En élevage de veaux lourds, la perte économique engendrée par un veau qui connaît un épisode de BPI est estimée à 40 € (soit 57\$ CAD) (Pardon et al., 2013). Or, la BPI est la maladie la plus fréquemment traitée en élevage de veaux lourds ce qui laisse également présager une perte non négligeable pour cette industrie (Pardon et al., 2012a, Lava et al., 2016b, Jarrige et al., 2017). À ces dépenses vient se rajouter le coût des vaccins pour prévenir la BPI.

Le moral, la manœuvre et le temps de travail : Hormis ces dépenses directes liées aux problèmes de BPI, l'impact négatif sur le moral des employés, le recrutement de main-d'œuvre supplémentaire pour réaliser les traitements ou encore le temps supplémentaire de travail sont d'autres réactions liées à la gestion du problème de BPI décrit chez les bovins de boucherie, mais généralisable à l'industrie de veaux lourds et aux génisses de remplacement (Smith et al., 2020).

II-2- Les impacts sur la santé de la bronchopneumonie infectieuse

Le bien-être animal : l'absence de douleur, de lésions et de maladie fait partie d'un des principes fondamentaux du bien-être animal défini par OMSA (Organisation Mondiale de Santé Animale) (OMSA, 2019). Or, récemment une étude a objectivé qu'un veau atteint de BPI présente de la douleur (Martin et al., 2022). Ainsi, si un veau atteint de BPI n'est pas traité, le veau peut ressentir un certain inconfort, impactant son bien-être. La forte proportion de veaux (69 à 87 %) présentant des lésions échographiques de BPI mais n'ayant jamais reçu de traitement est donc actuellement problématique pour le bien-être des veaux (Ollivett et al., 2011, Buczinski et al., 2014). Une meilleure détection et un meilleur traitement de la BPI s'imposent afin de diminuer les répercussions sur le bien-être des veaux de race laitière.

Les infections chroniques de BPI : Comme c'est le cas en élevage de boucherie, en élevage de veaux lourds et chez les génisses de remplacement, il n'est pas rare d'être face à des infections chroniques plus difficiles à traiter en raison d'une diffusion moins efficace

de l'antibiotique et l'implication fréquente de bactéries plus difficile à tuer comme *Mycoplasma bovis*, *Trueperella pyogenes* ou encore des bactéries anaérobies. C'est pourquoi ces cas chroniques sont souvent sujets à des récives qui va avoir des impacts importants sur la santé, et sur tous les aspects économiques cités précédemment (pertes visibles, pertes invisibles, et liées aux dépenses et réactions au problème) (Bookers et al., 2020). De plus, ces cas vont nécessiter des traitements avec une grande quantité d'antibiotiques avec parfois l'emploi de différentes familles lors de récive, ce qui va poser une problématique sur l'utilisation raisonnée des antibiotiques (Bookers et al., 2020). Développons ce point dans le paragraphe suivant.

II-3- Les impacts sur la consommation d'antibiotiques de la bronchopneumonie infectieuse

II-3-1- L'utilisation d'antibiotiques

La BPI et la diarrhée néonatale constituent les maladies principales traitées avec des antibiotiques chez les veaux de race laitière (Pardon et al., 2012a, Pereyra et al., 2015). À titre d'exemple, dans les fermes laitières en Pennsylvanie, Sawant et al. (2005) rapportent que 40 % et 57 % des antibiotiques utilisés chez les génisses de remplacement sont dédiés à la BPI et la diarrhée néonatale, respectivement (Sawant et al., 2005). Cette utilisation pour la BPI est parfois doublée chez les veaux lourds avec entre 50 à 80 % des traitements antibiotiques destinés à cette maladie (Pardon et al., 2012a, Fertner et al., 2016, Lava et al., 2016b, Schnyder et al., 2019). La consommation d'antibiotiques en prévention (prophylaxie ou métaphylaxie) constitue la principale source d'utilisation en élevage de veaux lourds (Pardon et al., 2012a, Lava et al., 2016b).

II-3-2- La contribution au développement de résistance

L'Organisation mondiale de la santé définit une résistance aux antibiotiques lorsqu'une bactérie acquiert au cours du temps la faculté de survivre à l'exposition d'antibiotiques (WHO, 2021). La transmission de bactéries résistantes de l'animal à l'homme est un problème de santé publique majeur (WHO, 2021). La surveillance de la résistance et l'emploi judicieux d'antibiotiques sont donc primordiaux chez les veaux de race laitière.

Cette résistance peut se mesurer de différentes manières (Schrenzel, 2018). Voici les deux méthodes les plus couramment utilisées:

- **Par détermination de la susceptibilité de la bactérie à l'antibiotique** soit (1) par diffusion sur disque qui renseigne une zone d'inhibition, ou soit (2) par microdilution qui permet d'obtenir une concentration minimale inhibitrice (**CMI**). Les valeurs obtenues par ces méthodes sont ensuite comparées à des seuils cliniques ou épidémiologiques prédéfinis, permettant de juger de la résistance de la bactérie ;
- **Par détermination des gènes de résistances** par PCR ou séquençage plus précise et exacte que la méthode précédente pour déterminer une résistance *in vivo*.

Ces dernières années, ces méthodes ont été utilisées pour mettre en évidence de la résistance aux antibiotiques chez les veaux de race laitière. Une résistance à plusieurs antibiotiques a été mise en évidence chez des bactéries de la flore commensale respiratoire par microdilution (Catry et al., 2005, Di Labio et al., 2007, Rérat et al., 2012), des bactéries digestives par disque de diffusion (Catry et al., 2007, Catry et al., 2016) et par les méthodes génotypiques pour *Escherichia Coli* (Massé et al., 2021), mais plus inquiétant, chez des bactéries zoonotiques (p. ex., Salmonelle ; *Staphylococcus aureus* résistant à la méthicilline) par des méthodes génotypiques (Graveland et al., 2010, Cook et al., 2011, Vandendriessche et al., 2013).

L'utilisation importante d'antibiotiques pour la BPI (Berge et al., 2006, Catry et al., 2016), leur administration par voie orale à des doses sous thérapeutiques lors de traitement en métaphylaxie rapportée chez les veaux lourds (Roberts et al., 2008, Pardon et al., 2012a, Bos et al., 2013, Burow et al., 2014, Schönecker et al., 2019), ou encore l'utilisation de combinaison d'antibiotiques ou de familles différentes lors de récurrence (Pardon et al., 2012a, Coetzee et al., 2019) sont toutes des pratiques réalisées chez les veaux de race laitière susceptibles de favoriser l'émergence de telles résistances. Si on veut diminuer l'émergence de résistance, il est primordial de raisonner l'utilisation des antibiotiques en (1) diminuant la quantité d'antibiotiques administrée notamment l'administration de groupe ; (2) en traitant uniquement les veaux qui en ont besoin ; (3) et en améliorant la durée et l'efficacité du traitement pour limiter les récurrences (Hubbuch et al., 2021).

III- L'intérêt d'améliorer la détection de la bronchopneumonie infectieuse

Comme le souligne le paragraphe précédent recensant les conséquences actuelles engendrées par la BPI, l'efficacité des mesures de prévention et de contrôle de la maladie est loin d'être optimale de nos jours. Une des stratégies pour diminuer ces conséquences serait d'améliorer la détection de la maladie en parallèle des mesures préventives instaurées dans chaque ferme, afin de traiter exclusivement, précocement et plus efficacement les veaux malades. Plusieurs effets sont attendus en améliorant la détection de la BPI :

III-1- Réduction des impacts économique et sur la santé de la bronchopneumonie infectieuse

L'absence de détection d'un veau atteint de BPI implique qu'il ne va pas recevoir de traitement, laissant alors la maladie évoluer soit : (1) en gravité et engendrer sa mort (perte directe) ; (2) guérir en laissant parfois des séquelles dans le poumon rendant le veau moins performant à court et long terme et plus sensible à une recrudescence de la maladie dans le futur ; (3) persister dans le temps sous forme d'infection chronique, rendant le veau encore moins performant, difficile à traiter avec un risque important de rechute et souvent un dépérissement après plusieurs traitements.

Ainsi, en améliorant la sensibilité (**Se**) de détection précoce de la BPI, c.-à-d., la faculté de détection des veaux malades, les veaux pourront être traités précocement contre la BPI et ainsi stopper plus efficacement l'évolution négative de la maladie. Les conséquences économiques et sur la santé de la BPI seront alors réduites.

III-2- Réduction de la consommation d'antibiotique pour la bronchopneumonie infectieuse

La mauvaise détection des veaux atteints de BPI engendre l'instauration d'un traitement antibiotique pour des veaux diagnostiqués faussement malades, engendrant une surutilisation d'antibiotique et contribuant à l'émergence de résistance. L'amélioration de la spécificité (**Sp**) de détection, c.-à-d. ; la faculté à détecter les veaux non malades permettra de traiter exclusivement les veaux malades et ainsi réduire la consommation d'antibiotiques. Chez les bovins de boucherie, l'administration d'antibiotique à l'échelle d'un groupe en métaphylaxie a un intérêt uniquement dans les groupes où la prévalence de BPI est ≥ 0.15 (Baptiste and Kyvsgaard, 2017). Ce point se retrouve chez les veaux lourds

avec une étude montrant que le traitement de groupe n'était pas bénéfique dans un lot de faible prévalence (0.05) (Berman et al., 2017). L'amélioration de la détection des lots au-dessus d'un certain seuil de prévalence de BPI (p. ex., ≥ 0.15 comme rapporté dans les bovins de boucherie [Baptiste and Kyvsgaard, 2017]) permettrait de justifier l'administration d'antibiotiques de groupe et ainsi réduire leur consommation.

Conclusion PARTIE II : Malgré des années de recherche, la BPI reste prévalente en élevage de génisses de remplacement et de veaux lourds. Les conséquences négatives de la maladie sont multiples, que ce soit sur l'économie, sur la santé des veaux ou encore sur la consommation d'antibiotiques. Ces impacts importants imposent une révision des mesures de contrôle actuelles de la BPI chez les veaux de race laitière. L'amélioration des stratégies diagnostiques pour détecter les veaux à traiter avec des antibiotiques permettrait de mieux estimer les réelles conséquences de la BPI sur l'industrie, mieux mesurer l'effet des facteurs de risque et d'intervention sur la maladie, et, ainsi, permettrait de diminuer les impacts de la BPI.

PARTIE III : Les tests diagnostiques actuels pour détecter la bronchopneumonie infectieuse chez les veaux de race laitière

L'amélioration de la détection de la BPI passe par l'amélioration des tests diagnostiques réalisables en *ante-mortem* chez un veau de race laitière. On distingue deux catégories de tests : (1) les principaux tests permettant de mettre en évidence l'agent microbien responsable de la BPI (**Tableau 3**) ; (2) les principaux tests permettant de mettre en évidence les lésions, les symptômes et les conséquences de la BPI (**Tableau 5**).

Test		Mesure
Score clinique respiratoire diagnostique		Signes cliniques
Observation et interprétation des signes cliniques		Signes cliniques
Auscultation thoracique		Bruits respiratoires
Tests d'imagerie médicale : échographie thoracique, radiographie thoracique, tomodensitométrie		Lésions pulmonaires
Thermographie infra-rouge		Température
Acéléromètre, données automatisées de la prise alimentaire		Comportement
Analyse sanguine	Hématologie, dosage des protéines de l'inflammation aiguës (ex : Haptoglobine)	Inflammation pulmonaire
	Gaz artériels et veineux	Échange gazeux

Tableau 5. – Principaux tests diagnostiques disponibles pour mettre en évidence les symptômes, les lésions et les conséquences de la bronchopneumonie infectieuse chez les veaux de race laitière.

Dans cette thèse, j'ai fait le choix d'améliorer les stratégies diagnostiques de la BPI à partir de quatre tests : le score clinique respiratoire diagnostique (**SCRD**), l'auscultation thoracique, l'échographie thoracique et la radiographie thoracique. Dans cette partie III, nous reverrons les avantages et inconvénients actuels de chaque technique, de l'interprétation et de l'accessibilité de chacun de ces tests avant de développer plus en détail leurs performances de détection actuelles dans la partie IV.

I- Les scores cliniques respiratoires diagnostiques

Par définition, un score clinique diagnostique est un outil d'aide à la décision médicale consistant à pondérer des signes pertinents (= prédicteurs) pour les agréger en une seule valeur numérique (= score total), représentative de la probabilité de la présence de la maladie (Hayes et al., 2010, Moons et al., 2015).

I-1- Les différents scores cliniques respiratoires diagnostiques rapportés chez les veaux de race laitière

Chez les veaux de race laitière, 5 principaux SCRD ont été décrits pour diagnostiquer spécifiquement la BPI : (1) le score CIS (*Clinical Illness System*) adapté chez les veaux de race laitière à partir du score décrit par Perino et Apley (1998) chez les bouvillons de boucherie (Amrine et al., 2013) ; (2) le score du Wisconsin (McGuirk, 2008) ; (3) le score Californien pour les veaux présevrés (Love et al., 2014) ; (4) le score Californien — Québec, dont les poids du score Californien ont été recalculés avec une technique statistique différente, dans une population différente (Buczinski et al., 2018c) ; et (5) le score Californien chez les veaux sevrés élevés en groupe (qu'on appellera score de Maier dans cette thèse) (Maier et al., 2019). La description plus détaillée de chacun de ces scores est présentée dans le **Tableau 6**. À noter que certains d'entre eux ont été illustrés sous forme de diagramme (Wisconsin, **Annexe 2** ; Californien, **Annexe 4** ; et Maier, **Annexe 5**). Des applications mobiles sont également disponibles pour faciliter leur utilisation. Tous ces scores incluent des prédicteurs individuels (p. ex. ; jetage nasal, jetage oculaire, etc.) à l'exception du score de Maier qui inclut également des prédicteurs plus généraux, comme le changement de température de l'environnement dans la journée précédant l'évaluation. Ces SCRD ont été élaborés pour les producteurs afin qu'ils puissent identifier les veaux malades et initier le traitement.

I-2- L'interprétation des scores cliniques respiratoires diagnostiques

Ces SCRD sont interprétés à partir d'un seuil décisionnel définissant un veau comme positif ou négatif. Deux paramètres sont à prendre en compte pour juger de la robustesse de ce seuil et de son application dans une population : (1) sa méthode de détermination qui peut être arbitraire *versus* statistiquement calculée pour : soit optimiser la performance (p. ex., Se + Sp optimal sur une courbe de ROC [Receiver Operating Characteristic]) (Šimundić, 2009) ou soit diminuer les coûts d'un mauvais diagnostic du score (p. ex., calcul du MCT [Misclassification Cost Term]) (Greiner, 1996) (voir **Partie IV, II-1-2** pour plus de détails) ; (2) sa flexibilité : variation selon l'ajout ou non d'un prédicteur (p. ex. ; la température rectale) (Maier et al., 2019) ou la prévalence de la maladie (Buczinski et al., 2018c).

Score clinique respiratoire diagnostique	Population de développement du score	Prédicteurs	Évaluation des prédicteurs	Poids ¹	Interprétation Du score total
SCORE CIS (Amrine et al., 2013)	Veaux laitiers sevrés expérimentalement inoculés avec <i>Mycoplasma bovis</i>	4 prédicteurs Comportement Dépression Toux Respiration anormale	1= comportement habituel 2= légèrement malade (léger signe de dépression ou toux) 3= modérément malade (gravement déprimé, respiration anormale ou toux) 4= moribond	-	Seuil arbitraire >1 (Amrine et al., 2013)
SCORE DU WISCONSIN (McGuirk, 2008)	Veaux laitiers présevrés logés en groupe	5 prédicteurs Température rectale Jetage nasal Jetage oculaire Toux Position des oreilles	Gravité (0-3)	♦ x 1 pour tous les prédicteurs	Seuil arbitraire Deux seuils décrits : ≥ 6 (Lago et al., 2006) ≥ 5 (McGuirk, 2008)
SCORE CALIFORNIEN (Love et al., 2014)	Veaux laitiers présevrés logés individuellement	6 prédicteurs Les 5 même prédicteurs que pour le score du Wisconsin + l'ajout de la respiration anormale aux prédicteurs inclus dans le score du Wisconsin	Présence ou absence	♦ x 2 Jetage oculaire, toux, respiration anormale, température rectale (seuil 39.2°C) ♦ x 4 Jetage nasal ♦ x 5 Position des oreilles	Seuil statistiquement calculé pour optimiser les performances Seuil ≥5
SCORE CALIFORNIEN QUÉBEC (Buczinski et al., 2018c)	Veaux laitiers présevrés logés individuellement ou en groupe	6 prédicteurs Les 6 mêmes prédicteurs que le score Californien	Présence ou absence	♦ x 20 : respiration anormale ♦ x 16 : position des oreilles, toux ♦ x 10 : Jetage nasal ♦ x 7 : Température rectale ♦ x-1 : Écoulement oculaire	Seuil statistiquement calculé pour optimiser les performances et variable selon la prévalence dans un groupe Seuil variable de 9 à 16 Plus la prévalence est élevée, plus le seuil est revu à la baisse
SCORE DE MAIER (Maier et al., 2019)	Veaux sevrés logés en groupe	6 Prédicteurs Toux Respiration anormale Note d'état corporel Yeux enfoncés Température rectale Température ambiante (24h)	Présence ou absence	♦ x 4 Yeux enfoncés ♦ x 5 Note d'état corporel faible ♦ x 2 Toux ♦ x 1 Respiration anormale, fluctuation de température diurne ♦ x 1 Température rectale	Seuil statistiquement calculé pour optimiser les performances ≥2 si inclusion de la température dans le score Seuil après l'ajout d'un prédicteur ≥1 sans la température

¹ dans cette colonne, chaque poids est indiqué par le signe multiplicateur (x) signifiant qu'il faut multiplier le nombre indiqué par 1 si le signe est présent ou par 0 si absent

Tableau 6. – Principaux scores cliniques respiratoires diagnostiques rapportés pour diagnostiquer la bronchopneumonie infectieuse chez les veaux de race laitière.

Actuellement, les seuils optimaux des SCRD pour définir un veau malade *versus* non malade sont discutables et varient parfois selon les publications (score du Wisconsin ≥ 6 (Lago et al. 2006) *versus* ≥ 5 (McGuirk et al. 2008) et les populations (Score Californien *versus* Score Californien Québec) (**Tableau 6**).

I-3- Les scores cliniques respiratoires diagnostiques *versus* l'observation et l'interprétation des signes cliniques

I-3-1- Variation d'expression et de détection des signes cliniques de la bronchopneumonie infectieuse

L'expression clinique de la BPI est variable pour chaque veau atteint (Apley, 2014) :

- **Dans le temps** en fonction du moment de l'infection (p. ex. ; la fièvre qui précède les signes cliniques) (Timsit et al., 2011a, Ollivett et al., 2013), ou au moment de l'évaluation (p. ex. ; le jetage nasal que le veau peut facilement masquer en se léchant) (McGuirk et Peek, 2014) ;
- **Parmi les individus**, les bovins domestiques ont conservé un comportement de proie propre aux herbivores faisant en sorte que certains individus ont tendance à masquer leurs signes cliniques (Weary et al., 2009). De plus, les mécanismes de défense respiratoire étant variables d'un individu à l'autre, l'expression clinique peut donc s'avérer différente (Caswell, 2014, Bassel et al., 2020) ;
- **Selon la gravité** de l'infection : plus l'infection est grave, plus les mécanismes de défense seront stimulés et le dysfonctionnement du poumon sera apparent. L'expression clinique sera alors plus évocatrice (Yates, 1982) ;
- **Selon l'agent étiologique** : viral *versus* bactérien (Baruch et al., 2019) ou encore entre les différents agents (De Oliveira et al., 2016). Cependant, même si des signes cliniques peuvent être évocateurs de certains agents (p. ex. ; *Mycoplasma bovis* et la position anormale des oreilles), étant donné que les signes cliniques manquent de spécificité, cela n'implique pas nécessairement que la maladie est secondaire à cet agent. Pour illustration, Stanton et al. (2013) ont détecté plus de veaux *Mycoplasma bovis* négatifs avec une position anormale des oreilles que de veaux *Mycoplasma bovis* positifs. La position anormale des oreilles peut en effet être causée par de la douleur ou de l'abattement ;

En plus de leur expression, la détection des signes cliniques de la BPI a également été rapportée comme variable **selon les opérateurs** notamment entre les médecins vétérinaires et les producteurs (Virtala et al., 1996, Svensson et al., 2003). Cette différence peut s'expliquer par le fait que les producteurs n'ont pas forcément une qualification spécifique pour détecter les maladies contrairement aux médecins vétérinaires. Cependant, le diagnostic de BPI est majoritairement réalisé par eux sans aucune formation au préalable, ce qui peut expliquer la mauvaise détection actuelle.

Tous ces facteurs de variabilité d'expression et de détection rendent la détection clinique de la BPI inconstante et subjective, si bien que certains auteurs assimilent l'observation et l'interprétation des signes cliniques à un art, dont la faculté va dépendre des talents et de l'expérience de l'examineur (Portillo, 2014).

Plusieurs études ont souligné les limites diagnostiques de l'observation et l'interprétation des signes cliniques dans différents élevages : (1) chez les bovins de boucherie, une méta-analyse rapporte une faculté sous-optimale à détecter les veaux atteints de BPI (sensibilité ($Se = 0.27$ (ICB95% : 0.12 ; 0.65)) basé uniquement sur l'observation et l'interprétation des signes cliniques (Timsit et al., 2016) ; (2) en élevage de génisses de remplacement, cette limite diagnostique de l'observation et d'interprétation des signes cliniques se traduit par la faible proportion de veaux avec des lésions dans le poumon ayant été détectées et traités dans le passé (31 % pour l'étude de Buczinski et al. (2014) et 13 % pour celle de Ollivett et al. (2011)) ; (3) En élevage de veaux lourds, on constate la même inefficacité de détection des signes cliniques pour détecter des lots de veaux lourds de fortes prévalences de BPI avec une différence importante entre la proportion de signes cliniques de BPI observés dans des lots (<10 %) *versus* la proportion de lésions de BPI de ces lots à l'abattage (50 %) (Leruste et al., 2012).

I-3-2- L'objectivité des scores cliniques respiratoires diagnostiques par rapport à l'observation et l'interprétation des signes cliniques

À l'inverse de l'observation et l'interprétation des signes cliniques qui peut s'avérer subjective, l'analyse systématique d'un nombre limité de prédicteurs apporte plusieurs avantages dans la détection d'une maladie (Hayes et al., 2010, Moons et al., 2015) :

- L'évaluation par les utilisateurs est plus **standardisée**, ce qui diminue la subjectivité liée à l'expérience et au talent de l'examineur ;
- Certains signes cliniques sont **activement évalués** plutôt qu'observés, ce qui peut permettre de détecter des signes plus précocement (p. ex. ; toux induite *versus* toux simplement notée lors de l'observation passive) ;
- L'apport de **mesures objectives qui peuvent être comparées dans le temps**.

I-4- L'utilisation des scores cliniques respiratoires diagnostiques chez les veaux de race laitière

Par son objectivité apparente, le SCRD est donc un test avantageux à utiliser sur un grand nombre de veaux de race laitière. Plus concrètement, ces avantages ont rendu populaire l'utilisation des SCRD en recherche, où l'utilisation d'un tel score permet de récolter des données plus précises et objectives (Hayes et al., 2010), ou encore de cibler une population en particulier (p. ex. ; une population sous-clinique ayant des lésions thoraciques sans symptôme apparent dans l'étude de Ollivett et al. (2015)). En pratique, les SCRD ont été développés pour les producteurs afin qu'ils détectent et traitent les animaux malades. Ces SCRD facilitent l'implantation de protocole de dépistage et de traitement à la ferme. À titre d'exemple, McGuirk et Peek (2014) préconisent sans calcul statistique de réaliser le score du Wisconsin sur un échantillon aléatoire de 50 veaux (tous les veaux si $n < 20$), deux fois par semaine et quotidiennement sur les veaux à risques présentant de l'abattement, de l'anorexie, de la toux, etc.. Si le score du Wisconsin est ≥ 5 , le veau est traité avec des antibiotiques qui sont choisis en fonction des agents microbiens prévalents dans l'élevage. Finalement, d'un point de vue préventif, cet outil utilisé seul (McGuirk, 2008) ou avec l'échographie thoracique (Ollivett et Buczinski, 2016), permet d'évaluer et d'améliorer les facultés de détection du producteur lors de visites préventives en mesurant le taux de détection (=nombre de veaux traités/nombres de veaux malades avec le score +/- l'échographie thoracique). En conclusion, le SCRD est un test parfaitement adapté aux conditions d'élevage actuelles en étant simple, rapide, peu coûteux et facilement utilisable par les producteurs. Les seuils d'interprétation restent cependant à mieux définir.

II- L'auscultation pulmonaire

L'auscultation pulmonaire est la science et l'art d'écouter et d'interpréter les sons provenant des poumons et des voies respiratoires (Curtis et al., 1986). Cet examen nécessite un stéthoscope qui permet la transmission des sons de la cavité thoracique aux oreilles. L'auscultation thoracique est un test simple et rapide demandant peu de matériel. Cet examen est réalisé depuis des années systématiquement lors d'un examen physique par les médecins vétérinaires (Curtis et al., 1986).

II-1- La technique d'auscultation pulmonaire

Pour réaliser l'auscultation pulmonaire d'un veau de race laitière, le stéthoscope est posé sur la totalité de l'aire pulmonaire des deux côtés (**Figure 1**). En raison de la propriété du son à s'irradier, la localisation des bruits respiratoires anormaux détectés avec le stéthoscope ne correspond pas forcément à la localisation de la lésion (Scott, 2013, Buczinski et al., 2014). Des bruits respiratoires anormaux sont cependant rarement audibles dans les portions dorsales du poumon en raison de la localisation préférentielle des lésions de BPI cranioventralement (voir **Partie I, I-2-3**). La zone d'auscultation peut alors être limitée aux portions moyenne et ventrale du poumon lors de l'auscultation d'un grand nombre de veaux (Buczinski et al., 2014). Pour améliorer l'amplitude respiratoire, un test au sac peut également être réalisé à l'aide d'un gant de fouille positionné sur le nez du veau (**Annexe 6**). En respirant à travers le gant, la quantité d'air inspiré est restreinte, favorisant une augmentation de l'amplitude respiratoire. L'augmentation de l'amplitude permet alors d'accentuer les turbulences du flux d'air et faciliter l'auscultation de bruits anormaux lors de maladie respiratoire (Wilson et al., 2015).

II-2- L'interprétation de l'auscultation pulmonaire

Le **Tableau 7** définit les différents sons détectables lors de l'auscultation pulmonaire. Les sons rencontrés lors de BPI sont : l'augmentation des bruits bronchiques, les crépitants, les sifflements, le frottement pleural ou la diminution des bruits respiratoires.

Son audible	Définition	Signification pathologique
Bruits respiratoires normaux	Vélocité normale du flux des gaz respiratoires.	Physiologique
Bruits respiratoires augmentés	Augmentation de la vélocité normale du flux des gaz respiratoires. Cette augmentation est secondaire à une augmentation de la fréquence respiratoire et de la profondeur de la respiration.	♦ Physiologique (exercice intense, excitation, température environnementale élevée) ♦ Pathologique (fièvre, acidose, début d'œdème pulmonaire)
Augmentation des bruits bronchiques	Sons audibles dans l'aire pulmonaire ressemblant à des bruits trachéaux. L'augmentation des bruits bronchiques survient lorsque les bronches sont entourées de parenchyme pulmonaire consolidé favorisant l'amplification du son.	Pathologiques (bronchopneumonie infectieuse avec des lésions de consolidation pulmonaire sans bronchogramme (c.à.d. ; bronches non remplies de liquides))
Crépitants	Sons de cliquetis ou d'éclats. Deux phénomènes peuvent engendrer ce type de bruits : (1) l'éclatement de bulles de liquide dans les voies respiratoires, ou (2) les voies respiratoires qui restent fermées pour une portion de l'inspiration et qui s'ouvrent soudainement.	Pathologiques (œdème pulmonaire, pneumonie interstitielle)
Sifflements	Vibration des parois des voies respiratoires ou air qui passe à travers des voies respiratoires rétrécies.	Pathologiques (pneumonie allergique, bronchopneumonie infectieuse, pneumonie interstitielle)
Frottement pleural	Son similaire à l'utilisation de papier de verre causé par une inflammation de la cavité pleurale et le frottement de la plèvre viscérale et pariétale.	Pathologiques (pleurobronchopneumonie)
Bruits respiratoires diminués	Diminution de la vélocité normale du flux des gaz respiratoires. Cette diminution peut être secondaire à : (1) une obstruction de la transmission des bruits pulmonaires (accumulation de fluides (effusion pleurale), air (pneumothorax), pus (abcès pleural ou pyothorax), obésité); (2) lors de consolidation pulmonaire sévère avec bronchogramme (c.à.d. ; bronches remplies de liquides); (3) lors de respiration superficielle (douleur, faiblesse, méningoencéphalite).	♦ Physiologique (obésité) ♦ Pathologique (bronchopneumonie infectieuse, pleurobronchopneumonie, douleur, faiblesse, méningoencéphalite)
Bruits divers dans la zone thoracique	Grognements respiratoires, bruits cardiaques, bruits ruminiaux, bruits de frottement cutané.	

Tableau 7. – Sons audibles lors de l'auscultation thoracique chez les veaux de race laitière, d'après Curtis et al. (1986).

Contrairement au SCRD dont l'interprétation est définie par un seuil, la définition d'un test normal/anormal à l'auscultation thoracique n'est pas clairement définie. C'est pourquoi Curtis et al. (1986) qualifient l'interprétation comme un art sujet à une importante subjectivité basé sur le talent et l'expérience de l'examineur. Bien que des scores d'auscultation aient été décrits pour en augmenter l'objectivité d'interprétation chez les veaux laitiers (Verhoeff et al., 1985) et les bouvillons de boucherie (DeDonder, 2008), ces scores sont rarement utilisés en pratique par manque de popularité et d'évidence d'efficacité.

Chez les bouvillons de boucherie, un stéthoscope électronique (*Whisper*®) a récemment été développé afin d'être plus objectif pour interpréter l'auscultation thoracique (Noffsinger et al., 2014, Mang et al., 2015, Zeineldin et al., 2016). Brièvement, le stéthoscope électronique est positionné pendant 8 secondes au même endroit, entre le 7^e et le 11^e EIC à droite, dorsalement à la base du cœur. Le stéthoscope est capable d'enregistrer les sons et de valider la qualité de l'enregistrement. Puis un algorithme commercial permet de donner directement un score de gravité de la BPI (1 — absence de BPI; 2 — BPI légère; 3 — BPI modérée; 4 — BPI aiguë; 5 — BPI chronique) (<https://www.youtube.com/watch?v=a76g4dYTXck>). Les performances de cet outil sont prometteuses chez les bouvillons de boucherie (Se de 0.92 [ICB95 % = 0.71 ; 0.99] et Sp de 0.90 [ICB95 % = 0.64 ; 0.99]) (Mang et al., 2015). Cet outil n'a cependant pas été validé chez les veaux de race laitière avec la plupart des études le concernant qui ont été réalisées sur des animaux de plus de 250 kg de poids vifs (veaux de race de boucherie entrant en parc d'engraissement).

III- Les tests d'imagerie médicale

Les tests d'imagerie médicale permettent de visualiser les lésions pulmonaires associées aux BPI. Parmi eux, la tomodesitométrie (**CT scan**) est le test le moins sujet à des erreurs d'interprétation par sa faculté à examiner l'ensemble du poumon sans superposition anatomique. C'est pourquoi il est souvent utilisé comme test de référence pour visualiser les lésions de BPI dans les animaux de compagnies ou en médecine humaine (Schwarz et Tidwell, 1999, Bourcier et al., 2014, Chavez et al., 2014). Bien que le test soit utilisé chez les veaux de race laitière (Lubbers et al., 2007, Ohlerth et al., 2014, Fowler et al., 2017), son utilisation est limitée en raison de son prix élevé et de son accessibilité restreinte au milieu hospitalier. En pratique, deux autres tests d'imagerie médicale sont utilisés chez les veaux de race laitière : l'échographie thoracique, accessible à la ferme, et la radiographie thoracique, en milieu hospitalier. Par rapport au CT scan, l'échographie thoracique a l'inconvénient de ne pas évaluer les lésions en profondeur du parenchyme pulmonaire (Babkine et Blond, 2009) et la radiographie thoracique est sujette à des superpositions anatomiques (Farrow, 1999). C'est pourquoi ces deux tests sont moins exacts que le CT scan. Cependant, ils sont moins coûteux et plus familiers pour les médecins vétérinaires

bovins, c'est pourquoi nous nous attarderons essentiellement sur ces deux tests dans les prochains paragraphes.

III-1- L'échographie thoracique

III-1-1- La technique d'échographie thoracique

i- Première description de la technique

Rabeling et al. (1998) décrit pour la première fois, la technique d'échographie thoracique chez des veaux sevrés de race laitière. Cette technique, similaire à ce qui est réalisé chez les adultes (Reef et al., 1991), consistait initialement à raser bilatéralement les hémithorax droit et gauche de la 13^e côte jusqu'au bord caudal de la scapula. Après l'application de gel, une sonde sectorielle 7,5 MHz était appliquée dans chaque EIC, du haut vers le bas. La limite dorsale correspondait aux processus transverses et la limite ventrale à une ligne diagonale du coude jusqu'à la 13^e côte (Rabeling et al., 1998). Cette même technique a ensuite été décrite dans des études impliquant des veaux en début de sevrage (âgé de 2 à 3 mois) (Reinhold et al., 2002) ou nouveau-nés (Jung et Bostedt, 2004).

ii- Simplification de la technique

Depuis cette première description, la technique a été simplifiée afin d'optimiser :

- 1- **Sa rapidité et son utilisation sur plusieurs veaux.** En effet, une version simplifiée de la technique du 8^e jusqu'au 3^e EIC bilatéralement a été proposée pour gagner en rapidité lors de l'évaluation d'un nombre important de veaux (Buczinski et al., 2014). Plus récemment, une étude chez des veaux de race laitière a montré un accord substantiel (kappa pondéré de 0.78) entre l'échographie limitée au 5^e et 4^e EIC bilatéralement et l'échographie complète pour détecter les lésions de consolidation pulmonaire (Pravettoni et al., 2021). Chez les veaux lourds, l'échographie thoracique du 1^{er} au 5^e EIC à droite et du 2^e au 5^e EIC à gauche a montré un accord presque parfait (kappa de 0.99) avec l'échographie thoracique complète (Masset et al., 2022). En revanche, la simplification de la technique à un seul côté ne permet pas de détecter adéquatement les lésions de consolidation pulmonaire (Buczinski et al., 2014).

2- **Sa facilité d'utilisation par un grand nombre de médecins vétérinaires.** La technique a, en effet, été ajustée à des conditions de terrain en remplaçant le rasage/gel par l'application d'alcool (Ollivett et al., 2011, Ollivett et al., 2013) ou d'huile végétale (Pravettoni et al., 2021). De plus, l'utilisation d'une sonde linéaire transrectale, accessible pour un grand nombre de médecins vétérinaires, a montré une qualité d'image satisfaisante (Ollivett et al., 2011, Ollivett et al., 2013). Finalement, Ollivett et al. (2011) a montré que des lésions échographiques étaient facilement détectables lorsque le veau était debout facilitant son utilisation sur des veaux avec un minimum de contention (**Annexe 7**).

Ainsi, les avancées actuelles ont rendu l'échographie thoracique pratique à utiliser en élevage de génisses de remplacement (Ollivett et Buczinski, 2016) et de veaux lourds (Masset et al., 2022). En effet, actuellement, l'échographie thoracique est : (1) **rapide**, allant d'un maximum de 4 min pour les opérateurs expérimentés (Ollivett et al., 2011) à 9 min pour les moins expérimentés (Buczinski et al., 2013) ; (2) **non invasive**, en se réalisant sur des veaux debout, sans sédation (Ollivett et al., 2011, Ollivett et Buczinski, 2016) ; et finalement, (3) **peu coûteuse** en matériel, en utilisant une sonde transrectale et une machine déjà accessible par la majorité des médecins vétérinaires bovins pour la reproduction (Ollivett et Buczinski, 2016).

iii- Les ajouts à la technique initiale

Comme mentionné précédemment, la technique décrite chez le veau a été adaptée de celle décrite initialement chez les adultes (Reef et al., 1991). Cependant, contrairement aux bovins adultes, la taille des veaux présevrés nous permet d'avoir accès plus facilement aux fenêtres échographiques des lobes crâniens des poumons situés entre 1-3^e EIC en avant du cœur en passant notamment la sonde échographique sous le membre antérieur (**Figure 1**). C'est pourquoi certains auteurs ont proposé l'ajout des sites crâniens dans la technique d'échographie thoracique, nécessitant de passer la sonde sous le membre thoracique (**Annexe 7**) (Ollivett et al., 2011; Pardon, 2019a). L'échographie du lobe crânial gauche est cependant controversée car certains auteurs ne recommandent pas son évaluation, le considérant rarement infecté et plus difficile à évaluer en raison de la présence du thymus (Ollivett et Buczinski, 2016). L'inconvénient de l'ajout des lobes crâniens est que leur

examen demande un savoir-faire supplémentaire (passage de la sonde sous la patte et détection anatomique des repères permettant de détecter les images échographiques du lobe) sans savoir si son ajout améliore les performances de la technique. En effet, même si des études montrent des lésions fréquentes du lobe crânial droit (Ollivett et al., 2015, Tejero et Bach, 2016), aucune étude ne permet actuellement de savoir si son ajout systématique à un examen échographique de routine permet de mieux détecter les veaux atteints de BPI.

III-1-2- L'interprétation de l'échographie thoracique

Le **Tableau 8** répertorie les images normales et anormales visibles à l'échographie thoracique lors de BPI (Rabeling et al., 1998, Flock, 2004, Jung et Bostedt, 2004). Afin d'optimiser son utilisation, de nombreuses études ont simplifié et standardisé l'interprétation des images échographiques pour détecter la BPI chez les veaux de race laitière. Cette simplification s'est réalisée en deux étapes : (1) en sélectionnant la lésion la plus fréquemment associée à la BPI ; (2) en déterminant une manière objective d'évaluer cette lésion afin de standardiser et simplifier son interprétation.

i- La lésion la plus évocatrice de bronchopneumonie infectieuse

Certaines lésions sont fréquentes, mais retrouvées chez les veaux sains et malades, comme les artéfacts de queues de comète ou les lignes B. D'autres lésions sont uniquement retrouvées chez les veaux malades, mais sont peu fréquentes, comme l'effusion pleurale ou le pneumothorax (Braun et al., 2018). La lésion de **consolidation pulmonaire** est la lésion la plus souvent retrouvée et spécifique de BPI (Buczinski et al., 2014, Braun et al., 2018). Sa présence est donc un bon indicateur de la maladie. De plus, ces lésions de consolidation pulmonaire ont l'avantage d'apparaître très tôt lors d'infection bactérienne avant les signes cliniques (apparition des lésions dans les 2 h post-infection expérimentale à *Mannheimia haemolytica* contre 6 h pour les signes cliniques) optimisant ainsi une détection précoce de la BPI (Ollivett et al., 2013). Autre propriété intéressante, la lésion de consolidation pulmonaire persiste dans le temps au-delà des signes cliniques favorisant la détection de veaux sous cliniques permettant de renseigner sur les capacités de détection des producteurs (les veaux sous cliniques devraient avoir reçu un traitement) (Ollivett et al., 2013, Hussein et al., 2018, Abdallah et al., 2019).

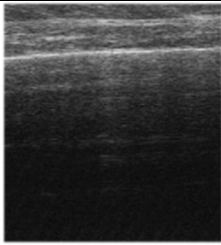
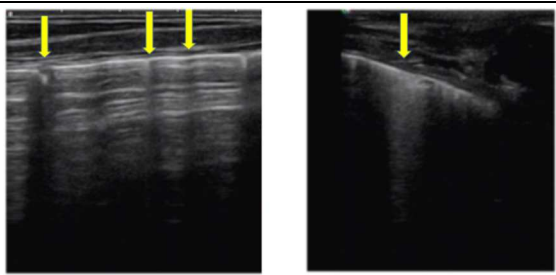

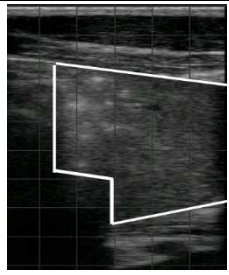

Image échographique	Définition échographique et étiologie	Illustration
Normale	Les plèvres apparaissent comme une ligne hyperéchogène en mouvement de glissement continu dû à la respiration. Le parenchyme pulmonaire n'est pas visible en raison de l'air qu'il contient qui provoque une réverbération du rayon échographique. Des lignes horizontales (ligne A) parallèles sont visibles et représentent la traversée dans le sens inverse du rayon réverbéré entre la plèvre et la peau.	 <p>Plèvres →</p> <p>← Lignes A</p>
Queue de comète ou ligne B	Artéfact secondaire à la réflexion des ultrasons entre deux surfaces réfléchissantes. Ces lésions peuvent être présentes en petit nombre chez des veaux sains ou en plus grand nombre chez les veaux atteints de bronchopneumonie infectieuse ou d'œdème pulmonaire.	
Pneumothorax	Les plèvres sont séparées l'une de l'autre par de l'air. Le mouvement de glissement des plèvres n'est plus visible ventralement au point de pneumothorax (site où la plèvre pariétale et viscérale se rejoignent à nouveau). La lésion est présente lors de bronchopneumonie infectieuse par BRSV; pneumonie par aspiration sévère; trauma (p.ex.; fracture de côte).	Image identique à l'image normale en statique
Effusion pleural	Les plèvres sont séparées l'une de l'autre par un fluide anéchogène (transsudat) ou échogène (exsudat). Cette lésion est présente lors de pleurobronchopneumonie infectieuse, hémothorax, chylothorax ou pyothorax traumatique.	
Consolidation pulmonaire	L'air dans les alvéoles a été remplacé par du matériel inflammatoire, tumoral ou cicatriciel. Un tissu hypo échotique dans le parenchyme pulmonaire similaire à l'aspect échographique du parenchyme hépatique est présent. Cette lésion est présente lors de bronchopneumonie; pneumonie par aspiration; moins fréquemment lors de contusion pulmonaire ou métastase pulmonaire. Elle peut être confondue avec des lésions d'atélectasie qui correspondent à une compression des alvéoles par du liquide ou lors d'immaturation pulmonaire.	 <p>Profondeur</p>
Abcès pulmonaire	Structure circulaire bien définie légèrement hypo échotique à l'intérieur du parenchyme pulmonaire. La lésion est présente lors de bronchopneumonie infectieuse chronique; pneumonie embolique (e.g.; syndrome de la veine cave caudale); tuberculose.	

Tableau 8. – Images échographiques normales et anormales détectables lors de bronchopneumonie infectieuse chez les veaux de race laitière.

Cependant, bien qu'évocatrices de BPI chez les veaux de race laitière, les connaissances actuelles ne nous permettent pas de savoir lesquelles des lésions de consolidation pulmonaire nécessitent d'être traitées ou non pour guérir la BPI. En effet, premièrement, les lésions de consolidation pulmonaire peuvent être non actives, c.-à-d. non infectées et non accompagnées d'un processus inflammatoire (cicatriciel). Mettre un traitement antibiotique ou anti-inflammatoire sur ce type de lésion est alors inutile et n'est donc pas une utilisation judicieuse des antibiotiques. Deuxièmement, dans les lésions actives nécessitant un traitement, la mise en évidence de lésions de consolidation pulmonaires ne nous renseigne pas sur la nature bactérienne ou virale de l'infection. Il est donc impossible actuellement de distinguer les lésions nécessitant un traitement antibiotique pour guérir.

ii- La mesure objective des lésions de consolidation pulmonaire pour détecter la bronchopneumonie infectieuse

Différents paramètres objectifs ont été étudiés afin de simplifier l'interprétation des lésions échographiques lors de BPI chez les veaux de race laitière incluant : la présence et le nombre de sites avec des queues de comètes, la présence d'effusion pleurale, l'irrégularité et/ou l'épaississement des plèvres, la profondeur de consolidation pulmonaire (définie comme l'épaisseur en centimètre de la lésion au départ de la plèvre illustrée dans le **Tableau 8**), et le nombre de sites de consolidation pulmonaire (Buczinski et al., 2014). Parmi ces paramètres, **la profondeur** des lésions de consolidation pulmonaire s'est montré le paramètre le plus simple et le plus performant pour détecter des veaux précédemment traités pour BPI par le producteur (Buczinski et al., 2013) ou prédire un pronostic vital négatif chez des veaux laitiers (Buczinski et al., 2014) et bovins de boucherie (Rademacher et al., 2014, Timsit et al., 2019). Une étude réalisée sur un faible nombre de veaux (n=25) sans signe clinique a également montré une possible corrélation entre la profondeur des lésions de consolidation pulmonaire et la nature virale (< 3 cm) ou bactérienne de l'infection (≥ 3 cm) (Ollivett et al., 2015).

Le **Tableau 9** répertorie une liste non exhaustive des différents seuils et sites échographiques rapportés chez les veaux de race laitière.

Étude	Population	Site échographié	Seuil de profondeur consolidation pulmonaire	Méthode de détermination des sites et du seuils
(Rabeling et al., 1998)	Veaux sevrés de 5 mois	Caudal au cœur	0 cm	Subjective
(Reinhold et al., 2002)	Veaux sevrés de 3 mois	Caudal au cœur	0 cm	Subjective
(Jung and Bostedt, 2004)	Veaux nouveau nés	Caudal au cœur	0 cm	Subjective
(Ollivett et al., 2011)	Veaux présevrés	Caudal + Crânial D au cœur	1 cm	Subjective - Seuil de 1 cm choisi pour différencier les lésions de la plèvre (irrégularité)
(Ollivett et al., 2013)	Veaux sevrés de 4 mois	Caudal + Crânial D au cœur	0 cm	Subjective
(Buczinski et al., 2013)	Veaux présevrés	Caudal au cœur	1 cm	Subjective - Seuil de 1 cm choisi pour différencier les lésions de la plèvre (irrégularité)
(Ollivett, 2014)	Veaux présevrés	Caudal + Crânial D au cœur	3 cm	Objective – Le seuil de 3 cm a été déterminé en sélectionnant la valeur entre le 90 et le 95 percentiles (2,25 cm et 3,5 cm) des lésions de consolidation chez les veaux avec une baisse de croissance
(Buczinski et al., 2014)	Veaux présevrés	Caudal au cœur	1 cm	Subjective - Seuil de 1 cm choisi pour différencier les lésions de la plèvre (irrégularité)
(Ollivett et al., 2015)	Veaux présevrés	Caudal + Crânial D au cœur	0 cm	Subjective
(Buczinski et al., 2015b)	Veaux présevrés de l'étude de Ollivett et al. (2011) et Buczinski et al. (2014)	Caudal au cœur	1 cm	Subjective - Seuil de 1 cm choisi pour différencier les lésions de la plèvre (irrégularité)
(Ollivett et Buczinski, 2016)	Veaux présevrés	Caudal + Crânial D au cœur	Score échographique	Subjective – Élaboration d'un score pour être plus clinique
(Tejero and Bach, 2016)	Veaux présevrés	Caudal + Crânial D au cœur	1 cm	Subjective - Seuil de 1 cm choisi pour différencier les lésions de la plèvre (irrégularité)
(Teixeira et al., 2017b)	Veaux présevrés	Caudal au cœur	0 cm	Subjective
(Dunn et al., 2018)	Veaux présevrés	Caudal + Crânial D au cœur	3 cm	Subjective- basé sur les travaux de Ollivett (2014).
(Braun et al., 2018)	Veaux présevrés et sevrés	Caudal au cœur	0 cm	Subjective
(Hussein et al., 2018)	Veaux sevrés	Caudal au cœur	Score échographique	Subjective - Score de Ollivett et Buczinski (2016)
(Abdallah et al., 2019)	Veaux présevrés	Caudal + Crânial D au cœur	1 cm et 3 cm	Objective – Comparaison des deux seuils - Pas de différence entre les deux seuils sur le gain moyen quotidien des veaux consolidés
(Cramer et Ollivett, 2019)	Veaux présevrés	Caudal + Crânial D au cœur	Score échographique	Subjective - Score de Ollivett et Buczinski (2016)
(Van leenen et al. 2020)	Veaux présevrés et sevrés	Caudal + Crânial D et G au coeur	1 cm, 3 cm et 6 cm	Subjective
(Masmeijer et al., 2020)	Veaux présevrés	Caudal + Crânial D et G au coeur	1 cm et 3 cm	Subjective
(Pravettoni et al., 2021)	Veaux présevrés et sevrés	Caudal + Crânial D au cœur	Score échographique	Subjective - Score de Ollivett et Buczinski (2016)
(Cuevas-Gómez et al., 2021)	Veaux présevrés	Caudal + Crânial D au cœur	Score échographique	Subjective - Score de Ollivett et Buczinski (2016)
(Rhodes et al., 2021)	Veaux présevrés	Caudal + Crânial D au cœur	Score échographique	Subjective - Score de Ollivett et Buczinski (2016)
(Jourquin et al. 2022)	Veaux sevrés	Caudal + Crânial D et G au cœur	0.5 cm	Subjective

Tableau 9. – Différents seuils et sites échographiques rapportés dans la littérature pour détecter la bronchopneumonie infectieuse chez les veaux de race laitière.

Brièvement, différents seuils de profondeur ont été définis pour interpréter les lésions de consolidation pulmonaire : 0 cm, 1 cm, 3 cm, ou encore, 6 cm (Pardon, 2019a). Dans la majorité des études, ces seuils ont été déterminés subjectivement et il n'existe actuellement, aucun consensus sur le seuil qui doit être utilisé pour définir un veau atteint de BPI et encore moins, un veau à traiter avec des antibiotiques.

Récemment, afin de généraliser l'interprétation des lésions de consolidation pulmonaire avec tous les types d'échographes et notamment ceux sans la possibilité de mesurer précisément la profondeur de consolidation, un **score échographique** a été décrit selon l'étendue des lésions (**Annexe 3**) (Ollivett et Buczinski, 2016). Cependant, tout comme les seuils de profondeur, aucune validation n'a été réalisée avec ce score.

III-2 La radiographie thoracique

Tout comme l'échographie thoracique, la radiographie thoracique permet de visualiser les lésions pulmonaires chez les veaux de race laitière. L'avantage de la radiographie par rapport à l'échographie thoracique est que ce test permet de visualiser le poumon dans son ensemble et de détecter des lésions plus en profondeur dans le parenchyme pulmonaire (Reef et al., 1991, Braun et al., 1995). Son inconvénient reste la superposition des tissus qui peut nuire à son interprétation particulièrement problématique chez les bovins adultes (Masseau et al. 2008). Cet inconvénient est cependant contournable chez le veau par sa facilité de manipulation en réalisant plusieurs vues ou en avançant les membres thoraciques (voir section **III-2-2**). Malgré ces avantages et l'accès à de meilleurs appareils portatifs, son utilisation actuelle sur le terrain est limitée et est souvent restreinte au milieu hospitalier (Fowler et al., 2017).

III-2-1 La technique de radiographie thoracique

i- La première description de la technique

Fondé sur ce qui était réalisé dans les autres espèces animales, Slocombe et al. (1988) propose une première description de la technique chez les veaux de race laitière (n=7) dans le but d'étudier les relations anatomiques du poumon selon l'inflation d'air. Les veaux de l'étude étaient sous anesthésie générale. Ils étaient positionnés en décubitus latéral gauche ou droit pour la vue latérale (VL), décubitus dorsal pour la vue ventro-dorsale (VD), et

décubitus ventral pour la vue dorso-ventrale (DV). La machine utilisée était réglée à une intensité constante de 200 mAs, l'énergie des rayons à 85kVp et le temps d'exposition de 33 ms pour la VL, et 100 kV et 50 ms pour les vues VD et DV. La cassette mesurait 35x43cm et était positionnée à 75 cm de la cathode.

ii- Simplification de la technique

Depuis sa première description, la technique a évolué afin d'optimiser son utilisation en pratique en favorisant :

- **Sa réalisation chez des veaux debout sans sédation** en effectuant uniquement la VL (Fowler et al., 2017, Shimbo et al., 2018b). Bien qu'aucune étude n'ait évalué l'intérêt de la VL seule *versus* l'ajout de la VD et DV, ces dernières sont souvent réalisées en deuxième intention en raison des manipulations supplémentaires nécessaires ;
- **Sa réalisation avec des appareils portatifs** pour optimiser son utilisation sur le terrain (Shimbo et al., 2018a,b). Les constantes utilisées dans ce genre d'appareil sont équivalentes aux constantes nécessaires aux appareils fixes pour radiographier le poumon à savoir : une intensité de 150 mAs, une énergie des rayons à 80kVp et un temps d'exposition de 0,13 ms. La cassette utilisée mesure 35x43cm et est placée à une distance de 100 cm. Une grille de rapport 3 : 1 permet de corriger le mieux la dispersion sur de tels appareils portatifs (Shimbo et al., 2018a, b) ;
- **Sa réalisation par une seule personne sur un grand nombre de veaux**, avec le veau maintenu dans une cage à contention où la cassette est fixée à une distance de 100 cm (Shimbo et al., 2018c) (**Annexe 8**).

iii- Les ajouts à la technique initiale

Une technique permettant l'évaluation des lobes crâniens, plus accessibles chez le veau que chez les adultes, a été décrite pour la première fois par Tegtmeier et Arnbjerg (2000). Cette technique permet d'éviter la superposition des membres en avant du cœur. L'inconvénient est que la technique de Tegtmeier et Arnbjerg (2000) nécessitait l'intervention de trois personnes et une sédation profonde du veau. Récemment, Shimbo et al. (2018c) a simplifié cette approche en nécessitant l'intervention d'une seule personne et comprenant une technique de contention que les auteurs jugent rapide et facile une fois son

apprentissage maîtrisé (avec une moyenne de 2 essais [1-7] et une durée de 263 secondes [105-488]) (Shimbo et al., 2018c) (**Annexe 8**). De plus, cette technique dégage les lobes crâniens et s'est montrée plus performante pour détecter les lésions de consolidation pulmonaire mises en évidence en nécropsie que la technique traditionnelle sans dégage des portions crâniées (aire sous la courbe de ROC de 0,884 [technique alternative] *versus* 0,694 [technique standard]) (Shimbo et al., 2018c). Cependant, comme on peut le constater sur la photo de l'**Annexe 8**, cette technique suggère son utilisation sur des veaux relativement calmes, permettant une manipulation facile et une tolérance à la position pour la prise de cliché. Cette technique peut s'avérer difficile sur des veaux plus agités pour toute sorte de raison comme le jeune âge, l'absence d'affection (pas assez malade) ou au contraire, une détresse respiratoire.

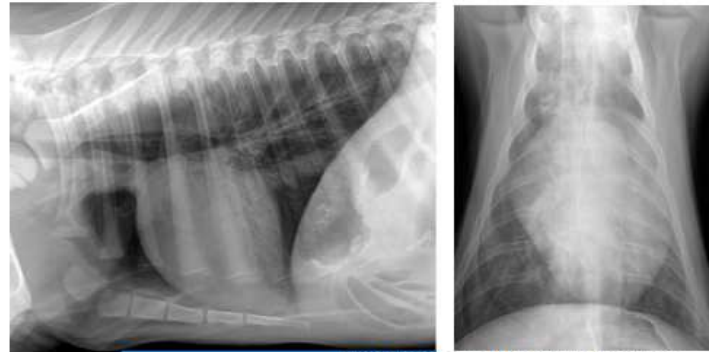
III-2-2- L'interprétation de la radiographie thoracique

La **Figure 2** répertorie les images normales et anormales visibles à la radiographie thoracique lors de BPI (Jung et Bostedt, 2004, Fowler et al., 2017).

i- La lésion la plus évocatrice de bronchopneumonie infectieuse

À la radiographie thoracique, le **patron alvéolaire** est la lésion la plus retrouvée lors de BPI (Jones et al., 1998, Tegtmeier et Arnbjerg, 2000, Fowler et al., 2017, Shimbo et al., 2018c). Le patron alvéolaire peut cependant être confondu avec des structures anatomiques superposées, particulièrement si le veau est debout (Tegtmeier et Arnbjerg, 2000, Shimbo et al., 2018c). Ceci peut être responsable du diagnostic de faux positif (veaux qui n'ont en réalité pas de lésions) et impacter la spécificité du test (Jones et al., 1998). La technique proposée par Shimbo et al. (2018c) permettrait d'éviter cet inconvénient en étendant le membre thoracique vers l'avant. Le rajout de la vue DV ou VD permet également d'éviter ces superpositions.

Images normales



Par son contenu aérien, le parenchyme pulmonaire apparaît noir à la radiographie thoracique. La silhouette cardiaque est visible dans le médiastin d'opacité tissulaire.

Lésions radiographiques

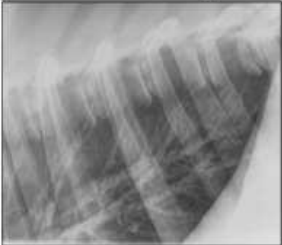

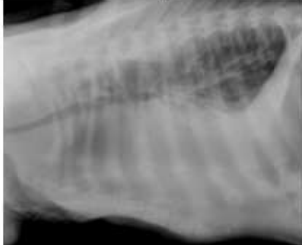


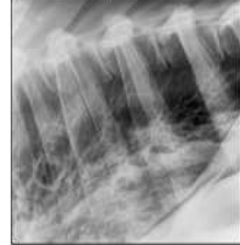
Patron bronchique	Patron interstitiel	Effusion pleurale	Pneumothorax	Patron alvéolaire	Abcès pulmonaire
					
Inflammation des bronches et des bronchioles : épaississement des parois bronchiques	Inflammation du tissu interstitiel : aire pulmonaire d'opacité augmentée accompagné de bords vasculaires floutés	Accumulation de liquide ou de pus entre les plèvres : rétraction des marges des lobe pulmonaire, d'une ligne de fissure pleurale, et la présence de fluide entre la paroi thoracique	Accumulation d'air entre les plèvres : Visualisation d'air (radio transparent; noir) entre les lobes pulmonaires et la paroi thoracique causant une rétraction	Les alvéoles sont remplies par du matériel inflammatoire, tumoral ou cicatriciel : présence d'une zone d'opacité de tissus mous, masquant les vaisseaux pulmonaires localisés à l'intérieur de cette zone et pouvant masquant les structures habituellement visible dans le médiastin (e.g.; silhouette cardiaque)	Présence d'une zone focale d'opacité de tissus mous aux bords relativement bien définis dans le parenchyme pulmonaire

Figure 2. – Images radiographiques normales et anormales détectables lors de bronchopneumonie infectieuse chez les veaux de race laitière.

ii- La mesure objective des lésions alvéolaires pour détecter la bronchopneumonie infectieuse

Contrairement à ce qui a été fait pour l'échographie thoracique, l'interprétation du patron alvéolaire à la radiographie thoracique n'a pas été simplifiée. Leur détection et leur interprétation nécessitent une expertise (Shimbo et al., 2018c). Afin de favoriser son utilisation universelle par les médecins vétérinaires sur le terrain, une simplification de l'interprétation est nécessaire afin d'en faciliter l'utilisation. De plus, on retrouve les mêmes inconvénients que pour l'échographie thoracique, à savoir le manque de distinction entre lésions actives *versus* non actives, virales *versus* bactériennes, ce qui limite, actuellement, son utilisation pour prendre une décision de traitement.

Conclusion PARTIE III : La technique, l'interprétation et l'accessibilité des tests diagnostiques couramment utilisés pour mettre en évidence la BPI présentent des limites résumées dans le Tableau 10. L'amélioration des techniques et des interprétations de ces tests permettrait de simplifier leur utilisation sur le terrain et d'élaborer des stratégies diagnostiques plus performantes pour détecter les veaux de race laitière atteints de BPI.

	SCRD	Auscultation thoracique	Échographie thoracique	Radiographie thoracique
Technique	Simple mais de nombreux scores décrits et aucun consensus sur lequel utiliser.	Simple.	Manque de standardisation sur l'ajout ou non des lobes crâniens.	Demande des installations particulières.
Interprétation	Simple mais seuil parfois déterminé de façon discutable.	Manque d'objectivité – absence d'une définition claire d'un test normal <i>versus</i> anormal.	Manque de standardisation sur le seuil d'interprétation des profondeurs de consolidation (0 <i>versus</i> 1 <i>versus</i> 3 cm).	Manque de standardisation sur les lésions significatives et la manière de les interpréter.
Accessibilité	Accessible à tous.	Non enseigné aux producteurs.	Non enseigné aux producteurs.	Difficilement accessible actuellement en pratique.

Tableau 10. – Limites des techniques, des interprétations et de l'accessibilité des tests diagnostiques couramment utilisés pour diagnostiquer la bronchopneumonie infectieuse chez les veaux de race laitière.

PARTIE IV : Les limites de performance des tests actuels pour détecter la BPI chez les veaux de race laitière

Les tests diagnostiques actuels étudiés dans cette thèse présentent des limites de performance qui peuvent être réelles, c.-à-d., réellement limitées pour détecter la BPI, ou biaisées par la manière dont les performances ont été déterminées. En effet, pour valider un test diagnostique afin de détecter des maladies infectieuses, l'**OMSA** recommande la succession de trois étapes (Diaz 2014) : (1) l'évaluation de la faisabilité du test, étape durant laquelle la technique est standardisée et optimisée (étape de développement) ; (2) la détermination des caractéristiques diagnostiques du test dans différentes populations ; (3) l'étude de la répétabilité et reproductibilité du test par différents opérateurs. Aucune réglementation n'exige actuellement la réalisation de ces étapes en médecine vétérinaire pour mettre un test sur le marché. Ces étapes sont alors rarement réalisées ou de façon incomplète. Dans cette prochaine partie, je reviendrai sur les limites de réalisation de ces étapes pour les tests décrits en **Partie III** visant à détecter les BPI chez les veaux de race laitière.

I- Les limites de développement des tests pour détecter la bronchopneumonie infectieuse

I-1- Les étapes générales nécessaires pour développer un test

I-1-1- Le développement d'un test

Avant de développer un test, plusieurs paramètres doivent être préalablement définis : (1) la définition des individus cibles ; (2) la définition précise de la maladie ; (3) la définition des facteurs de variations pouvant affecter les prélèvements et les performances du test (Diaz, 2014). Une fois ces paramètres définis, les caractéristiques intrinsèques au test sont améliorées (étape de validation interne), à savoir sa capacité à donner une mesure exacte (l'exactitude), sa capacité à donner la même valeur (la précision), et sa capacité à ne pas être affecté par des variations mineures des facteurs de variations pouvant affecter les prélèvements et les performances du test (la robustesse). La technique est également standardisée afin d'optimiser ces caractéristiques (Diaz, 2014). Puis dans un second temps, ces caractéristiques intrinsèques au test sont évaluées dans différentes populations (étape

de validation externe). À noter que lorsqu'un test est déjà développé et standardisé dans une espèce animale, la phase de développement consiste en une nouvelle validation interne et externe de la technique pour cette nouvelle espèce animale (Diaz, 2014).

I-1-2- Le développement d'un score clinique diagnostique

En médecine humaine, le développement d'un score clinique diagnostique suit les recommandations du **TRIPOD** (*Transparence Reporting of multivariable prediction model for Individual Prognosis or Diagnosis*) (Moons et al., 2015). Brièvement, le score est réalisé à partir de la formule mathématique suivante :

Pour Y = variable à prédire, dichotomique (malade/non malade) ou stratifiée (p. ex. ; haut *versus* moyen *versus* faible *versus* très faible), et p le nombre de prédicteurs :

$$F(Y) = W_1X_1 + W_2X_2 + \dots + W_pX_p$$

Où F correspond à la fonction *logit* qui permet de définir directement les poids attribués à chaque prédicteur (W). Pour aboutir à cette équation, l'élaboration d'un score nécessite au préalable :

- (1) **Une sélection des prédicteurs (X) et leurs techniques d'évaluation :** Cette sélection peut être objective (p. ex. ; incorporation de prédicteur statistiquement différent entre des individus malades *versus* sains) ou subjective (p. ex. ; incorporation de prédicteur pratique à évaluer dans un contexte donné) ;
- (2) **La mesure statistique de la relation de ces prédicteurs avec la variable à prédire et l'attribution de poids relatif (W) pour chaque prédicteur ;**
- (3) **La validation interne du score** qui permet l'optimisation de son exactitude et de sa robustesse. Cette étape permet notamment d'ajuster le score au surajustement (*overfitting*) c.-à-d. ; la production d'un score qui correspond trop étroitement ou exactement aux données de la population de développement, et peut donc ne pas s'adapter à des données supplémentaires ou prédire de manière aussi fiable les observations futures. Des exemples de méthodes pour limiter ce surajustement sont de calculer les poids sur des sous-échantillons différents de la population de développement (technique de validation croisée) ou sur de nouveaux échantillons

de la population de développement obtenus après réplification multiple des données (technique du *bootstrapping*) ;

- (4) **La validation externe du score** qui consiste à évaluer les performances du modèle dans des populations différentes. À cette étape, le score peut être à nouveau ajusté en rajoutant un prédicteur, ou recalibré en recalculant les poids. La validation externe peut être réalisée dans des données incluant : une population similaire, mais échantillonnée à une période différente (validation temporelle) ou par d'autres investigateurs dans un autre pays (validation externe géographique); une population incluant les mêmes participants d'un autre type de production (p. ex. ; génisses de remplacement *versus* veaux lourds); ou une population incluant d'autres types de participants (p. ex. ; veaux de boucherie).

I-2- Le développement des tests pour détecter la bronchopneumonie infectieuse

I-2-1- L'inexactitude de la définition de la bronchopneumonie infectieuse

Une des premières étapes de développement d'un test est de définir avec exactitude la maladie ou l'affection qu'il est censé mesurer (Diaz, 2014). On parle de condition cible (*target condition*) (Leeflang et al., 2013). En absence de définition exacte, les tests diagnostiques développés pour détecter la maladie peuvent ainsi se révéler inexacts. Or, la définition de BPI est actuellement inexacte (Buczinski et Pardon, 2020). L'amélioration du développement et des performances des tests pour détecter la BPI chez les veaux de race laitière passe donc par une amélioration de la définition de la maladie et notamment, dans le contexte d'élevage actuel, la définition visant à diagnostiquer les veaux à traiter de BPI (définition diagnostique). Lorsqu'un test de référence définit clairement la maladie et son absence (on parle de test étalon or ou test *gold standard*), on peut l'utiliser pour évaluer un nouveau test (index test). Cependant, en l'absence de *gold standard*, il est important de tenir compte de l'imperfection des tests disponibles avec lesquels le nouveau test va être comparé (voir section **II-1-1**).

I-2-2- Les limites de développement

L'auscultation thoracique, l'échographie thoracique et la radiographie thoracique ont été développées chez d'autres espèces animales. Le SCRD, spécifique de l'espèce bovine, a nécessité une phase de développement complète dans les veaux de race laitière. Le **Tableau 11** résume les étapes de développement des principaux SCRD rapportés pour détecter la BPI chez les veaux de race laitière.

On constate que sur les 5 principaux SCRD rapportés, seulement 3 d'entre eux ont été développés. Sur les 3 SCRD développés, la mesure de la BPI est variable : définie parfois par des tests imparfaits sujets à des erreurs de classification des malades *versus* non malades, et parfois par des tests ne permettant pas de se limiter aux cas de BPI nécessitant un traitement comme l'échographie thoracique qui ne distingue pas les lésions cicatricielles ou en voie de guérison (non actives) *versus* des lésions inflammatoires et infectées (actives). Les performances des scores du Wisconsin et Californien ont été testés dans différentes populations. La plupart de ces validations externes impliquaient un test comparateur imparfait (voir **Tableau 15**) (Love et al., 2016, Lowie et al., 2022).

I-2-3- Les limites de standardisation

L'auscultation thoracique, l'échographie thoracique et la radiographie thoracique ont été développées dans d'autres espèces animales. Cependant, aucune validation interne et externe de ces tests n'a été réalisée spécifiquement pour son utilisation aux veaux de race laitière. Cette absence de validation soulève d'importantes limites dans leurs utilisations actuelles. Pour l'auscultation thoracique, bien que la technique soit bien définie chez les veaux de race laitière (l'aire d'auscultation, l'emploi d'un stéthoscope), la définition d'un veau positif/malade et négatif/non malade n'est pas clairement définie et variable selon les études (Buczinski et al., 2014) *versus* (Buczinski et al., 2016b) (voir **Tableau 16**). Ce manque de standardisation rend l'interprétation de l'auscultation thoracique difficile et explique les désaccords possibles entre plusieurs opérateurs (Pardon et al., 2019b). On retrouve ce manque de standardisation pour l'échographie thoracique où les sites à échographier et le seuil d'interprétation ne font pas consensus (**Tableau 9**). Il en est de même pour la radiographie thoracique concernant l'extension ou non des membres thoraciques lors de la prise des clichés et les vues à réaliser (VL *versus* VL+VD ou DV).

Score clinique	Étape 1		Étape 2		Étape 3 : Validation interne	Étape 4 : Validation externe	Limites
	1a : Sélection des prédicteurs	1b : Évaluation des prédicteurs	2a : Mesure de la bronchopneumonie infectieuse	2b : Calcul des poids			
SCORE CIS (Amrine et al., 2013)	Sélection subjective	Évaluation subjective par une échelle de gravité.	Aucune mesure statistique de la relation des prédicteurs choisis avec la maladie.		Aucune	Aucune	Absence de développement et de validation interne et externe
SCORE DU WISCONSIN (McGuirk, 2008)	Sélection subjective	Évaluation subjective par une échelle de gravité.	Aucune mesure statistique de la relation des prédicteurs choisis avec la maladie.		Aucune	Validé dans différentes populations (Buczinski et al., 2015, Love et al., 2016) et races (Lowie et al., 2022)	<ul style="list-style-type: none"> ◆ Absence de développement et de validation interne ◆ Validation externe utilisant un test imparfait pour définir la maladie (biais de mauvaise classification) Love et al., (2016) et Lowie et al., (2022)
SCORE CALIFORNIEN (Love et al., 2014)	Sélection objective : choix des prédicteurs selon leur relation linéaire avec la présence de la maladie.	Sélection objective : choix de la technique d'évaluation selon la simplicité et les résultats de validation des modèles statistiques.	1. PCR positive pour BRSV, IBR-1, ou BVD-1 et BVD-2 Et/ou 2. Score du Wisconsin ≥ 5 + culture positive de bactéries aérobie ou <i>Mycoplasma spp.</i>	Analyse de régression logistique multivariée.	Aucune	Validé dans différentes populations (Love et al., 2016) et races (Lowie et al., 2022)	<ul style="list-style-type: none"> ◆ Biais de classification lors du développement et de la validation externe (définition de la maladie avec des tests imparfaits) ◆ Non limité au statut à traiter de la bronchopneumonie infectieuse ◆ Biais d'incorporation (utilisation du score du Wisconsin pour définir la maladie ce qui peut surestimer les poids due à la colinéarité entre le test utilisé pour définir la maladie et les mêmes signes cliniques évalués) ◆ Absence de validation interne
SCORE CALIFORNIEN Québec (Buczinski et al., 2018c)	Sélection subjective : choix des mêmes prédicteurs que le score Californien.	Sélection subjective avec le choix de la même technique d'évaluation que le score Californien.	Analyse bayésienne de classe latente à partir des résultats de l'échographie thoracique.	Analyse de régression logistique multivariée avec des <i>priors</i> sur les coefficients selon un sondage d'expert.	Par analyse Bayésienne	Aucune	<ul style="list-style-type: none"> ◆ Non limité au statut à traiter de la bronchopneumonie infectieuse ◆ Absence de validation externe
SCORE DE MAIER (Maier et al., 2019)	Sélection objective : choix des prédicteurs selon leur relation linéaire avec la présence de la maladie.	Sélection subjective : Présence-Absence.	Utilisation en parallèle de l'auscultation thoracique + l'échographie thoracique.	Analyse de régression logistique multivariée.	Par validation croisé	Aucune	<ul style="list-style-type: none"> ◆ Biais de classification (définition de la maladie avec des tests imparfaits) ◆ Non limité au statut à traiter de la bronchopneumonie infectieuse ◆ Absence de validation externe

Tableau 11. – Étapes de développement des scores cliniques respiratoires diagnostiques rapportés pour détecter la bronchopneumonie infectieuse chez les veaux de race laitière.

II- Les limites de détermination des performances des tests diagnostiques pour détecter la bronchopneumonie infectieuse

II-1- Procédure générale de détermination des performances d'un test diagnostique

On entend par performances diagnostiques d'un test sa capacité à détecter la maladie. Pour déterminer ces performances, on compare un test d'intérêt ou test index, à un test comparateur (test standard de référence ou *reference standard test*). À noter que les tests étudiés dans cette thèse sont tous des tests dichotomiques, c.-à-d. ; dont le résultat est soit positif, soit négatif. Les paragraphes qui suivront se restreindront donc à ce type de test.

II-1-1- Test *gold standard* versus test standard de référence

Selon les performances de détection de la maladie, le test comparateur peut être :

- **Un test *gold standard***, c.-à-d. ; un test 100 % exact pour dire si un animal est malade ou non. Pour la BPI, seules la nécropsie et l'histologie des lésions permettent d'avoir de telles performances. Cet examen se réalisant en post-mortem, cela est contraignant lors d'évaluation d'un grand nombre de veaux (Buczinski et O'Connor, 2016c). En raison de ses excellentes performances, le CT scan est une alternative comme test *gold standard* ante-mortem pour détecter les lésions pulmonaires de BPI chez les veaux de race laitière (Lubbers et al., 2007, Ohlerth et al., 2014). Cependant, il ne permet pas de savoir si ces lésions sont cicatricielles (non actives) ou enflammées et infectées (actives). Actuellement, il n'existe donc pas de test *gold standard* pratique pour détecter les BPI à traiter du vivant de l'animal.
- **Un test standard de référence** constitue une alternative en l'absence ou l'impossibilité d'utilisation d'un test *gold standard*. Ce test est construit à partir de plusieurs tests basés soit : (1) sur une règle prédéfinie (test standard composite) ; (2) sur un consensus parmi des experts (panel diagnostique) ; (3) ou par un modèle statistique où le statut de la maladie non observé (appelé classe latente ou variable latente) est estimé à partir des mesures communes des tests diagnostiques imparfaits (analyse de classe latente) (Reitsma et al., 2009). Les avantages et inconvénients de ces méthodes d'élaboration d'un test standard de référence sont résumés dans le **Tableau 12** (Pepe, 2003a, Reitsma et al., 2009).

	Construction	Avantages	Inconvénients
Test standard composite	Selon une règle prédéfinie d'utilisation de plusieurs tests (en série, en parallèle).	Possède une meilleure propriété discriminatoire que chaque test isolé.	Le principal problème est de savoir si la combinaison de test de référence est un moyen significatif de définir la condition ciblée; s'il n'y a pas de consensus sur la règle prédéfinie pour définir la condition, il est plus difficile à appliquer.
Panel diagnostique	C'est un panel d'expert qui juge de la présence ou de l'absence de la condition à l'aide de toutes les informations pertinentes disponibles.	Par l'utilisation de toutes les informations (historiques, examen physique, etc.), cette approche est plus proche de ce qui est fait en pratique et donc de la condition ciblée.	Absence de définition stricte de la condition engendrant différents points de vue de la part des membres du panel. Pour pallier cela, l'accord entre les experts doit être rapporté comme outil de validation du panel diagnostique.
Analyse de classe latente	Il s'agit d'un modèle qui relie les résultats de différents tests à une classe latente dont le statut est estimé à partir des observations réalisées avec les tests.	Méthode plus objective car basé sur les données actuelles de performances des tests prenant ainsi en compte leurs imperfections.	La condition ciblée n'est pas définie de manière clinique ce qui peut manquer de clarté sur ce que signifient les résultats en pratique; L'autre défi est dans l'évaluation de la robustesse des résultats lorsqu'ils sont basés sur diverses suppositions (optimiste ou pessimiste), particulièrement quand le nombre d'animaux est faible (Lambert et al., 2005).

Tableau 12. – Avantages et inconvénients des méthodes d'élaboration d'un test standard de référence.

Parmi ces méthodes, aucune n'est donc parfaite. Le choix de la méthode dépend de l'exactitude de la définition de la condition cible (pour plus de détails, voir l'algorithme décisionnel présenté par Reitsma et al. [2009] en **Annexe 9**). Brièvement, lorsque la maladie est clairement définie par les résultats de tests diagnostiques, le test standard composite est utilisé. Sinon, ce sont soit l'analyse de classe latente ou le panel diagnostique pour une vision plus clinique qui sont utilisés. Étant donné que la définition diagnostique de la BPI est inexacte (voir **Partie I, III**), l'utilisation de l'analyse de classe latente et du panel diagnostique comme choix de test standard de référence sont les méthodes les plus appropriées.

II-1-2- Les paramètres de performances individuelles

À la suite de la comparaison du test index et du test comparateur, lorsque la maladie est considérée dichotomique (malade/non malade), on obtient la table 2x2 suivante :

	Test comparateur +	Test comparateur -	
Test index +	VP	FP	VP+FP
Test index -	FN	VN	FN+VN
	VP+FN	FP+VN	n

VP : vrai positif; FP : faux positif; FN : faux négatif; VN : vrai négatif; n : nombre total d'individu testé

À partir de cette table, deux paramètres objectifs sont calculés pour décrire les performances d'un test diagnostique chez un individu :

- **La sensibilité (Se)** = la proportion d'animaux malades qui ont un test positif (VP/VP+FN). La Se exprime le potentiel d'un test à reconnaître les veaux malades (Šimundić, 2009, Buczinski et O'Connor, 2016c) ;
- **La spécificité (Sp)** = la proportion d'animaux non malades qui sont déclarés négatifs (VN/FP+VN). La spécificité exprime la faculté d'un test à reconnaître les veaux non malades (Šimundić, 2009, Buczinski et O'Connor, 2016c).

La Sp est complémentaire à la Se et permet de décrire l'exactitude du test à diagnostiquer la maladie (Šimundić, 2009, Buczinski et O'Connor, 2016c). La Se et Sp vont dépendre : (1) des caractéristiques de la population (âge, sexe, race, etc.) ; (2) de l'environnement d'application du test (condition d'élevage, groupe ou individuel, etc.) ; (3) du spectre de gravité de la maladie. En revanche, la Se et Sp sont relativement indépendants de la prévalence de la maladie, bien que le seuil d'interprétation puisse varier selon cette dernière (Dohoo et al., 2012). La principale limite de ces deux paramètres est leur manque d'application en pratique. En effet, la Se (Sp) permet de répondre à la question : « si un veau est malade (non-malade), quelle est la probabilité que le test soit positif (négatif) ». En pratique, il est plus intéressant de raisonner à l'inverse, à savoir : si un veau teste positif (négatif) quelle serait sa probabilité d'être réellement malade (non malade)? (Timsit et al., 2018). C'est pourquoi, à partir de la Se et Sp, un certain nombre de paramètres dérivés peuvent être calculés pour exprimer différemment les performances d'un test. Le **Tableau 13** répertorie les paramètres les plus utilisés pour décrire les performances de détection individuelle des tests diagnostiques dichotomiques.

Pour optimiser les performances d'un test diagnostique dichotomique, différents seuils d'interprétation peuvent être définis selon l'intérêt du test : (1) l'intérêt diagnostique, le seuil d'interprétation est alors choisi afin de minimiser les proportions de FP et FN (p. ex. ; la courbe de ROC [Se par rapport à 1-Sp], l'index de Youden [Se + Sp – 1]) (Šimundić, 2009) ; (2) l'intérêt économique, le seuil d'interprétation est alors choisi afin d'avoir le moins d'impact économique possible. Une des méthodes est alors de déterminer la valeur du coût de mauvaise classification (*Misclassification Cost Term*, MCT) (Greiner, 1996). Ce terme ne considère pas seulement la Se et la Sp du test, mais aussi la prévalence de la maladie et le ratio du coût (r) relatif entre les FN et les FP selon la formule :

$$\text{MCT} = (1 - \text{Prévalence}) * (1 - \text{Sp}) + r * \text{Prévalence} * (1 - \text{Se})$$

Pour la BPI, ce terme a déjà été rapporté basé sur l'avis d'expert considérant les FN comme plus coûteux que les FP (r variant de 3 : 1, 8 : 1, et 20 : 1) (Buczinski et al., 2015b, Buczinski et al., 2018c). Que ce soit dans un intérêt diagnostique ou économique, le seuil d'interprétation peut ainsi être ajusté à la prévalence de la BPI dans une population (Buczinski et al., 2018c).

II-1-3- Les paramètres de performances de groupe

On définit un test de groupe comme l'évaluation d'un échantillon d'animaux dans un groupe (n), auquel on applique des règles de décision qui classifient un groupe comme positif ou négatif basé sur les résultats des tests réalisés individuellement sur n animaux (Christensen et Gardner, 2000). Un test de groupe dépend donc : (1) du nombre d'animaux à échantillonner (n) ; (2) du nombre d'animaux positifs au test à partir duquel on considère le groupe positif (seuil de positivité : k) ; (3) de la Se et Sp individuelles du test ; et (4) de la prévalence de la maladie dans les groupes infectés. Les performances d'un test de groupe peuvent être exprimées de la même manière qu'un test individuel, mais au niveau du troupeau (*Herd*) ou du groupe. On parle alors de Se de groupe (**HSe**), Sp de groupe (**HSp**), et des valeurs prédictives associées (**Tableau 13**). Lors du développement d'un test de groupe, le nombre d'animaux à échantillonner (n) et le seuil de positivité (k) sont ainsi déterminés afin d'optimiser la HSe et HSp. Dans un contexte d'élevage de veaux de race laitière, un test pourra s'appliquer sur un lot de veaux, la totalité des veaux dans un troupeau, un rassemblement de veau du même âge ou un rassemblement de veau dans une même localisation.

Paramètres de performance individuelle	Calculs	Signification	Avantages	Inconvénients
Sensibilité (Se)	Se = nombre de veaux positifs au test index et au test comparateur (VP) / nombre de veaux positifs au test comparateur (VP+FN)	Probabilité d'avoir un résultat positif chez un veau malade.	♦ Indépendant de la prévalence (bien que le seuil d'interprétation puisse varier selon la prévalence) ♦ Comparaison facilitée entre plusieurs populations et études	♦ Dépend du spectre de la maladie ♦ Dépend des caractéristiques de la population (e.g., sexe, âge, race etc.) et l'environnement d'application ♦ Plus difficilement interprétable en pratique
Spécificité (Sp)	Sp = nombre de veaux négatifs au test index et au test comparateur (VN) / nombre de veaux négatifs au test comparateur (VN+FP)	Probabilité d'avoir un résultat négatif chez un veau non -malade.		
Valeur prédictive positive (VPP)	VPP = proportion de veaux malades avec un test positif (VP) / le total de veaux avec un résultat positif (VP+FP)	Probabilité d'avoir la maladie chez un veau avec un résultat positif.	Interprétation plus intuitive que la Se et Sp dans la prise de décision clinique dans une population avec une certaine prévalence	♦ Dépend de la Se et Sp ♦ Dépend de la prévalence de la maladie dans la population ♦ Comparaison difficile entre les populations et études
Valeur prédictive négative (VPN)	VPN = proposition de veaux non malades avec un test négatif (VN) / le total de veaux avec un test négatif (VN+FN)	Probabilité de ne pas avoir la maladie chez un veau avec un résultat négatif.		
Rapport de vraisemblance positif (LR+)	LR+ = pourcentage de veaux malades avec un test positif (Se = (VP/(VP+FN)) / le pourcentage de veaux non malades avec un test positif (1-Se = (FP/(FP+VN))	Combien un veau malade a plus de chance d'avoir un test positif comparativement à un veau non-malade.	♦ Interprétation plus intuitive que la Se et la Sp dans la prise de décision clinique ♦ Seuils d'interprétation pré définis : LR+ > 10 utile pour confirmer une maladie LR- < 0.1 utile pour infirmer une maladie ♦ Lié aux côtes (odds) pré-test et post-test via le théorème de Bayes : Odds post-test = LR*Odds pré-test	Tout comme la Se et Sp les LR dépendent du spectre de la maladie et des caractéristiques de la population (e.g., sexe, âge, race etc.)
Rapport de vraisemblance négatif (LR-)	LR- = pourcentage de veaux malades avec un test négatif (1-Sp = (FN/(VP+FN)) / le pourcentage de veaux non malades avec un test négatif (Sp = (VN/(FP+VN))	Combien un veau malade à plus de chance d'avoir un test négatif, comparativement à un veau non-malade.		
Paramètres de performance de groupe	Calculs	Signification	Avantages	Inconvénients
Sensibilité (HSe)	$HSe = \sum_0^{k-1} C_{k-1}^n (AP_{pos})^{k-1} * (1 - AP_{pos})^{n-(k-1)}$ Avec AP _{pos} la prévalence apparente obtenue considérant l'imperfection d'un test individuel AP _{pos} = Prévalence réelle*Se + (1-Prévalence réelle) *(1-Sp)	Probabilité d'avoir un résultat de groupe positif chez un groupe positif.	. Permet de définir le nombre d'individu à échantillonner et le seuil de positivité	♦ Dépend de Se et Sp du test ♦ Dépend de la prévalence ♦ Plus difficilement interprétable en pratique
Spécificité (HSp)	$HSp = \sum_0^{k-1} C_{k-1}^n (Sp)^{n-(k-1)} * (1 - Sp)^{k-1}$	Probabilité d'avoir un résultat de groupe négatif chez un groupe négatif.		
Valeur prédictive positive (HPV+)	HPV+ = (HSe * HTP) / (HSe * HTP + (1-HTP) * (1-HSp)) Avec HTP la probabilité pré-test qu'un groupe soit positif	Probabilité qu'un groupe soit positif avec un résultat positif.	. Plus facilement interprétable en pratique	♦ Dépend de tous les facteurs affectant HSe et HSp ♦ Dépend de la vraie proportion de groupe positif (HTP) (peut dépendre du type de groupe, taille du groupe, autres facteurs de risques)
Valeur prédictive négative (HPV-)	HPV- = HSp * (1- HTP) / (HSp * (1- HTP) + HTP * (1-HSe)) Avec HTP la probabilité pré-test qu'un groupe soit positif	Probabilité qu'un groupe soit négatif avec un résultat négatif.		

Tableau 13. – Paramètres couramment utilisés pour décrire les performances individuelles et de groupe d'un test diagnostique dichotomique.

II-1-4- Les paramètres de performances de plusieurs tests

L'idéal pour diagnostiquer une maladie est d'avoir un test dont la Se et la Sp se rapproche du test *gold standard* 100 % exact. En pratique, ces tests sont rares. L'un des moyens pour augmenter les performances d'un test est de le combiner à un autre test dont les performances sont complémentaires, comme coupler un test sensible, mais peu spécifique, à un test spécifique mais peu sensible. L'utilisation d'un test par rapport à l'autre peut être soit : (1) en série : c.-à.-d. que l'individu ou le groupe est déclaré positif si les deux tests sont positifs ; (2) soit en parallèle : c.-à.-d. que l'individu ou le groupe est déclaré positif si l'un des deux tests est positif. Les formules suivantes permettent d'estimer les performances de ces stratégies (Dohoo et al., 2012) :

- Utilisation en série d'un test 1 et 2 :

$$Se_{\text{série}} = Se_1 * Se_2 + covar(+)$$

$$Sp_{\text{série}} = 1 - (1 - Sp_1) * (1 - Sp_2) - covar(-)$$

- Utilisation en parallèle d'un test 1 et 2 :

$$Se_{\text{parallèle}} = 1 - (1 - Se_1) * (1 - Se_2) - covar(+)$$

$$Sp_{\text{parallèle}} = Sp_1 * Sp_2 + covar(-)$$

Avec $Covar(+)$ = $p_{111} - Se_1 * Se_2$ avec p_{111} = proportion d'individus malades et positifs aux deux tests; et $Covar(-)$ = $p_{000} - Sp_1 * Sp_2$ avec p_{000} = proportion d'individus non malades et négatifs aux deux tests. À noter que si les deux tests sont indépendants, on a alors $Covar(+)$ = $Covar(-)$ = 0.

Le calcul de ces performances prend en compte la possible covariance entre les tests, c.-à.-d. ; le lien entre le résultat des deux tests (quand le résultat d'un test est associé au résultat de l'autre test). D'un point de vue pratique, si le but est de confirmer une maladie, on privilégiera la stratégie, en série ou en parallèle, avec la meilleure Sp pour limiter la proportion de FP. Inversement, si le but est d'infirmer une maladie, on privilégiera la stratégie avec la meilleure Se pour limiter la proportion de FN (Dohoo et al., 2012).

II-1-5- Les facteurs affectant la détermination des performances d'un test

Dans certaines circonstances, la détermination des performances d'un test diagnostique manque de précision et d'exactitude (Schmidt et Factor, 2013). La connaissance des facteurs affectant la détermination de ces performances est primordiale pour juger de l'efficacité d'un test.

i- Les facteurs affectant la précision

Mathématiquement, les paramètres de performance sont des proportions. Pour pouvoir exprimer ces proportions avec précision, il faut donc avoir un certain nombre d'individus malades pour

estimer la Se et non malades pour la Sp. Plus ce nombre est grand, plus l'intervalle de confiance (**IC**) ou de crédibilité lors d'analyse bayésienne (**ICB**) autour de la proportion estimée est étroit, et plus l'estimation est précise. Ainsi, l'estimation de la Se et de la Sp dépend : du nombre total d'individus dans l'étude mais également de la prévalence de la maladie. En effet, si la prévalence de la maladie est faible, la précision sera bonne pour la Sp (beaucoup d'individus non malades), au détriment de la Se (moins d'individus malades). À l'inverse, si la prévalence est élevée, la précision sera bonne pour la Se (beaucoup d'individus malades) au détriment de la Sp (moins d'individus non malades). Dans le cas de la comparaison de deux tests, la puissance statistique est donc à prendre en compte et dépend du nombre d'individus nécessaires pour détecter une différence entre deux Se ou Sp. La covariance entre les deux tests est également à prendre en compte dans la comparaison (Dohoo et al., 2012).

ii- Les facteurs affectant l'exactitude

Le manque d'exactitude des performances d'un test diagnostique est causé par les biais des études les rapportant. Les valeurs estimées de Se et Sp, sont déviées de leurs valeurs réelles. Le **Tableau 14** répertorie une liste non exhaustive des biais les plus communément rencontrés lors d'étude diagnostique (Schmidt et Factor, 2013 ; Buczinski et O'Connor, 2016c ; Buczinski et Pardon, 2020).

Pour détecter les biais de spectre, la connaissance du devis expérimental (*design*) de l'étude est nécessaire. Deux types de *design* peuvent être utilisés pour sélectionner la population : (1) *two-gate design*, qui correspond à un *design* cas-témoin favorisant la sélection d'un spectre de la maladie restreint ; (2) le *one-gate design*, qui correspond à une sélection d'une population dont le statut de maladie est inconnu au moment du test et donc avec un spectre de maladie plus large (Rutjes et al., 2005, Buczinski et O'Connor, 2016c).

Biais	Définition	Impact potentiel	Exemple appliqué à la bronchopneumonie infectieuse
Biais de spectre	Survient si la population ne représente pas le spectre complet de la maladie (cas référés, sélection de cas extrêmes).	Surestimation de la Se et Sp, les cas où le diagnostic est plus difficile étant exclus.	La sélection de veaux malades et non malades selon leur score clinique qui n'inclue pas les veaux malades mais sous-cliniques.
Biais de classification	Survient quand le test comparateur n'est pas un <i>gold standard</i> 100% exact.	Sous-estimation potentielle; dépend de la corrélation entre le test index et le test comparateur.	Déterminer les performances de l'échographie thoracique chez des veaux définis malades selon leur score clinique sans prendre en compte les performances imparfaites du score clinique.
Biais de revue diagnostique	Survient quand il y a connaissance des résultats du test index lors de la réalisation et l'interprétation du test comparateur.	Généralement une surestimation des performances du test index; la connaissance des résultats du test index peut influencer l'interprétation du test comparateur.	Confirmer la présence de lésions au CT scan en sachant qu'il y a des lésions à l'échographie thoracique.
Biais d'incorporation	Survient quand il y a incorporation du test à investiguer dans la construction du test comparateur.	Généralement une surestimation des performances du test index.	Quand l'échographie thoracique est le test index et que ces résultats sont utilisés par les experts pour déterminer si un veau est malade ou non par panel diagnostique.
Biais de vérification partielle	Survient quand la totalité des individus ne sont pas testés. Les résultats du test index a un impact sur la probabilité de réaliser le test comparateur.	L'absence de test comparateur chez les individus non testés peut sur ou sous-estimer les performances du test index.	Réaliser le CT scan uniquement sur les veaux positifs à l'échographie thoracique et ne pas prendre en compte de la possible inexactitude à définir les veaux négatifs dans les analyses de détermination des performances.
Biais de vérification différentielle	Survient quand différents tests comparateurs sont appliqués aux individus.	L'exactitude des différents tests comparateurs peuvent différer et doivent être pris en compte dans les analyses d'exactitude du test index.	Les veaux sévèrement atteints vont être vérifiés en nécropsie mais les veaux avec des signes cliniques moins graves sont seulement suivis cliniquement.
Biais de revue clinique	Survient quand les informations cliniques de l'individu recevant le test sont manquantes.	Dans ce cas-là, les résultats du test index ont besoin d'un contexte spécifique.	Détermination des performances d'un test sans conséquences des caractéristiques de la population (âge, production, sexe).
Biais de résultats indéterminés ou d'exclusion d'individu	Survient quand certains individus sont exclus de l'étude à cause de l'impossibilité de réaliser ou d'interpréter le test index ou le test comparateur.	Tendance à une surestimation des performances.	Exclusion des veaux dont la portion crâniale droite n'a pu être évaluée lors de l'examen échographique.

Tableau 14. – Principaux biais affectant l'exactitude des performances d'un test diagnostique.

II-2- Les performances actuelles des tests diagnostiques pour détecter la bronchopneumonie infectieuse chez les veaux de race laitière

II-2-1- Les performances individuelles

i- Les scores cliniques respiratoires diagnostiques

Le **Tableau 15** répertorie les Se et Sp des principaux SCRD utilisés chez les veaux de race laitière pour détecter la BPI. Pour certains scores, on distingue les performances de dépistage, c.-à-d. quand le score est appliqué sur n'importe quel veau, et de confirmation, c.-à-d. quand le score est appliqué sur des veaux suspects. La précision des performances est estimée selon l'étendue de l'IC (fréquentiste) ou du ICB (Bayésien) autour de la valeur estimée de Se et Sp. Les biais permettant de juger de l'exactitude de ces valeurs sont rapportés dans la dernière colonne.

Score clinique	Étude/Design	Population de veaux	Confirmation de la BPI (test comparateur)	Performances	Limites
SCORE DE CIS	(Amrine et al., 2013) / Expérimentale	Génisses de remplacement sevrées expérimentalement inoculées avec <i>Mycoplasma bovis</i> .	Test <i>gold standard</i> = Nécropsie 48 heures après la réalisation du score; Des seuils de consolidation ont été définis (5, 10, 15, 20, 15 et 30%) selon la surface consolidée.	Selon les seuils de consolidation et les opérateurs Senbilité (Se) = 0.82 à 0.99 Spécificité (Sp) = 0.81 à 0.95	Biais de spectre – <i>Design</i> expérimental
SCORE DU WISCONSIN	(Buczinski et al., 2015a) / <i>one-gate</i>	Génisses de remplacement présevrées logées en groupe ou individuellement provenant de populations à prévalence élevée (Ollivett et al., 2011) et faible (Buczinski et al., 2014).	Analyse Bayésienne de classe latente utilisant les résultats imparfaits de l'échographie thoracique (positif si consolidation ≥ 1 cm).	Se = 0.62 (ICB95% : 0.48; 0.76) Sp = 0.74 (ICB95% : 0.65; 0.83) *ICB95% : intervalle de crédibilité à 95%	♦ Imprécisions autour des valeurs de Se et Sp ♦ Pas limité au statut à traiter
	(Love et al., 2016) / <i>two-gate</i>	Génisses de remplacement présevrées logées individuellement avec des veaux sélectionnés aléatoirement (dépistage) et des veaux suspectés malades (confirmation).	Test standard composite = 2 tests imparfaits interprétés en parallèle : ♦ Échographie thoracique (positif si consolidation extensive, focale, abcès > 2cm, effusion pleurale) ♦ Auscultation thoracique (positif si présence de bruits anormaux).	Se (Dépistage) = 0.46 (IC95% : 0.39; 0.53) Se (Confirmation) = 0.71 (IC95% : 0.64; 0.78) Sp = 0.91 (IC95% : 0.87; 0.94) *IC95% : intervalle de confiance à 95%	♦ Biais de spectre – <i>two-gate design</i> ♦ Biais de classification – test comparateur imparfait ♦ Pas limité au statut à traiter
	(Lowie et al., 2022) / <i>one-gate</i>	Utilisation de base de données de 297 génisses de remplacement et 399 veaux lourds présevrées et sevrées.	Test comparateur unique : échographie thoracique (positif si consolidation ≥ 1 et ≥ 3 cm).	Veaux laitiers présevrés Se = 0.27 (seuil ≥ 1 cm) et 0.22 (seuil ≥ 3 cm) Sp = 0.94 (seuil ≥ 1 cm) et 0.86 (seuil ≥ 3 cm) Veaux laitiers sevrés Se = 0.25 (seuil ≥ 1 cm) et 0.18 (seuil ≥ 3 cm) Sp = 0.89 (seuil ≥ 1 cm) et 0.83 (seuil ≥ 3 cm)	♦ Pas de calcul IC jugeant la précision ♦ Biais de classification – test comparateur imparfait ♦ Pas limité au statut à traiter
SCORE CALIFORNIEN	(Love et al., 2016) / <i>two-gate</i>	Génisses de remplacement présevrées logées individuellement avec des veaux sélectionnés aléatoirement (dépistage) et des veaux suspectés malades (confirmation).	Test standard composite = 2 tests imparfaits interprétés en parallèle : ♦ Échographie thoracique (positif si consolidation extensive, focale, abcès > 2cm, effusion pleurale) ♦ Auscultation thoracique (positif si présence de bruits anormaux).	Se (Dépistage) = 0.47 (IC95% : 0.40; 0.54) Se (Confirmation) = 0.73 (IC95% : 0.65; 0.79) Sp = 0.87 (IC95% : 0.83; 0.91)	♦ Biais de spectre – <i>two-gate design</i> Biais de classification – test comparateur imparfait ♦ Pas limité au statut à traiter
	(Lowie et al., 2022) / <i>one-gate</i>	Utilisation de base de données de 297 génisses de remplacement et 399 veaux lourds présevrées et sevrées.	Test comparateur unique : échographie thoracique (positif si consolidation ≥ 1 et ≥ 3 cm).	Veaux laitiers présevrés Se = 0.27 (seuil ≥ 1 cm) et 0.23 (seuil ≥ 3 cm) Sp = 0.94 (seuil ≥ 1 cm) et 0.86 (seuil ≥ 3 cm) Veaux laitiers sevrés Se = 0.33 (seuil ≥ 1 cm) et 0.20 (seuil ≥ 3 cm) Sp = 0.83 (seuil ≥ 1 cm) et 0.75 (seuil ≥ 3 cm)	♦ Pas de calcul IC jugeant la précision ♦ Biais de classification – test comparateur imparfait ♦ Pas limité au statut à traiter
SCORE CALIFORNIEN-Québec	(Buczinski et al., 2018c) / <i>one-gate</i>	Génisses de remplacement présevrées logées individuellement ou en groupe.	Analyse Bayésienne de classe latente utilisant les résultats imparfaits de l'échographie thoracique (positif si consolidation ≥ 1 cm).	Se variant de 0.83 à 0.67 pour un seuil de 9 à 13 respectivement, à définir selon la prévalence Sp variant de 0.69 à 0.83 pour un seuil de 9 à 13 respectivement, à définir selon la prévalence	♦ Pas limité au statut à traiter
SCORE DE MAIER	(Maier et al., 2019) / <i>two-gate</i>	Génisse de remplacement sevrées logées en groupe avec des veaux sélectionnés aléatoirement (dépistage) et des veaux suspectés malades (confirmation).	Test standard composite = 2 tests imparfaits interprétés en parallèle : ♦ Échographie thoracique (positif si consolidation extensive, focale, abcès > 2cm, effusion pleurale) ♦ Auscultation thoracique (positif si présence de bruits anormaux).	♦ Sans température rectale: Se (Dépistage) = 0.77 (IC95% : 0.62; 0.87) Se (Confirmation) = 1.0 Sp = 0.62 (IC95% : 0.51; 0.72) ♦ Avec température rectale: Se (Dépistage) = 0.65 (IC95% : 0.49; 0.77) Se (Confirmation) = 0.77 (IC95% : 0.43; 0.94) Sp = 0.77 (IC95% : 0.67; 0.84)	♦ Biais de spectre – <i>two-gate design</i> ♦ Biais de classification - test comparateur imparfait ♦ Imprécisions autour des valeurs de Se et Sp

Tableau 15. – Sensibilités et spécificités des scores cliniques respiratoires diagnostiques (test index) pour détecter la bronchopneumonie infectieuse (BPI) chez les veaux de race laitière.

Limites des études : Dans le **Tableau 15**, on constate que des imprécisions (larges IC95% et ICB95% autour des valeurs estimées) et des biais sont présents dans les études décrivant les performances des principaux SCRD utilisés chez les veaux de race laitière. Les performances sont donc à interpréter en conséquence. Autre constatation, la définition de la BPI utilisée dans ces études ne se limite pas au statut à traiter. Les performances des SCRD pour justifier un traitement antibiotique et/ou anti-inflammatoire sont donc inconnues.

Limites des performances du SCRD : Si l'on calcule la moyenne des bornes inférieures et supérieures des différentes performances des SCRD disponibles, la Se moyenne du test varie entre 0.54 et 0.73 et la Sp entre 0.73 et 0.88. Concrètement, si on considère une population avec 0.20 de prévalence de BPI (prévalence rapportée comme problématique par Leruste et al. [2012]), entre 10 et 30 % des veaux malades ne vont pas être diagnostiqués par le SCRD et possiblement avoir un impact négatif sur leurs productions futures (FN). Dans un même temps, entre 13 % et 30 % des veaux vont être considérés comme malades, et donc traités avec des antibiotiques alors qu'ils n'en ont pas besoin (FP). Pour limiter ces conséquences de mauvais diagnostic, une amélioration des performances du SCRD est donc nécessaire et permettra de mieux traiter les veaux de race laitière.

ii- L'auscultation thoracique

Le **Tableau 16** résume les Se et Sp de l'auscultation pulmonaire chez les veaux de race laitière.

Étude/ Design	Population de veaux	Définition du test index positif	Confirmation de la BPI (Test comparateur)	Performances	Limites
(Buczinski et al., 2014) / <i>one-gate</i>	Génisses de remplacement présevrées.	Sifflement, crépitant, absence de bruits respiratoires, frottement pleural.	Test isolé imparfait = Échographie thoracique (positif si consolidation ≥ 1 cm) utilisé comme gold standard.	Se = 0.06 variant de 0 à 0.17 selon l'aire d'auscultation Sp = 0.99 variant de 0.97 à 1.0 selon l'aire d'auscultation	♦ Biais de classification : test comparateur imparfait ♦ Pas limité au statut à traiter
(Buczinski et al., 2016b) / <i>one-gate</i>	Veaux lourds d'environ 3 semaines d'âge.	Sifflement, crépitant, absence de bruits respiratoires, frottement pleural et bruits respiratoires augmentés.	Analyse Bayésienne de classe latente utilisant les résultats de l'échographie thoracique (positif si consolidation ≥ 1 cm) comme test comparateur imparfait.	Se = 0.73 (ICB95% : 0.50; 0.96) Sp = 0.53 (ICB95% : 0.43; 0.64)	♦ Imprécision sur la Se et Sp ♦ Pas limité au statut à traiter
(Pardon et al., 2019b) / <i>one-gate</i>	3 sessions de 8-10 veaux lourds âgés de 5, 10, 13 semaines.	Laisser libre à l'opérateur.	Test isolé imparfait = Échographie thoracique (positif si consolidation ≥ 1 cm) utilisé comme gold standard.	Se = 0.63 (SD = 0.2) Sp = 0.46 (SD = 0.3)	♦ Biais de classification : test comparateur imparfait ♦ Imprécision sur la Se et Sp ♦ Pas limité au statut à traiter

BIC95% : intervalle de crédibilité à 95%; BPI : Bronchopneumonie infectieuse; SD : déviation standard; Se : sensibilité; Sp : Spécificité.

Tableau 16. – Sensibilités et spécificités de l'auscultation thoracique (test index) pour détecter la bronchopneumonie infectieuse chez les veaux de race laitière.

Limites des études : Tout comme pour le SCRD, la fiabilité des études est variable et la définition de la maladie ne se restreint pas au statut à traiter.

Limites de standardisation du test : La définition d'un test positif est variable selon les études faisant varier les performances de l'auscultation thoracique. En effet, si l'on compare les performances de l'auscultation thoracique dans les deux études de Buczinski et al. (2014) et Buczinski et al. (2016b), on constate que l'ajout des bruits augmentés a rendu l'auscultation thoracique plus sensible (passant de 0.17 à 0.73) au détriment de la Sp (passant de 1 à 0.53). Une définition de ce que représente un résultat positif est donc nécessaire pour juger des performances de ce test. Finalement, si l'on compare les performances de l'auscultation thoracique du **Tableau 16** à celles rapportées en utilisant une méthode d'auscultation plus objective comme le stéthoscope électronique développé dans les bovins de boucherie (*Whisper®*)(Se = 0.92 [ICB95%: 0.71 ; 0.99] et Sp = 0.90 [ICB95%: 0.64 ; 0.99] (Mang et al., 2015) ; **Partie III, II-2**), on constate que les performances de l'auscultation thoracique sont nettement inférieures, sans doute par manque de standardisation.

Limites des performances de l'auscultation thoracique : Prenons à titre d'exemple les performances décrites par Buczinski et al. (2016b) qui rapporte une Se de 0.73 et une Sp de 0.53. Concrètement, si on utilise uniquement l'auscultation thoracique dans une population à 0.20 de prévalence de BPI, 25 % des veaux malades ne vont pas être diagnostiqués. Cette proportion de FN se rapproche de celle du SCRD compris entre 10 et 30 %. En revanche, 46 % des veaux vont être déclarés comme malades, et donc traités avec des antibiotiques alors qu'ils n'en ont pas besoin. Cette proportion est quasiment le double de celle rapportée avec le SCRD. Les performances actuelles de l'auscultation thoracique sont donc inférieures aux performances moyennes du SCRD. Ces performances remettent en question son utilité lors du diagnostic de veau en BPI, son ajout n'améliorant pas les performances du SCRD. Son amélioration est donc primordiale pour justifier son utilisation dans des stratégies diagnostiques de BPI chez les veaux de race laitière.

iii- Les tests d'imagerie médicale

Le **Tableau 17** répertorie les Se et Sp de l'échographie thoracique pour détecter la BPI chez les veaux de race laitière.

Étude/ Design	Population de veaux	Définition d'un test index positif	Confirmation de la BPI (Test comparateur)	Performances	Limites
Rabeling et al. (1998) / <i>two-gate</i>	Génisses de remplacement sevrées atteintes de polyarthrite ou de BPI grave.	Queue de comète, Abcès, Consolidation (seuil 0 cm).	Test <i>gold standard</i> = Nécropsie	Se = 0.85 Sp = 0.98	◆ Biais de spectre, les cas étant des cas graves et chroniques de la BPI
(Jung and Bostedt, 2004) / <i>two-gate</i>	Génisses de remplacement de 0 à 14 jours (veaux nouveau- nés).	Queue de comète, abcès, effusion pleurale, pneumothorax, consolidation (> 0 cm).	Test isolé imparfait = Radiographie thoracique utilisé comme <i>gold standard</i>	Se = 0.77 Sp = 1.0	◆ Biais de classification : test comparateur imparfait ◆ Biais de spectre – <i>two- gate design</i>
(Ollivett et al., 2015) / <i>one-gate</i>	Veaux mâles présevrés sans signe clinique (score du Wisconsin < 5).	Queue de comète, abcès, effusion pleurale, pneumothorax, consolidation (> 0 cm).	Test <i>gold standard</i> = Nécropsie	Se = 0.94 (IC95% : 0.69; 1) Sp = 1.0 (IC95% : 0.64; 1)	◆ Biais de spectre incluant uniquement des veaux sous cliniques ◆ Imprécision sur les estimations de Se et Sp
(Buczinsk i et al., 2015b) / <i>one-gate</i>	Génisses de remplacement présevrées provenant de deux localisations (New York, prévalence élevée (Ollivett et al., 2011); Québec, prévalence faible (Buczinski et al., 2014).	Consolidation (≥ 1 cm)	Analyse Bayésienne de classe latente réalisée pour estimer les performances du score du Wisconsin et en utilisant des <i>priors</i> d'expert sur les performances de l'échographie thoracique	Se = 0.79 (ICB95% : 0.66; 0.91) Sp = 0.94 (ICB95% : 0.88; 0.98)	◆ Les <i>priors</i> sur les performances de l'échographie thoracique et le faible nombre de veaux ont pu influencer les résultats (n= 191) ◆ Imprécision sur l'estimation de la Se
(Buczinsk i et al., 2016b) / <i>one-gate</i>	209 veaux lourds âgés d'environ 3 semaines d'âge.	Consolidation (≥ 1 cm)	Analyse Bayésienne de classe latente réalisée pour estimer les performances de l'auscultation thoracique et en utilisant des <i>priors</i> d'expert sur les performances de l'échographie thoracique	Se = 0.77 (ICB95% : 0.60; 0.89) Sp = 0.93 (ICB95% : 0.87; 0.97)	◆ Les <i>priors</i> sur les performances de l'échographie thoracique et le faible nombre de veaux ont pu influencer les résultats (n=209) ◆ Imprécision sur l'estimation de la Se

BPI : Bronchopneumonie infectieuse; IC95% : intervalle de confiance à 95%; ICB95% : intervalle de crédibilité à 95%; Se : sensibilité; Sp : Spécificité.

Tableau 17. – Sensibilités et spécificités de l'échographie thoracique (test index) pour détecter la bronchopneumonie infectieuse chez les veaux de race laitière.

Limite des études : Tout comme pour le SCRD et l'auscultation thoracique, la fiabilité des études est variable et la définition de la maladie ne se restreint pas au statut à traiter.

Limites des performances de l'échographie thoracique : D'après les études, la Se de l'échographie thoracique varie en moyenne entre 0.71 (moyenne des bornes inférieures) et 0.88 (moyenne des bornes supérieures). La Sp varie en moyenne entre 0.87 (moyenne des bornes inférieures) et 0.99 (moyenne des bornes supérieures). Ainsi, dans une population de 0.20 de prévalence de BPI, la proportion de veaux malades non diagnostiqués varie de 10 à 30 % avec l'échographie thoracique, ce qui est similaire au SCRD et à l'auscultation thoracique. Cependant, avec l'échographie thoracique, seulement 1 à 13 % des veaux vont être déclarés comme malades,

et donc traités avec des antibiotiques alors qu'ils n'en ont pas besoin. Cette proportion de FP est donc divisée par 5 par rapport aux SCRD. Les performances de détection de l'échographie thoracique de la BPI chez un veau de race laitière sont donc supérieures aux SCRD et à l'auscultation thoracique.

Limites de connaissances sur la radiographie thoracique : Les performances de la radiographie thoracique pour détecter la BPI chez les veaux de race laitière sont inconnues. Seule l'étude de Jung et Bostedt (2004) rapporte une Se supérieure de la radiographie à l'échographie thoracique sans pour autant en estimer ces performances. De plus, cette comparaison est discutable considérant que la radiographie thoracique était le test de référence dans cette étude engendrant un biais d'incorporation. Au contraire, les résultats chez les vaches adultes (Reef et al., 1991, Braun et al., 1997), d'autres espèces animales (Partlow et al., 2017, Ward et al., 2018) ou en médecine humaine (Ye et al., 2015) sont en faveur d'une supériorité de la Se de l'échographie thoracique. La description des performances de la radiographie thoracique permettrait de comparer les deux tests et de pouvoir juger de la meilleure stratégie d'utilisation des tests d'imagerie médicale.

II-2-2- Les performances de groupe

Considérant la forte consommation d'antibiotiques que représentent les traitements de groupe, la détection de groupe de veaux à traiter ou non avec des antibiotiques est une des solutions pour raisonner et diminuer leur utilisation pour la BPI (Pardon et al., 2012a, Lava et al., 2016b). Cependant, actuellement, aucune stratégie diagnostique n'a été développée pour savoir si un groupe nécessite d'être traité ou non. Le développement d'un test de groupe, simple à utiliser et performant pour détecter la BPI permettrait d'améliorer la productivité future d'un groupe de veaux de race laitière tout en réduisant l'utilisation des antibiotiques.

II-2-3- Les performances de plusieurs tests

L'excellente Sp des tests d'imagerie médicale rapportée **Tableau 17** en font des tests de choix à utiliser en combinaison avec des tests moins spécifiques, comme le SCRD ou l'auscultation thoracique. Les avantages d'une telle stratégie ont été chiffrés par Buczinski et al. (2016b) montrant que de rajouter l'échographie thoracique après l'auscultation thoracique augmentait considérablement les performances diagnostiques de BPI (**Annexe 10**). Les données sont manquantes concernant la combinaison d'autres tests comme SCRD + Échographie thoracique, SCRD + Radiographie thoracique, etc.

En milieu hospitalier, la radiographie thoracique est souvent utilisée en complément de l'échographie thoracique (Fowler et al., 2017). Cependant, en l'absence de connaissance des performances de la radiographie thoracique pour détecter la BPI, l'intérêt de cette combinaison n'est, à l'heure actuelle, pas avéré. Des études visant à comparer ces deux tests d'imagerie médicale sont nécessaires afin de juger de la meilleure stratégie d'utilisation : l'un ou l'autre, les deux en série ou les deux en parallèle. La détermination des performances de la radiographie thoracique par rapport à l'échographie thoracique permettrait également de juger de la pertinence de développer son utilisation sur le terrain.

III- Les limites de répétabilité et reproductibilité des tests pour détecter la bronchopneumonie infectieuse chez les veaux de race laitière

La répétabilité d'un test sur un individu se définit comme sa faculté à donner un résultat similaire entre différents opérateurs (répétabilité interopérateurs) et entre différentes mesures répétées dans un court intervalle de temps (répétabilité intraopérateurs). La reproductibilité d'un test correspond à sa répétabilité dans différentes populations et correspond à la validation externe de la répétabilité.

III-1- Procédure générale d'évaluation de la répétabilité d'un test diagnostique

III-1-1- Les indicateurs de répétabilité d'un test

Pour évaluer la répétabilité d'un test, un certain nombre d'opérateurs applique le test sur un certain nombre d'individus (n). Les résultats de chaque opérateur sont ensuite comparés pour déterminer la répétabilité interopérateurs (Kottner et al., 2011). À titre d'exemple, pour deux opérateurs, lorsque le test est dichotomique (résultat soit positif ou négatif), on obtient la table 2x2 suivante, aussi appelée matrice d'accord :

Opérateur A	Opérateur B		
	Test +	Test -	
Test +	n^{++}	n^{+-}	n^+_A
Test -	n^{-+}	n^{--}	n^-_A
	n^+_B	n^-_B	n

n^{++} : nombre d'individu positif identifié par les deux opérateurs; n^{--} : nombre d'individu négatif identifié par les deux opérateurs; n^{-+} : nombre d'individu identifié négatif par l'opérateur A mais positif par l'opérateur B; n^{+-} : nombre d'individu identifié positif par l'opérateur A mais négatif par l'opérateur B; n^+_A : nombre d'individu positif identifié par l'opérateur A; n^-_A : nombre d'individu négatif identifié par l'opérateur A; n^+_B : nombre d'individu positif identifié par l'opérateur B; n^-_B : nombre d'individu négatif identifié par l'opérateur B.

Pour évaluer la répétabilité intraopérateurs d'un test dichotomique, le test est répété identiquement par le même opérateur sur les mêmes individus, mais à différents intervalles de temps (Kottner et al., 2011). Une matrice d'accord similaire peut alors être construite entre un opérateur A à temps 1 et à temps 2.

À partir de cette matrice d'accord, différents indicateurs peuvent être estimés pour calculer l'accord inter et intraopérateurs. Le **Tableau 18** répertorie les indicateurs couramment utilisés pour quantifier la répétabilité des tests catégoriques.

Parmi eux, le pourcentage d'accord (**Pa**) et le kappa de Cohen (**κ**) sont des indicateurs de répétabilité interopérateurs et intraopérateurs. Si le nombre d'opérateurs est > 2 , le Pa médian et le κ -médian peuvent être calculés. Certains indicateurs ont, quant à eux, l'avantage de pouvoir calculer directement l'accord entre plusieurs opérateurs (≥ 2) (p.ex. ; coefficient de Fleiss, α — Krippendorff, Gwet AC1) (Gwet, 2008). Lorsque la classification est ordinale (c.à.d. ; 1,2,3,4), les indicateurs sont pondérés afin de donner des poids relatifs à l'importance des désaccords. Prenons l'exemple d'un score comprenant 4 catégories de gravité de normal (1) à grave (4), il est plus facile d'être d'accord sur les différences entre 1 et 4 qu'entre 1 et 2, 2 et 3, ou 3 et 4. Des poids différents doivent donc être donnés à chaque catégorie afin de souligner cette différence d'accord (Gwet, 2014).

Comme l'illustre le **Tableau 18**, les indicateurs sont nombreux et calculés différemment. Il n'existe pas un indicateur parfait et l'expression de plusieurs indicateurs est conseillée pour estimer au mieux la répétabilité d'un test (Morris et al., 2008, O'Leary et al., 2014, Walsh et al., 2014).

Indicateur	Calcul	Signification	Avantages	Inconvénients
Pourcentage d'accord (Pa)	Nombre d'individu positif et négatif classé par deux opérateurs ($n^{++} + n^{-}$) / nombre total (n)	Pourcentage d'accord entre deux opérateurs.	<ul style="list-style-type: none"> ◆ Calcul simple des accords entre deux opérateurs ◆ Indicateur intra et interopérateurs 	<ul style="list-style-type: none"> ◆ Ne prend pas en compte la chance ◆ Restreint à deux opérateurs mais le pourcentage d'accord médian peut être rapporté en cas d'un nombre d'opérateur > 2
Kappa de Cohen (κ)	$\kappa = (Pa - Pe) / (1 - Pe)$ Pa = le pourcentage d'accord Pe = probabilité que l'opérateur classe un individu positif ($n^+/n * n^+/n$) + la probabilité qu'un opérateur classe un individu négatif ($n^-/n * n^-/n$).	Estimation du pourcentage d'accord au-delà de la chance.	<ul style="list-style-type: none"> ◆ Prend en compte la chance ◆ Indicateur le plus familier en médecine vétérinaire donc universellement interprétable ◆ Indicateur intra et interopérateurs 	<ul style="list-style-type: none"> ◆ Sujet à une imprécision de calcul (paradoxe) dans certaines conditions : (1) si les opérateurs ont plus de chance de classer un individu dans une catégorie; (2) si les nombres d'individu positif ou négatif par un opérateur (nombre marginal = n^+_A; n^+_B; n^-_A; n^-_B) sont déséquilibrés ◆ Restreint à deux opérateurs mais le Kappa de Cohen médian peut aussi être rapporté en cas d'un nombre d'opérateur > 2
Kappa ajusté à la prévalence et aux biais = PABA κ (identique à l'indicateur de Brennan-Prediger)	$PABA\kappa = (Pa - 1/q) / (1 - 1/q)$ Pa = Pourcentage d'accord corrigé par la chance = $1/q$ (q = nombre de catégorie), ce qui signifie que l'individu sera classé pour les opérateurs dans n'importe quel catégorie avec une probabilité de $1/q$	Un ajustement du kappa de Cohen au paradoxe.	Résistant au paradoxe ce qui permet d'améliorer l'exactitude du Kappa de Cohen	<ul style="list-style-type: none"> ◆ Restreint à deux opérateurs mais PABAκ médian peut être rapporté en cas d'un nombre d'opérateur > 2 ◆ Le pourcentage d'accord corrigé par la chance est calculé comme si les individus étaient classés aléatoirement, ce qui n'est pas le cas en réalité et ce qui peut amener à une correction trop importante de la chance
α - Krippendorff (α_K)	$\alpha_K = (Pa_K - Pe_K) / (1 - Pe_K)$ Pa $_K$ = le pourcentage d'accord ajusté aux données manquantes $(1 - \epsilon_n)Pa + \epsilon_n$ avec $\epsilon_n = 1/2n$ Pe $_K$ = la probabilité qu'un individu soit classé dans la même catégorie (positif ou négatif) par les deux opérateurs $((n^+_A/n + n^+_B/n)/2)^2 + ((n^-_A/n + n^-_B/n)/2)^2$ ou plusieurs	Estimation du pourcentage d'accord au-delà de la chance et prenant en compte les individus non évalués par les deux opérateurs.	<ul style="list-style-type: none"> ◆ Prend en compte les données manquantes (c.à.d ; les individus pas nécessairement évalués par tous les opérateurs) ◆ Utilisé pour calculer l'accord entre plusieurs opérateurs 	<ul style="list-style-type: none"> ◆ Sujet au paradoxe du Kappa ◆ Manque de popularité en médecine vétérinaire donc plus difficile à comprendre par les lecteurs
Gwet AC1	$AC1 = (Pa - Pe_{AC1}) / (1 - Pe_{AC1})$ Pa = le pourcentage d'accord Pe $_{AC1}$ = la probabilité qu'un individu soit classé dans une catégorie (positif ou négatif) par un opérateur au hasard soit $2 * ((n^+_A/n + n^+_B/n)/2) * ((n^-_A/n + n^-_B/n)/2)$. Le calcul est réalisable pour plusieurs opérateurs	Estimation du pourcentage d'accord au-delà de la chance.	<ul style="list-style-type: none"> ◆ Résistant au paradoxe ◆ Utilisé pour calculer l'accord entre plusieurs opérateurs 	<ul style="list-style-type: none"> ◆ Manque de popularité en médecine vétérinaire donc plus difficile à comprendre par les lecteurs
Fleiss	$F = (Pa - Pe_F) / (1 - Pe_F)$ Pa = le pourcentage d'accord Pe $_F$ = la probabilité qu'un individu soit classé dans la même catégorie (positif ou négatif) par les deux opérateurs $((n^+_A/n + n^+_B/n)/2)^2 + ((n^-_A/n + n^-_B/n)/2)^2$ ou plusieurs	Estimation du pourcentage d'accord au-delà de la chance pour plusieurs opérateurs.	<ul style="list-style-type: none"> ◆ Utilisé pour calculer l'accord entre plusieurs opérateurs 	<ul style="list-style-type: none"> ◆ Sujet au paradoxe du Kappa ◆ Uniquement utiliser pour calculer l'accord entre plusieurs opérateurs ◆ Manque de popularité en médecine vétérinaire donc plus difficile à comprendre par les lecteurs

Tableau 18. – Principaux indicateurs utilisés pour quantifier la répétabilité inter et intra opérateur d'un test diagnostique catégorique.

III-1-2- L'interprétation des indicateurs de répétabilité

La principale problématique des indicateurs est la difficulté de traduire cliniquement la valeur trouvée (Hilden, 2014). Afin de faciliter leur interprétation, des échelles d'interprétation ont été décrites par différents auteurs (**Tableau 19**). Bien que nombreuses, ces échelles n'ont pas été validées et le choix d'une échelle par rapport à l'autre est réalisé de manière subjective selon les objectifs de l'analyse des auteurs. Par exemple, si on veut juste savoir si l'accord est excellent ou pauvre on utilise l'échelle de Fleiss et al. (1981), alors que si on veut une catégorisation plus raffinée on va opter pour l'échelle de Landis et Koch (1977). Finalement, certains auteurs vont encore plus loin dans la simplification en utilisant subjectivement un seuil ($P_a \geq 0.75$ et $\kappa \geq 0.4$) jugé cliniquement utile (Burn et Weir, 2011, Schlageter-Tello et al., 2014). Aucun de ces seuils et échelles d'interprétation n'a été validé objectivement permettant de prouver leur pertinence d'utilisation.

Nom	Échelle d'interprétation		Avantage	Inconvénient
(Landis et Koch, 1977)	< 0 0 à 0.2 0.21 à 0.40 0.41 à 0.60 0.61 à 0.80 0.81 à 1	Pauvre Léger Juste Modéré Substantiel Presque parfait	Simple à utiliser pour interpréter les indicateurs	◆ Développé pour le kappa de Cohen ◆ Subjectif
(Fleiss et al., 1981)	< 0.40 0.40 à 0.75 ≥ 0.75	Pauvre Intermédiaire à bon Excellent	Version simplifiée de Landis et Koch (1977)	◆ Développé pour le kappa de Cohen ◆ Subjectif
(Altman, 1991)	< 0.2 0.21 à 0.40 0.41 à 0.60 0.61 à 0.80 0.81 à 1	Pauvre Juste Modéré Bon Très bon		◆ Développé pour le kappa de Cohen ◆ Très similaire à Landis et Koch (1977) ◆ Subjectif
(Gwet, 2008)	Calcul de la probabilité cumulative et attribution du qualificatif des catégories d'une des échelles précédentes si 95% excède ces catégories		Permet de prendre en compte l'imprécision dans l'estimation de l'indicateur (nombre insuffisant d'opérateur ou d'individu)	◆ Subjectif

Tableau 19. – Échelles d'interprétation des indicateurs de répétabilité intra et inter opérateurs.

III-1-3- Les facteurs affectant l'estimation des indicateurs de répétabilité :

i- Les facteurs affectant la précision

La précision des indicateurs de répétabilité d'un test dépend : (1) du nombre d'individus, (2) du nombre d'opérateurs et (3) du nombre de catégories d'évaluation (Gwet, 2008, Kottner et al., 2011)

. Ces paramètres doivent être détaillés afin de juger de la précision des indicateurs. Un calcul de puissance peut être réalisé en considérant ces paramètres. Même si certains auteurs critiquent la pertinence de calculer un IC95 % (Hilden, 2014), celui-ci permet de visualiser cette imprécision autour de la valeur estimée. L'échelle de Gwet (2008) est un outil pour prendre en compte la précision des indicateurs dans leurs interprétations cliniques.

ii- Les facteurs affectant l'exactitude

Le **paradoxe du Kappa** est le principal biais pouvant affecter la répétabilité. Ce biais se traduit par une valeur de l'indicateur corrigée pour la chance anormalement basse par rapport à la valeur du Pa, ne prenant pas en compte la chance. Ce biais intervient dans deux conditions : (1) si les opérateurs ont plus de chance de classer un individu dans une catégorie, comme lors de forte prévalence d'une maladie dans une population où les opérateurs auront plus tendance à classer les individus comme positifs ; (2) si le nombre d'individus positif ou négatif par un opérateur (nombre marginal = n^+_A ; n^+_B ; n^-_A ; n^-_B) sont déséquilibrés (biais des opérateurs). Ces biais peuvent être quantifiés *via* des outils, comme, par exemple l'index de prévalence dans l'étude de Buczinski et al. (2016a) (défini comme la différence des individus positifs entre les catégories d'accord d'opérateur, divisée par le nombre total d'accords possibles concluant que plus la valeur est proche de 1, plus la matrice d'accord est déséquilibrée), ou encore par la réalisation d'un test statistique comme le test McNemar pour les données appariées, pour savoir si deux opérateurs classent approximativement la même proportion d'individus positifs (Dohoo et al., 2012). Les autres moyens de limiter ce paradoxe sont d'ajuster la prévalence de la maladie à 0.5 afin de limiter son influence (Dohoo et al., 2012) ou d'utiliser des indicateurs paradoxe-résistants (p. ex. ; PABAκ; Gwet AC1) (Gwet, 2008). L'inconvénient de ces deux dernières méthodes est que l'extrapolation à des conditions où la prévalence est différente devient discutable.

iii- Les facteurs affectant l'interprétation des indicateurs de répétabilité

Différents facteurs peuvent influencer la répétabilité d'un test comme : (1) les caractéristiques des individus testés (p. ex. ; âge, sexe, race, etc.) ; (2) les caractéristiques des opérateurs (p. ex. ; l'expérience, l'entraînement) ; (3) les conditions d'évaluations (p. ex. ; indépendant entre les opérateurs, les conditions de mesures, etc.) ; ou encore (4) l'intervalle de temps entre deux mesures pour les indicateurs intraopérateurs. Tous ces facteurs doivent être mentionnés afin de juger de l'applicabilité de la répétabilité estimée. L'utilisation des lignes directrices du **GRAAS** (*Guidelines*

for Reporting Reliability and Agreement Studies) permet d'avoir toutes ces informations (Kottner et al., 2011). La validation externe de répétabilité dans d'autres populations permet de juger de la reproductibilité du test.

III-2- La répétabilité et la reproductibilité des tests diagnostiques pour détecter la BPI chez les veaux de race laitière

Le **Tableau 20** répertorie les accords inter et intra opérateurs du SCRD, de l'auscultation thoracique, de l'échographie thoracique et de la radiographie thoracique chez les veaux de race laitière pour détecter la BPI. Aucune étude ne rapporte la reproductibilité d'un de ces tests. Les limites de précision, l'exactitude ou l'absence d'information concernant les facteurs influençant les paramètres de répétabilité sont mentionnées dans la dernière colonne.

III-2-1- Les scores cliniques respiratoires diagnostiques

D'après le **Tableau 20**, on constate que la répétabilité des SCRD a été estimée uniquement chez les médecins vétérinaires alors que c'est un outil qui est utilisé par les producteurs. De plus, cette répétabilité est faible. Cliniquement, cela signifie que deux médecins vétérinaires différents ne classifient pas un veau malade de la même manière. À l'échelle d'un groupe de veau, si un médecin utilise l'un de ces SCRD pour décider de traiter un veau, cela signifie que deux lots ou groupes de mêmes prévalences peuvent avoir une proportion différente de traitement. Inversement, la proportion de veaux non diagnostiqués et impactant la productivité future sera également différente. Si on veut diminuer ces impacts, il est essentiel de fournir aux médecins vétérinaires, mais surtout aux producteurs, une méthode diagnostique répétable et reproductible.

III-2-2- L'auscultation thoracique

Tout comme les SCRD, l'auscultation thoracique n'est actuellement pas un test répétable (**Tableau 20**), c.-à-d. ; que son résultat va dépendre de l'opérateur qu'il l'utilise. Ainsi, la proportion de veaux diagnostiqués par l'auscultation thoracique sera variable d'un médecin vétérinaire à un autre. L'amélioration de la standardisation de ce test permettrait d'améliorer sa répétabilité et sa reproductibilité.

III-2-3- Les tests d'imagerie médicale

Grâce aux travaux de simplification de la technique et d'interprétation (détection des lésions de consolidation pulmonaire et calcul de leurs profondeurs), l'échographie thoracique est actuellement

un test répétable par différents opérateurs d'expérience variable (**Tableau 20**). Cette particularité rend son utilisation avantageuse dans les stratégies diagnostiques pour améliorer la détection de la BPI chez les veaux de race laitière.

En ce qui concerne la radiographie thoracique, la répétabilité a été étudiée dans une seule étude impliquant uniquement des médecins vétérinaires expérimentés (Shimbo et al., 2018c). Des données sont manquantes sur son utilisation par des médecins vétérinaires non expérimentés pour juger de sa répétabilité sur le terrain.

Conclusion PARTIE IV : Les tests diagnostiques actuels incluant les SCRD, l'auscultation thoracique, l'échographie et la radiographie thoracique, présentent des limites de développement, de performance et de répétabilité. Une amélioration de ces caractéristiques permettrait de mieux détecter les veaux de race laitière nécessitant un traitement pour la BPI.

Études	Individus	Opérateurs	Évaluation	Indicateurs	Interprétation		Limites	
					Valeur de l'indicateur (intervalle de confiance à 95%)			
SCRD								
Score CIS (Amrine et al., 2013)	178 génisses de remplacement sevrées expérimentalement inoculés par <i>Mycoplasma bovis</i> .	♦ 9 opérateurs (médecins vétérinaires) d'expériences variables ♦ Évaluation le même jour.	Binaire – CIS positif si malade (≥ 1), négatif si non malade ($=0$)	Kappa de Cohen médian (κ)	$\kappa = 0.16$ (0.10; 0.24) Échelle de Landis et Koch (1977) : Accord pauvre à léger		♦ Utilisation d'un seul indicateur ♦ Pas d'ajustement au paradoxe du Kappa ♦ Restreint aux médecins vétérinaires	
Score du Wisconsin (Buczinski et al., 2016a)	Veaux lourds présevrés en trois visites (V) : V1 : 10 veaux aléatoires à l'arrivée en parc engraissement ; V2 : 10 veaux aléatoires à 12 jours d'engraissement ; V3 : 20 veaux à 21 jours d'engraissement ; 10 veaux avec des lésions échographiques de consolidation pulmonaire et 10 veaux aléatoires.	♦ 2 (V1 et V2) ou 3 opérateurs (V3) (un médecin vétérinaire et deux étudiants vétérinaires), tous peu expérimentés ♦ Évaluation le même jour à 4 heures d'intervalle.	Score du Wisconsin, ordinal pour chaque signe clinique de 0 à 3 et considéré malade selon différents seuils (≥ 4 , ≥ 5 , et ≥ 6)	♦ Pourcentage d'accord médian (Pa) ♦ Kappa de Cohen pondéré médian ($w\kappa$) ♦ Fleiss (F) pour opérateurs multiples	Échelle de Landis et Koch (1977)		♦ Pas de moyen de juger de la précision autour de la valeur de l'indicateur ♦ Biais du paradoxe du Kappa exprimé statistiquement par l'index de prévalence et jugé important ♦ Restreint aux médecins vétérinaires	
					Signes cliniques : accord pauvre à modéré			
					Toux	Pa=0.78		$w\kappa = 0.1$
					Température	Pa = 0.63		$w\kappa = 0.6$
Jetage nasal	Pa = 0.62	$w\kappa = 0.24$						
Jetage oculaire	Pa=0.63	$w\kappa = 0.11$						
Oreilles	Pa=0.85	$w\kappa = -0.04$						
Les différents seuils : accord pauvre à juste								
≥ 4 : léger	F = 0.35							
≥ 5 : pauvre	F = 0.06							
≥ 6 : pauvre	F = 0.13							
Score Californien (Aly et al., 2014)	100 génisses de remplacement présevrées suspectées malades par auscultation thoracique et/ou échographie thoracique.	Un seul opérateur (médecin vétérinaire) expérimenté qui réalise le score Californien et le score du Wisconsin.	Accord comparé avec le score du Wisconsin	Kappa de Cohen (κ)	$\kappa = 0.85$ (Déviation standard : 0.099) Échelle de Fleiss et al. (1981) : Accord presque parfait entre le score Californien et Wisconsin		♦ Score Californien développé à partir du score du Wisconsin donc accord anticipé ♦ Pas de calcul de la répétabilité inter opérateur du score Californien ♦ Un seul opérateur médecin vétérinaire	
Auscultation thoracique								
(Pardon et al., 2019b)	3 sessions de 8 à 10 veaux lourds logés en groupe avec une historique de traitement pour BPI dans le groupe.	49 opérateurs (médecins vétérinaires) d'expériences variables.	Aucun critère objectif définis. Les opérateurs devaient dire subjectivement si le veau avait une auscultation thoracique compatible avec la présence d'une BPI.	α -Kripendorff car présence de données manquantes	$\alpha\text{-}\kappa = 0.18$ (-0.01; 0.38) Échelle non précisée mais accord jugé pauvre		♦ Utilisation d'un seul indicateur et imprécision autour de son estimation ♦ Pas d'ajustement au paradoxe du Kappa ♦ Opérateurs pas à l'aveugle (possibilité d'être influencés par les autres opérateurs ou de voir les signes cliniques)	

Études	Individus	Opérateurs	Évaluation	Indicateurs	Interprétation Valeur de l'indicateur (intervalle de confiance à 95%)	Limites										
Échographie thoracique																
(Buczinski et al., 2013)	10 veaux lourds présevrés, 6 traités par le producteur, 4 supposés sains, sans signes cliniques apparents.	3 opérateurs (médecins vétérinaires) d'expérience variable.	Consolidation pulmonaire (profondeur \geq 1cm)	Kappa de Cohen (κ)	κ variant 0.6 à 1.0 selon les opérateurs Échelle de Landis et Koch (1977) : Accord substantiel à presque parfait	<ul style="list-style-type: none"> ◆ Utilisation d'un seul indicateur ◆ Faible nombre d'opérateurs et d'individus 										
(Buczinski et al., 2018b)	50 vidéos échographiques de lésions thoraciques sélectionnées au hasard parmi 210 bouvillons de boucherie malades et 107 non malades selon leur signes cliniques.	6 opérateurs (médecins vétérinaires) d'expériences variables ayant reçu une formation de 30 minutes pour évaluer les vidéos.	Caractériser la lésions (queue de comète, irrégularité de la plèvre, consolidation pulmonaire, effusion pleurale et lésions cavitaires).	<ul style="list-style-type: none"> ◆ Kappa de Cohen médian (κ) ◆ Pourcentage d'accord médian (Pa) ◆ Gwet (AC1) pour opérateurs multiples ◆ Fleiss (F) pour opérateurs multiples 	Échelle de Landis and Koch (1977) <table border="1"> <tr> <td>Consolidation Pulmonaire</td> <td>Pa médian = 0.84 (0.80;0.89) κ médian = 0.68 (0.60; 0.78) F = 0.67 (0.49; 0.86) AC1 = 0.68 (0.51; 0.86)</td> </tr> <tr> <td>Queue de comète</td> <td>Pa médian = 0.82 (0.80;0.87) κ médian = 0.54 (0.50; 0.66) F = 0.56 (0.33; 0.80) AC1 = 0.73 (0.58; 0.89)</td> </tr> <tr> <td>Irrégularité de la plèvre</td> <td>Pa médian = 0.62 (0.53;0.67) κ médian = 0.25 (0.08; 0.33) F = 0.20 (-0.05; 0.44) AC1 = 0.21 (-0.01; 0.44)</td> </tr> <tr> <td>Effusion pleurale</td> <td>Pa médian = 0.82 (0.75;0.86) κ médian = 0.34 (0.21; 0.48) F = 0.36 (0.10; 0.61) AC1 = 0.71 (0.51; 0.92)</td> </tr> <tr> <td>Lésions cavitaires</td> <td>Pa médian = 0.94 (0.84;1.0) κ médian = 0.2; 0.56; 0.23 pour 3 paires d'opérateurs</td> </tr> </table>	Consolidation Pulmonaire	Pa médian = 0.84 (0.80;0.89) κ médian = 0.68 (0.60; 0.78) F = 0.67 (0.49; 0.86) AC1 = 0.68 (0.51; 0.86)	Queue de comète	Pa médian = 0.82 (0.80;0.87) κ médian = 0.54 (0.50; 0.66) F = 0.56 (0.33; 0.80) AC1 = 0.73 (0.58; 0.89)	Irrégularité de la plèvre	Pa médian = 0.62 (0.53;0.67) κ médian = 0.25 (0.08; 0.33) F = 0.20 (-0.05; 0.44) AC1 = 0.21 (-0.01; 0.44)	Effusion pleurale	Pa médian = 0.82 (0.75;0.86) κ médian = 0.34 (0.21; 0.48) F = 0.36 (0.10; 0.61) AC1 = 0.71 (0.51; 0.92)	Lésions cavitaires	Pa médian = 0.94 (0.84;1.0) κ médian = 0.2; 0.56; 0.23 pour 3 paires d'opérateurs	Accord déterminé à partir de vidéos sans possibilité de s'ajuster à la position de la sonde ou de reprendre les images si nécessaire
Consolidation Pulmonaire	Pa médian = 0.84 (0.80;0.89) κ médian = 0.68 (0.60; 0.78) F = 0.67 (0.49; 0.86) AC1 = 0.68 (0.51; 0.86)															
Queue de comète	Pa médian = 0.82 (0.80;0.87) κ médian = 0.54 (0.50; 0.66) F = 0.56 (0.33; 0.80) AC1 = 0.73 (0.58; 0.89)															
Irrégularité de la plèvre	Pa médian = 0.62 (0.53;0.67) κ médian = 0.25 (0.08; 0.33) F = 0.20 (-0.05; 0.44) AC1 = 0.21 (-0.01; 0.44)															
Effusion pleurale	Pa médian = 0.82 (0.75;0.86) κ médian = 0.34 (0.21; 0.48) F = 0.36 (0.10; 0.61) AC1 = 0.71 (0.51; 0.92)															
Lésions cavitaires	Pa médian = 0.94 (0.84;1.0) κ médian = 0.2; 0.56; 0.23 pour 3 paires d'opérateurs															
Radiographie thoracique																
(Shimbo et al., 2018c)	Clichés radiographiques de 19 génisses de remplacement présevrées et sevrées, référées dans un centre hospitalier.	7 opérateurs (médecins vétérinaires) expérimentés.	Score interprétation des clichés radiographiques 1-Définitivement absent 2-Probablement absent 3-Incertain 4-Probablement présent 5-Définitivement présent	Coefficient intra classe	ICC (coefficient de corrélation intra-classe) = 0.66 (0.54; 0.77) <i>Il s'agit de la mesure de l'homogénéité des observations au sein des classes d'un effet aléatoire par rapport à la dispersion de ces observations entre classes.</i> Échelle de Landis et Koch (1977) : Accord substantiel	<ul style="list-style-type: none"> ◆ Aucun renseignement sur les opérateurs à part qu'ils sont expérimentés ◆ Un seul indicateur, habituellement utilisé pour les données quantitatives 										

Tableau 20. – Répétabilité des tests diagnostiques couramment utilisés pour détecter la bronchopneumonie infectieuse (BPI) chez les veaux de race laitière.

PARTIE V : Conclusion et objectif général

Cette revue de littérature a permis de souligner les principales problématiques actuelles de la BPI chez les veaux de race laitière, à savoir :

- 1- **L'inexactitude de la définition de la BPI** pour détecter les veaux à traiter ;
- 2- **Le manque de développement et de standardisation des tests** couramment utilisés pour détecter la maladie ;
- 3- **Le manque de répétabilité entre les opérateurs et les performances sous-optimales** des tests couramment utilisés pour détecter et traiter la maladie.

L'**objectif général** de cette thèse était donc de pallier ces trois points majeurs et d'améliorer les stratégies de détection de la BPI des veaux de race laitière nécessitant un traitement afin de mieux diagnostiquer les veaux malades et, ultimement, diminuer les impacts économiques de la maladie, améliorer le bien-être des veaux et favoriser l'emploi judicieux des antibiotiques en traitant uniquement les veaux ou groupe de veaux qui en ont besoin.

Pour répondre à cet objectif général, trois sous-objectifs spécifiques ont donc été définis :

Sous-objectif 1 : Développement d'un **SCRD**, en identifiant, et validant les signes cliniques associés spécifiquement au statut à traiter de la BPI, ayant une bonne répétabilité interopérateurs, pour diagnostiquer la maladie à l'échelle individuelle et populationnelle (CHAPITRE 2).

Hypothèse 1 : Certains signes cliniques sont plus spécifiques de BPI et plus répétables que d'autres et pourraient constituer des marqueurs diagnostiques dans un SCRd.

Sous-objectif 2 : Amélioration des stratégies diagnostiques des tests d'imagerie médicale, en comparant **l'échographie et la radiographie thoracique** pour détecter les lésions et le statut actif, c. à d. ; à traiter, de la BPI (CHAPITRE 3).

Hypothèse 2 : L'échographie thoracique est suffisante pour diagnostiquer la BPI chez les veaux de race laitière.

Sous-objectif 3 : Amélioration de **l'échographie thoracique** en estimant les performances de l'échographie thoracique pour détecter le statut actif, c. à d. ; à traiter, de la BPI et en standardisant la technique (ajout ou non du lobe crânial droit) et l'interprétation (seuil de profondeur) afin d'en optimiser ses performances (CHAPITRE 4).

Hypothèse 3 : La Se et la Sp de l'échographie thoracique est adéquate pour diagnostiquer les BPI chez les veaux de race laitière.

Le **Tableau 21** détaille la composition de chacun de ces chapitres. Nous invitons le lecteur à s'y référer afin de connaître la contribution de chaque auteur dans l'élaboration de chacun des articles qui vont être présentés dans les prochaines pages.

Chapitre	Article	Auteurs	Contributions
CHAPITRE 2	ARTICLE 1 ARTICLE 2	Julie Berman	Élaboration du projet; récolte et analyse des données et rédaction de l'article
		David Francoz	Élaboration du projet et relecture de l'article
		Abdelmonem Mohamed Abdallah	Récolte des données et relecture de l'article
		Simon Dufour	Élaboration du projet et relecture de l'article
		Sébastien Buczinski	Élaboration du projet, récolte des données et relecture de l'article
CHAPITRE 3	ARTICLE 3 ARTICLE 4	Julie Berman	Élaboration du projet, récolte et analyse des données et rédaction de l'article
		Isabelle Masseau	Élaboration du projet, récolte des données et relecture de l'article
		Gilles Fecteau	Élaboration du projet, récolte des données et relecture de l'article
		Sébastien Buczinski	Élaboration du projet, récolte des données et relecture de l'article
		David Francoz	Élaboration du projet, récolte des données et relecture de l'article
CHAPITRE 4	ARTICLE 5	Julie Berman	Élaboration du projet, récolte et analyse des données et rédaction de l'article
		David Francoz	Élaboration du projet et relecture de l'article
		Simon Dufour	Élaboration du projet et relecture de l'article
		Sébastien Buczinski	Élaboration du projet, récolte des données et relecture de l'article

Tableau 21. – Répartition des articles scientifiques constituant cette thèse ainsi que les différents auteurs et leurs contributions respectives.

Chapitre 2 – Développement d’un score clinique respiratoire diagnostique pour détecter la bronchopneumonie infectieuse chez les veaux lourds

ARTICLE 1

Evaluation of inter-rater agreement of the clinical signs used to diagnose bovine respiratory disease in individually housed veal calves

Publié dans le *Journal of Dairy Science*. 2021 Nov 1;104(11):12053-65

J. Berman,^{1*} D. Francoz,¹ A. Abdallah,^{1,2} S. Dufour,³ and S. Buczinski.¹

¹ From the Département des sciences cliniques, Faculté de médecine vétérinaire, Université de Montréal, Saint-Hyacinthe, QC, Canada.

² From the Department of Animal Medicine, Faculty of Veterinary Medicine, Zagazig University, Zagazig, Egypt.

³ From the Département de Pathologie et Microbiologie, Faculté de médecine vétérinaire, Université de Montréal, St-Hyacinthe, QC, Canada.

Abbreviation: AC1 = Gwet’s agreement coefficient type 1; BRD = Bovine Respiratory Disease; CA = California score; CI = Confidence Intervals; GRRAS = Guidelines for Reporting Reliability and Agreement Studies; IMP = Interval Membership Probability; κ = Cohen’s kappa; Pa = Raw percentage of agreement; Pc = Percentage of agreement corrected for agreement due to chance; WI = Wisconsin scoring system chart; $w\kappa$ = weighted kappa.

ABSTRACT

Objectives: In dairy calves raised for veal, typical clinical signs of bovine respiratory disease (BRD) are ocular discharge, nasal discharge, ear droop/head tilt, abnormal respiration, cough, and increased rectal temperature. Despite the existence of several clinical scoring systems, the variability of human recognition of individual BRD clinical signs has been uncommonly studied. The objective of this study was therefore to assess the inter-rater agreement of BRD clinical signs in veal calves. We hypothesized that BRD clinical signs were not detected equally between

veterinarians, technicians, and producers of the veal industry and that some clinical signs have higher inter-rater agreement than others.

Methods: During 2017-2018, we prospectively recorded 524 videos of physical examinations of random veal calves from 48 different batches in Quebec, Canada. A researcher, not involved in the inter-rater assessment, classified each video as presence/absence of each BRD clinical sign except rectal temperature. For each of the 5 clinical signs, 15 videos with and 15 videos without the clinical signs were randomly selected to avoid kappa paradoxes. Those 30 videos were then presented in a random order to experienced raters of BRD in veal calves: 6 veterinarians, 6 technicians, and 6 producers. The raters assessed the clinical signs using a modified Wisconsin and California scoring system (0=absent, 1=mild, 2=moderate, 3=severe for nasal discharge, ocular discharge, and ear droop/head tilt; and 0=absent, 1=moderate, 2=severe for abnormal respiration and induced cough). We used median percentage agreement (Pa), median Cohen's kappa (κ), and Gwet's agreement coefficient 1 (AC1) to assess inter-rater agreement. The effect of scale combination was also tested to determine the optimal combination (4-scale 0/1/2/3 vs 3-scale 0/1/2 vs 2-scale 0/1,2,3; 0,1/2,3; or 0/1,2). The differences of inter-rater agreement between veterinarians, technicians, and producers were estimated by a Wilcoxon rank-sum test.

Results: The 2-scale combination (0,1/2,3 or 0/1,2) had the highest inter-rater agreement for all clinical signs. With this combination, induced cough was the clinical sign with the highest inter-rater agreement (Pa=0.93; κ =0.79; AC1=0.87) and abnormal respiration was the sign with the lowest inter-rater agreement (Pa=0.77; κ =0.20; AC1=0.74). According to Pa and AC1 values, the 2-scale inter-rater agreement of the 5 clinical signs was good (value >0.6). According to κ , only ear droop/head tilt and induced cough had a substantial 2-scale inter-rater agreement (κ >0.6). In general, the 2-scale inter-rater agreement was better among veterinarians than among technicians and producers, except for the ear droop/head tilt, where agreement was better among producers.

Conclusion: We concluded that with severity scores assessed on a scale of 2 (0,1/2,3 or 0/1,2), the inter-rater agreement of BRD clinical signs was variable according to the sign but clinically acceptable in veal calves. BRD clinical signs were, however, not detected equally between veterinarians, technicians, and producers of the veal industry. Future research could determine if this discrepancy could be improved by standardization training.

Key words: reliability; inter-operator; pneumonia; kappa; cattle.

I- INTRODUCTION

Veal production is a major outlet for dairy or crossbreed calves from the dairy industry (Sans et Fontguyon, 2009). Bovine respiratory disease (**BRD**) is a multifactorial infectious disease of the upper and lower respiratory tract (Woolums, 2015) highly prevalent in this industry (Van der Mei et Van den Ingh, 1987, Leruste et al., 2012, Pardon et al., 2012b). Economically, BRD in veal calves is responsible for mortality (Sargeant et al., 1994b, a, Pardon et al., 2013, Lava et al., 2016a), lighter carcass weight and lower carcass quality at slaughter (Van der Mei et Van den Ingh, 1987, Pardon et al., 2013). Importantly, BRD is also a major cause of antimicrobial use in veal calves (Pardon et al., 2012a, Pardon et al., 2012b, Pardon et al., 2015). An accurate and reliable BRD individual and group-level diagnosis is therefore essential to reduce economic losses and use more judiciously antimicrobial.

Currently, BRD diagnosis and treatment decisions in the veal industry are generally based on examination for BRD clinical signs, performed mostly by producers on a daily basis, along with supervision and occasional intervention of veterinarians and technicians (i.e., veal calf experts that visits the farm each week to perform a quality inspection of the animals) (Leruste et al., 2012, Pardon et al., 2012a, Lava et al., 2016b). However, disagreement currently exists between BRD clinical signs detection during the fattening period and lung lesion prevalence at slaughter in veal calves (5.0% of batches with BRD clinical signs *vs* 51.7% of batches with lesions at slaughter, respectively) (Leruste et al., 2012). If a BRD clinical sign has a poor inter-rater agreement, it cannot be used to make an accurate diagnosis that is generalizable across examiners and practice settings (Kottner et al., 2011). Moreover, previous works in dairy cattle have shown differences of clinical examination findings between veterinarians and producers (Virtala et al., 1999, Olson et al., 2019). It is therefore essential to define reliable BRD clinical signs for producers, technicians and veterinarians working in the veal industry.

Currently, common specific BRD clinical signs are ocular discharge, nasal discharge, ear droop/head tilt, cough, abnormal respiration and hyperthermia (Lago et al., 2006, McGuirk, 2008, Love et al., 2014). To date, previous studies reporting inter-rater agreement for BRD detection were slight to fair (Amrine et al., 2013, Buczinski et al., 2016a). Robust individual BRD clinical signs inter-rater agreement estimations are needed to define which BRD clinical signs are reliable

enough among veterinarians, producers, and technicians to be incorporated in a scoring system and provide clear guidance for a practical examination process to standardize case definition.

The main objective of this study was, therefore, to evaluate the inter-rater agreement of clinical signs used to diagnose BRD in individually housed veal calves. Our hypothesis was that the different BRD clinical signs (ocular discharge, nasal discharge, ear droop/head tilt, abnormal respiration, induced cough) are not detected equally between different examiners (i.e., veterinarians, technicians, and producers) working in veal farming, but that some clinical signs are robust enough (i.e., have high inter-rater agreement) to be included in a clinical scoring system for BRD monitoring.

II- MATERIALS AND METHODS

To estimate the inter-rater agreement of ocular discharge, nasal discharge, ear droop/head tilt, abnormal respiration and spontaneous or induced cough, we designed a simulation study where videos of those BRD clinical signs were assessed by producers, technicians, and veterinarians. This project was designed according to the guidelines for reporting reliability and agreement studies (**GRRAS**) (Kottner et al., 2011). This study was approved by the Comité d'Éthique de l'utilisation des animaux de l'Université de Montréal (17_rech-1898).

II-1- Collection of the videos

II-1-1 Selection of calves

During a cross-sectional study performed from October 1, 2017, to December 20, 2018, 524 calves were prospectively and randomly recruited from 48 batches of 36 commercial rosé or white veal calf fattening units in Québec (DéliMAX, Veaux Lourds Ltée, Saint-Hyacinthe, Québec, Canada). A batch was defined as a group of calves that arrived at one farm at the same time to be fattened and slaughtered together (Leruste et al., 2012), and the selected batches were distributed across all four seasons. The veal calves were mostly obtained from multiple local dairy farms through auction markets in the first two weeks of life. The calves were housed individually during the early feeding period (up to 2 months). Since the number of calves present in every unit was known before the visit and each individual box was numbered on the farm, 10 individual veal calves were randomly chosen (RANDOM function in Excel (Microsoft, Redmond, WA)). Whenever a box was empty due to early mortality, the next random number was selected. During the visit, if a non-randomly

selected calf had potentially interesting relevant clinical signs on the farm, this additional calf was also filmed to potentially adjust for prevalence of normal/abnormal BRD clinical signs and to avoid prevalence bias impact on agreement measures (Byrt et al., 1993).

II-1-2- Recording of videos

Data were collected two weeks after the calves' arrival at the fattening unit in order to intervene during the early main clinical BRD peak incidence (Miller et al., 1980, Pardon et al., 2012b). At that time, calves were housed in individual duckboard pens (baby boxes). Data were collected in the morning (between 9:00 AM and 11:00 AM), between meals. Each calf was filmed with a portable camera (GoPro Hero 5, San Mateo, CA, USA) for a period of 2 to 3 min in order to assess the BRD clinical signs previously reported in the literature, i.e. nasal discharge, ocular discharge, ear droop/head tilt, abnormal respiration and spontaneous or induced cough (McGuirk, 2008, Love et al., 2014). Precisely, nasal discharge, ocular discharge, and ear droop/head tilt were recorded by filming the head of each calf when the calf was alone in its box. Abnormal respiration was recorded by focusing on the border between the abdominal and thoracic walls of each calf. To maintain a short recording period similar to that of a day-to-day screening examination process, spontaneous cough could not be consistently recorded. An operator (A.A. or S.B.) entered the baby box and induced cough was recorded by filming the calf's response to manual compression of the trachea. Videos had a resolution of 1,920 x 1,080 pixels at a frame rate of 60 frames per setting.

II-1-3- Selection of videos

Figure 3 shows the selection steps of the videos. First, a single operator (J.B.), not included as a future rater, classified each video in the following categories (**Figure 3; Step 1**): normal (with no abnormal BRD clinical signs), presence of nasal discharge, presence of ocular discharge, presence of ear droop/head tilt, presence of abnormal respiration or presence of induced cough. One video could be in several categories.

Next, 100 videos were randomly selected in each sign's category using the RANDOM function in Excel (Microsoft, Redmond, WA) (**Figure 3; Step 2**). If the number of videos was less than 100 in one category, all the videos were kept for the next step.

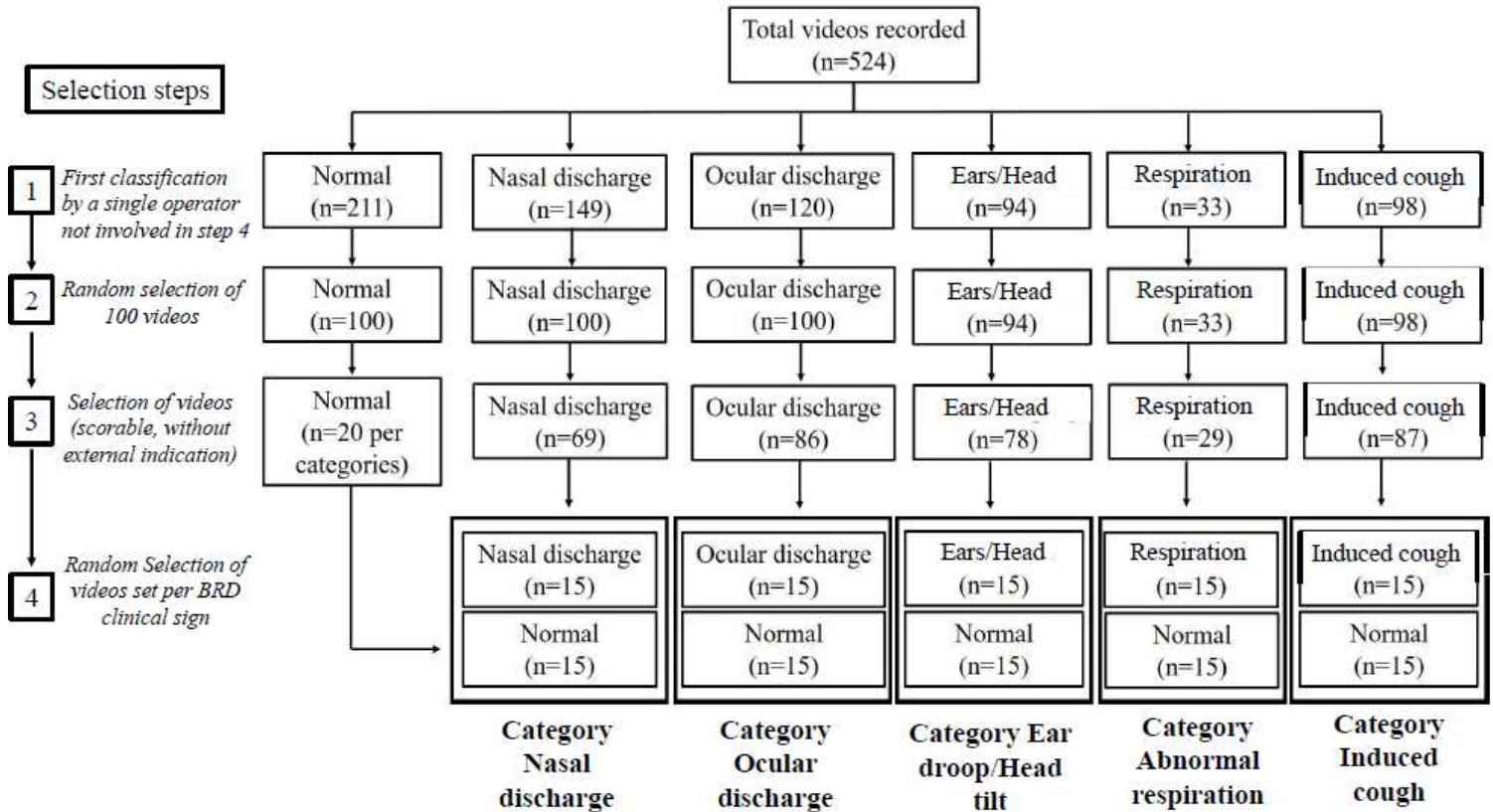


Figure 3. – Flowchart of the selection (4 steps) of the videos from the 254 filmed veal calves per category of BRD clinical signs: nasal discharge, ocular discharge, ear droop/head tilt, abnormal respiration, and induced cough.

Each selected video was then shortened with movie software (Windows Movie Maker, Windows 8, Microsoft, Redmond, WA) to isolate the targeted BRD clinical sign of the category. The 100 videos in the “normal” category were similarly shortened for each clinical sign and divided into groups of 20 control videos (supposedly negative based on initial screening) to assess the presence of nasal discharge, 20 control videos to assess the presence of ocular discharge, 20 control videos to assess the presence of ear droop/head tilt, 20 control videos to assess the presence of abnormal respiration, and 20 control videos to assess the presence of induced cough. Sound was preserved for all videos, and final videos for a specific clinical sign were 10 to 30 seconds in duration.

Finally, each video was scored by the same experienced operator (J.B.), who was not included as a future rater, using the score presented in **Table 22**.

Clinical signs	Score			
	0	1	2	3
Nasal discharge	Normal/serous discharge	Small amount of unilateral cloudy discharge	Bilateral, cloudy or excessive mucus discharge	Copious bilateral mucopurulent discharge
Ocular discharge	Normal	Small amount of ocular discharge	Moderate amount of bilateral discharge	Heavy ocular discharge
Ear droop and head tilt	Normal	Ear flick or head shake	Slight unilateral droop	Head tilt or bilateral droop
Abnormal respiration	Normal	Increase of the respiratory frequency without BRD	Pathologic respiratory pattern	
Cough	None	Induced single cough	Induced repeated coughs	

¹ This score was adapted from the Wisconsin score (McGuirk, 2008) and the California score (Love et al., 2014).

Tableau 22. – Scoring system used by the 18 raters (6 veterinarians, 6 technicians, and 6 producers) to assess the severity of clinical signs of bovine respiratory disease (BRD) of 524 filmed veal calves.¹

A video was included if the scoring of the target BRD clinical sign was judged possible by the operator (J.B.) (e.g., visualization of two eyes for the scoring of ocular discharge or two nostrils for the scoring of nasal discharge; **Figure 3; Step 3**). Moreover, videos with excessive darkness, contrast problems or excessive movement of the calf were not included in this study since it was anticipated that it could impact the scoring ability of the raters. A video was also excluded if the video was already selected in other category, and if external elements could influence the assessment of the BRD clinical signs (written indication of treatment on the box, other apparent clinical signs impossible to dissociate with the target BRD clinical sign on the video (e.g., presence of nasal discharge in the ocular discharge category) or oral indication (i.e., examiners talking about a particular clinical sign) during the physical examination).

II-2- Assessment of the videos by the raters

Videos were randomly selected within the database for a total of 30 videos per clinical sign (**Figure 3; Step 4**). Because inter-rater agreement estimated by Cohen’s kappa (κ) can be affected when one of the scoring categories being rated has low prevalence (so-called κ paradox), the 2-by-2 table margins were initially set to have a prevalence of 0.5 as initially classified (i.e., each set of 30 videos per BRD clinical sign included approximately 15 videos initially classified by J.B. without and 15 videos classified with the clinical sign (score 1, 2 or 3)).

II-2-1- Selection and training of the raters

A convenience sample was made up of 6 veterinarians, 6 technicians, and 6 producers. All the raters had various practical and relevant experience in the assessment of BRD in veal calves (**Table 23**).

Group	Rater	Experience (years)
Veterinarians	V1	5 to <10
	V2	+10
	V3	+10
	V4	+10
	V5	+10
	V6	1 to <5
Technicians	T1	+10
	T2	1 to <5
	T3	+10
	T4	5 to <10
	T5	+10
	T6	1 to <5
Producers	P1	1 to <5
	P2	1 to <5
	P3	+10
	P4	5 to <10
	P5	5 to <10
	P6	+10

Tableau 23. – Experience years (1 to <5; 5 to <10; +10) of work in calves or veal calves’ health of the 6 veterinarians (V1 to V6), the 6 technicians (T1 to T6) and the 6 producers (P1 to P6) assessing videos of clinical signs associated with bovine respiratory disease of 524 veal calves.

Before presenting the videos, each rater received basic training (slide show presentation) with specific explanations of how to score the severity of each BRD clinical sign (**Table 22**). Briefly, the assessment of nasal discharge, ocular discharge and ear droop/head tilt was performed using the 4-level scale of the Wisconsin scoring system chart (**WI**) as published by McGuirk (2008) The cough assessment was modified from the WI score because of the difficulty recording grade 3

(spontaneous cough) in this study using a relatively short duration of observation. The California score (CA; Love et al. (2014)) was modified for assessing abnormal respiration consisting of a score of 0 if the respiratory pattern was judged normal, of 1 if tachypnea without dyspnea was present, and of 2 if dyspnea was present.

II-2-2- Assessment of the videos

Each video was then individually scored by each recruited rater via the tutorial software Edpuzzle (Mischel, 2019). The videos were available online, and each rater could individually visualize them anytime and anywhere via his/her computer. The assessment of videos was consequently performed independently and blindly of the other raters' responses.

Briefly, each rater chose a category including the videos of an *a priori* targeted BRD clinical sign among nasal discharge, ocular discharge, ear droop/head tilt, abnormal respiration, or induced cough. Information concerning the farm, season, and treatment history was not available. Each video was then individually scored by each rater. Before scoring the next video, the score of the current video had to be validated in the software to avoid changing answers. At any time, the participants could save and end their scoring session, and the assessment of the remaining videos could be completed later. The time required to score a video was not recorded and there was no time limit to complete the scoring of all the videos. All raters were aware that the clinical assessment should be compatible with practical use in veal farm units.

II-3- Statistical analyses

All statistical analyses were performed using SAS v 9.4, (SAS, Cary, NC).

II-3-1- Power sample size justification of number of videos and rater

In the absence of data on inter-rater agreement for BRD clinical signs in cattle, we relied on sample size calculation for a traditional Cohen's kappa (κ) to detect a moderate agreement between raters for each BRD clinical sign, i.e. a $\kappa \geq 0.6$ (Landis and Koch, 1977). Sample size was consequently determined to obtain the minimum number of videos per BRD clinical sign and the minimum number of raters per group (veterinarians, technicians or producers) that provide 80% power for testing $H_0 = \kappa < 0.6$ vs $H_1 = \kappa \geq 0.6$ at the 5% level of significance. A sample size of 30 videos per BRD clinical sign and 6 raters per group was determined by using the Donner and Eliasziw (1987) method.

II-3-2- Assessment of different scoring combinations

Different scoring combinations from the scoring system in **Table 22** were assigned to each clinical sign. Briefly, inter-rater agreement was calculated for the original 4-level scale (0/1/2/3) or 3-level scale (0/1/2) and after merging different combinations of adjacent levels to create the 2-level scale (0/1,2,3; 0,1/2,3; 0,1/2; 0/1,2). The following combinations 0,1,2/3 and 0,1/2/3 were not assessed since a limited number of calves were scored 3 by the raters (**Table 24**). Agreement between each pair of raters was assessed in a k-by-k table, k being the number of the level scale – 2 or 3.

II-3-3- Assessment of the inter-rater agreement

Inter-rater agreement was determined from the k-by-k table using the AgreeStat macro and the function KAPPA from PROC FREQ in SAS. A variety of methods have been described for measuring inter-rater agreement in the literature (Cohen, 1960, Gwet, 2014). To date, no single indicator has been considered as the most appropriate measure. We decided, therefore, to estimate inter-rater agreement with the three following complementary statistical approaches:

- Raw percentage of agreement (**Pa**): noted as a crude marker of concordant pairs of ratings that does not account for agreement arising from chance alone. It is the most fundamental and intuitive approach. Briefly, Pa is defined as the total number of examinations where agreement is met by 2 raters divided by the total number of videos. Clopper-Pearson 95% confidence intervals (95% **CI**) were calculated for Pa (Brown et al., 2001). Because of the intervention of multiple raters in this study, the median Pa for each BRD clinical sign was calculated as the median of all Pa values calculated between 2 raters.
- Agreement beyond chance was assessed using Cohen's kappa test (κ) between pairs of raters. The kappa reports Pa corrected for agreement due to chance (**Pc**) corresponding to the probability that a rater classified one video in one category:
 $\kappa = (Pa - Pc) / (1 - Pc)$ (Cohen, 1960). For the combination including 3 or 4 levels (0/1/2 and 0/1/2/3), the penalty for disagreement was weighted according to the number of categories by which the raters disagreed and the weighted kappa (**w κ**) was calculated using ordinal weights (Gwet, 2014). The 95% CIs were calculated for both κ and w κ using the function KAPPA PROC FREQ in SAS (Fleiss et al., 1969). Because of the intervention of multiple raters in this study, the median κ was calculated as the median of all κ values calculated between 2 raters.

- Finally, general inter-rater agreement (multiple raters) was evaluated using the Gwet's agreement coefficient type 1 (**AC1**). Gwet's AC1 provides a chance-corrected agreement coefficient, in line with the percentage level of agreement (Gwet, 2014). This parameter allowed estimation of inter-rater agreement of raters in general, rather than for any specific rater pair. It differs from κ mainly in the way the percent chance of agreement is calculated and appears more stable than κ in a homogenous population (paradox resistant) and consequently is recommended by the authors for use in the medical literature (Wongpakaran et al., 2013). The 95% CIs were calculated for AC1 as described by Gwet (2008) using the AgreeStat macro in SAS and by assuming AC1 unconditional to raters or videos.

II-3-4- Interpretation of the inter-rater agreement

The most reliable combination for each BRD clinical sign was defined as the combination with the highest inter-rater agreement for at least two statistical coefficients (Pa, κ or AC1).

We used the probabilistic statistical approach proposed by Gwet (2014) to interpret the magnitude of the inter-rater agreement for BRD clinical signs. Briefly, for any given agreement coefficient, we first quantified the probability that the extent of agreement falls into each of the benchmark intervals defined by Landis and Koch (1977) (poor agreement for values <0.20 ; slight agreement for values between 0.21 and 0.40; moderate agreement for values between 0.41 and 0.60; substantial agreement for values between 0.61 and 0.80; and almost perfect agreement for values between 0.81 and 1) (Step 1). For each interval, an interval membership probability (**IMP**) was calculated. Briefly, it is the probability that a standard normal variate Z falls within the interval, that is: $IMP = P((X - b)/SE) \leq Z \leq (X - a)/SE)$

Where X corresponds to median Pa, median κ or AC1 of a BRD clinical sign, respectively and SE corresponds to their standard errors; a and b correspond to the lower and upper limits of the intervals, respectively.

We then computed the cumulative probability for each interval starting from the top and going down to the bottom (Step 2). That is, the cumulative probability associated with the top interval is identical to the probability calculated in Step 1 for that interval. For the second interval in descending order, the cumulative probability is the sum of probabilities calculated in Step 1 for the two top intervals. The final benchmark level was determined by the interval associated (poor,

moderate, substantial, or almost perfect) with the smallest cumulative probability that exceeds 0.95. We therefore concluded with 95% certainty or more that the extent of agreement among raters was poor, moderate, substantial, or almost perfect.

II-3-5- Differences between veterinarians, technicians, and producers

The inter-rater agreements of the different groups (veterinarians, technicians, and producers) were calculated for each BRD clinical sign. The inter-rater agreement differences between groups were globally estimated by a nonparametric test for dependent samples: the Kruskal–Wallis test. Specific differences between groups (veterinarians *vs* technicians, technicians *vs* producers and veterinarians *vs* producers) were estimated by a nonparametric test for dependent samples: the Wilcoxon signed-rank test. All comparisons were performed by using PROC NPAR1WAY from SAS. A *p*-value <0.05 was considered significant to exclude the null hypothesis of equality between groups.

III- RESULTS

III-1- Participants, videos, and raters

Among the 524 calves, 95.8% (n=502) were male. The breeds represented in the population were: Holstein (95%; n=497), Red Holstein (3%; n=16), Jersey (1%; n=5), Ayrshire (1%; n=5) and Brown Swiss (<1%; n=1). Overall, 149 videos were assessed by 18 raters after the selection process. One video included in the ear droop/head tilt category was not downloaded correctly on the web platform and was not available for the raters. The total video scoring time varied between two days and two weeks for the different raters.

The approximative previous experience in calf and/or veal calf health of the 6 veterinarians, the 6 technicians and the 6 producers is presented in **Table 22**. Seven raters were women (7/18). Most raters (13/18) had more than 5 years of experience working with veal calves and/or calf health.

III-2- Overall inter-rater agreement for the different scoring combinations

Table 24 shows the distribution of severity scores assigned to each BRD clinical sign using the original 3- and 4-level scales by the 18 raters. Despite initially fixing the prevalence of videos to 0.5 (absence–presence of the BRD clinical sign; **Figure 3; Step 1**), the distribution of severity score (1 and 2, or 1, 2, and 3) obtained was more variable.

Nasal discharge																		
	<i>Veterinarians</i>						<i>Technicians</i>						<i>Producers</i>					
	<i>V1</i>	<i>V2</i>	<i>V3</i>	<i>V4</i>	<i>V5</i>	<i>V6</i>	<i>T1</i>	<i>T2</i>	<i>T3</i>	<i>T4</i>	<i>T5</i>	<i>T6</i>	<i>P1</i>	<i>P2</i>	<i>P3</i>	<i>P4</i>	<i>P5</i>	<i>P6</i>
0	16	18	19	14	17	17	7	17	14	15	19	16	11	7	15	2	15	7
1	7	8	7	13	6	8	13	10	10	7	7	10	6	9	10	15	11	16
2	5	3	4	3	7	4	4	2	5	7	2	4	10	10	5	12	4	6
3	2	1	0	0	0	1	6	1	1	1	2	0	3	4	0	1	0	1

Ocular discharge																		
	<i>Veterinarians</i>						<i>Technicians</i>						<i>Producers</i>					
	<i>V1</i>	<i>V2</i>	<i>V3</i>	<i>V4</i>	<i>V5</i>	<i>V6</i>	<i>T1</i>	<i>T2</i>	<i>T3</i>	<i>T4</i>	<i>T5</i>	<i>T6</i>	<i>P1</i>	<i>P2</i>	<i>P3</i>	<i>P4</i>	<i>P5</i>	<i>P6</i>
0	24	16	17	12	16	10	13	19	11	14	19	17	12	17	20	7	22	13
1	3	8	7	10	8	13	6	7	11	8	4	9	11	5	6	13	3	8
2	2	5	4	4	5	6	7	2	7	7	6	3	5	6	2	8	3	4
3	0	0	1	3	0	0	3	1	0	0	0	0	1	1	1	1	0	4

Ear droop and head tilt																		
	<i>Veterinarians</i>						<i>Technicians</i>						<i>Producers</i>					
	<i>V1</i>	<i>V2</i>	<i>V3</i>	<i>V4</i>	<i>V5</i>	<i>V6</i>	<i>T1</i>	<i>T2</i>	<i>T3</i>	<i>T4</i>	<i>T5</i>	<i>T6</i>	<i>P1</i>	<i>P2</i>	<i>P3</i>	<i>P4</i>	<i>P5</i>	<i>P6</i>
0	9	9	14	13	9	10	13	19	11	14	19	17	12	13	14	5	12	6
1	8	7	3	3	2	3	6	7	11	8	4	9	3	3	2	11	3	9
2	4	4	6	7	10	6	7	2	7	7	6	3	5	6	5	11	6	6
3	9	10	7	7	9	11	3	1	0	0	0	0	10	8	9	3	9	9

Abnormal respiration																		
	<i>Veterinarians</i>						<i>Technicians</i>						<i>Producers</i>					
	<i>V1</i>	<i>V2</i>	<i>V3</i>	<i>V4</i>	<i>V5</i>	<i>V6</i>	<i>T1</i>	<i>T2</i>	<i>T3</i>	<i>T4</i>	<i>T5</i>	<i>T6</i>	<i>P1</i>	<i>P2</i>	<i>P3</i>	<i>P4</i>	<i>P5</i>	<i>P6</i>
0	15	13	24	17	22	18	9	19	10	18	21	16	14	16	17	3	22	9
1	8	11	4	7	7	10	10	7	11	11	6	7	10	6	8	17	8	14
2	7	6	2	6	1	2	11	4	9	1	3	7	6	8	5	10	0	7

Induced cough																		
	<i>Veterinarians</i>						<i>Technicians</i>						<i>Producers</i>					
	<i>V1</i>	<i>V2</i>	<i>V3</i>	<i>V4</i>	<i>V5</i>	<i>V6</i>	<i>T1</i>	<i>T2</i>	<i>T3</i>	<i>T4</i>	<i>T5</i>	<i>T6</i>	<i>P1</i>	<i>P2</i>	<i>P3</i>	<i>P4</i>	<i>P5</i>	<i>P6</i>
0	15	13	24	17	22	18	9	19	10	18	21	16	14	16	17	3	22	9
1	8	11	4	7	7	10	10	7	11	11	6	7	10	6	8	17	8	14
2	7	6	2	6	1	2	11	4	9	1	3	7	6	8	5	10	0	7

Tableau 24. – Number of videos of the 254 filmed veal calves for each score assigned to each Bovine Respiratory Disease clinical sign by 18 raters: 6 veterinarians (V1 to V6); 6 technicians (T1 to T6), and 6 producers (P1 to P6).

Table 25 shows the overall inter-rater agreement of each scoring combination of each BRD clinical sign. Whatever the combination, induced cough was the BRD clinical sign with the highest inter-rater agreement. Abnormal respiration was the BRD clinical sign with the lowest inter-rater agreement.

BRD clinical sign scoring	Median Pa	Median κ	AC1
Nasal discharge			
0/1/2/3	0.53 (0.34; 0.72)	0.43 (0.21; 0.66)	0.40 (0.26; 0.54)
0/1,2,3	0.73 (0.54; 0.88)	0.47 (0.16; 0.76)	0.45 (0.29; 0.61)
0,1/2,3	0.83 (0.65; 0.94)	0.45 (0.07; 0.84)	0.71 (0.57; 0.85)
Ocular discharge			
0/1/2/3	0.53 (0.34; 0.72)	0.39 (0.15; 0.64)	0.45 (0.31; 0.59)
0/1,2,3	0.67 (0.47; 0.83)	0.37 (0.08; 0.70)	0.38 (0.22; 0.54)
0,1/2,3	0.80 (0.61; 0.92)	0.44 (0.07; 0.86)	0.77 (0.65; 0.89)
Ear droop or head tilt			
0/1/2/3	0.57 (0.37; 0.75)	0.59 (0.36; 0.78)	0.43 (0.29; 0.57)
0/1,2,3	0.77 (0.58; 0.90)	0.49 (0.19; 0.81)	0.56 (0.36; 0.76)
0,1/2,3	0.83 (0.65; 0.94)	0.67 (0.42; 0.92)	0.60 (0.44; 0.76)
Abnormal respiration			
0/1/2	0.57 (0.37; 0.75)	0.34 (0.11; 0.57)	0.43 (0.29; 0.57)
0,1/2	0.77 (0.58; 0.90)	0.20 (-0.13; 0.62)	0.74 (0.62; 0.85)
0/1,2	0.70 (0.51; 0.85)	0.42 (0.13; 0.71)	0.59 (0.47; 0.71)
Induced cough			
0/1/2	0.83 (0.65; 0.94)	0.76 (0.59; 0.96)	0.76 (0.64; 0.88)
0,1/2	0.93 (0.78; 0.99)	0.79 (0.52; 1.00)	0.87 (0.73; 1.00)
0/1,2	0.90 (0.73; 0.98)	0.80 (0.58; 1.00)	0.85 (0.75; 0.95)

Tableau 25. – Overall inter-rater agreements (and 95% CI) of the different scoring combination of each bovine respiratory disease (BRD) clinical sign from 254 filmed veal calves; inter-rater agreement was estimated with the median percentage agreement (Pa), the median weighted (scoring categories > 2) or unweighted (scoring categories < 2) kappa (κ), and the Gwet’s agreement coefficient type 1 (AC1). The scoring combination with the higher inter-rater agreement are highlighted in grey for each BRD clinical sign.

Within each BRD clinical sign, the 2-level scoring combination 0,1/2,3 (absence-mild/moderate-severe) for nasal discharge, ocular discharge, and ear droop/head tilt, and 0/1,2 (absence/presence) for abnormal respiration and induced cough had the highest inter-rater agreement.

Table 26 shows the magnitude of the inter-rater BRD clinical sign agreements for the best combination (2-levels). The cumulative probability for each interval was different according to statistical coefficient (median Pa, median κ , and AC1) (**Table 26; B**). The BRD clinical signs

showing a more than 95% probability of belonging in an interval between 0.4 and 1 (i.e., moderate or higher inter-rater agreement) were ear/head carriage and induced cough.

		Nasal discharge			Ocular discharge			Ear droop/Head tilt			Abnormal respiration			Induced cough		
Benchmark Range		Median Pa	Median κ	AC1	Median Pa	Median κ	AC1	Median Pa	Median κ	AC1	Median Pa	Median κ	AC1	Median Pa	Median κ	AC1
0.8 to 1.0	Almost perfect	0.66	0.02	0.01	0.50	0.03	0.31	0.66	0.15	0.01	0.11	0.01	0	0.95	0.47	0.84
0.6 to 0.8	Substantial	0.33	0.17	0.84	0.50	0.17	0.69	0.33	0.55	0.49	0.79	0.11	0.43	0.02	0.47	0.16
0.4 to 0.6	Moderate	0	0.43	0.06	0.002	0.38	0	0	0.23	0.49	0.11	0.44	0.57	0	0.03	0
0.2 to 0.4	Slight	0	0.31	0	0	0.31	0	0	0.02	0.01	0	0.38	0	0	0	0
< 0.2	Poor	0	0.07	0	0	0.09	0	0	0	0	0	0.07	0	0	0	0

*For each interval, interval membership probability ($IMP = P((COEFF - b) / SE) \leq Z \leq (COEFF - a) / SE$) was calculated. Where COEFF corresponds to median Pa, median κ or AC1 of a BRD clinical sign respectively and SE their standards errors; a and b corresponds to the lower and upper limit of the intervals, respectively.

		Nasal discharge			Ocular discharge			Ear droop/Head tilt			Abnormal respiration			Induced cough		
Benchmark Range		Median Pa	Median κ	AC1	Median Pa	Median κ	AC1	Median Pa	Median κ	AC1	Median Pa	Median κ	AC1	Median Pa	Median κ	AC1
0.8 to 1.0	Almost perfect	0.66	0.02	0.10	0.50	0.03	0.31	0.66	0.15	0.01	0.11	0.01	0.002	0.95	0.47	0.84
0.6 to 0.8	Substantial	0.99	0.19	0.94	1.0	0.20	1.0	0.99	0.70	0.50	0.89	0.12	0.43	0.98	0.93	1.0
0.4 to 0.6	Moderate	0.99	0.62	1.0	1.0	0.58	1.0	0.99	0.98	0.99	1.0	0.55	1.0	0.98	0.97	1.0
0.2 to 0.4	Slight	0.99	0.93	1.0	1.0	0.90	1.0	0.99	0.99	1.0	1.0	0.93	1.0	0.98	0.97	1.0
< 0.2	Poor	0.99	0.96	1.0	1.0	0.99	1.0	0.99	0.99	1.0	1.0	1.0	1.0	0.98	0.97	1.0

Tableau 26. – Probabilistic statistical approach to interpret the magnitude of the inter-rater bovine respiratory disease (BRD) clinical signs agreements from 254 filmed veal calves, for the 2-levels scoring combination: 0,1/2,3 (absence-mild/moderate-severe) for nasal discharge, ocular discharge, and ear droop/head tilt; or 0/1,2 (absence/presence) for abnormal respiration and induced cough. For any given agreement coefficient (median percentage agreement (Pa), median weighted (scoring categories > 2) or unweighted (scoring categories < 2) kappa (κ), and Gwet’s agreement coefficient type 1 (AC1)), the probability that the extent of agreement falls into each of the benchmark intervals defining by Landis and Koch (1977) was quantified (A). The cumulative probability for each interval was then computed (B). The final benchmark level was determined by the interval associated with the smallest cumulative probability that exceeds 0.95 (highlighted in grey).

III-3- Inter-rater agreement between veterinarians, technicians and producers

Table 27 shows the inter-veterinarian, -technician, and -producer agreement of each BRD clinical sign, for the most reliable scoring combinations (i.e., 0,1/2,3 (absence-mild/moderate-severe) for nasal discharge, ocular discharge, and ear droop/head tilt, and 0/1,2 (absence/presence) for abnormal respiration and induced cough). Inter-veterinarian agreement was significantly higher than inter-technician and inter-producer agreement for all BRD clinical signs except ear droop/head tilt, where inter-producer agreement was higher (**Figure 4**).

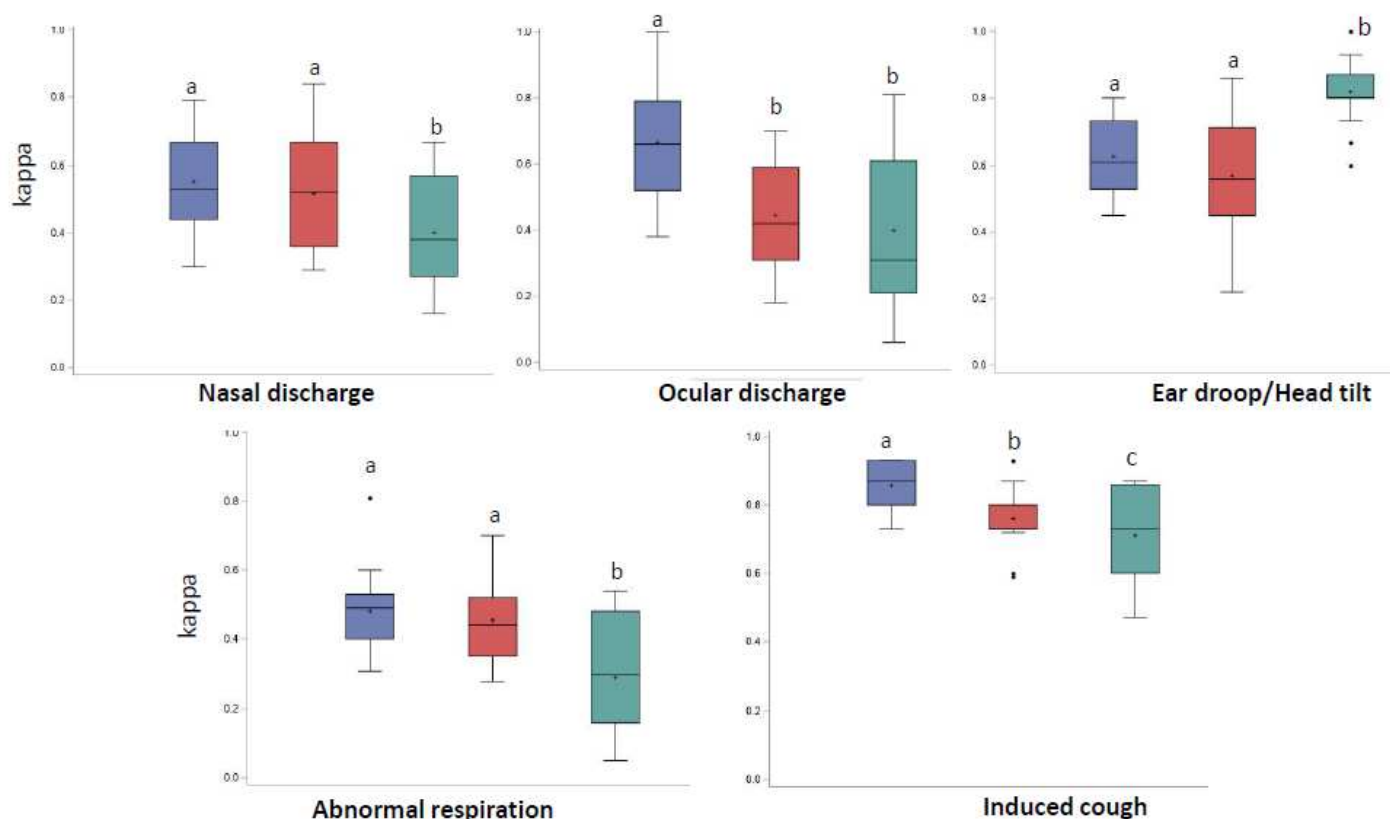


Figure 4. – Boxplot showing the distribution of kappa value calculated between pairs of raters between the 6 veterinarians (blue), the 6 technicians (red), and the 6 producers (green) for the different clinical signs of bovine respiratory disease (BRD). The distribution of the kappa is displayed using seven values: minimum and maximum (whisker); first quartile (bottom of box), median (horizontal bar), and third quartile (upper of box); mean (+); and outliers (dots). Columns that do not share at least one letter are significantly different (Wilcoxon; p -value < 0.05).

BRD clinical sign scoring	Veterinarians			Technicians			Producers		
	Median Pa	Median κ	AC1	Median Pa	Median κ	AC1	Median Pa	Median κ	AC1
Nasal discharge	0.90 (0.73; 0.98)	0.53 (0.16; 0.99)	0.52 (-0.17; 0.68)	0.87 (0.69; 0.96)	0.52 (0.11; 0.93)	0.48 (-0.19; 0.71)	0.70 (0.51; 0.85)	0.38 (0.07; 0.72)	0.34 (-0.31; 0.93)
Ocular discharge	0.87 (0.69; 0.96)	0.66 (0.31; 1.00)	0.55 (-0.16; 0.66)	0.80 (0.61; 0.92)	0.42 (0.08; 0.79)	0.45 (-0.22; 0.76)	0.77 (0.58; 0.90)	0.31 (-0.06; 0.71)	0.43 (-0.24; 0.80)
Ear droop or head tilt	0.80 (0.61; 0.92)	0.61 (0.36; 0.88)	0.40 (-0.31; 0.96)	0.83 (0.65; 0.94)	0.56 (0.30; 0.81)	0.36 (-0.33; 0.99)	0.90 (0.73; 0.98)	0.80 (0.59; 1.00)	0.49 (-0.27; 0.93)
Abnormal respiration	0.77 (0.58; 0.90)	0.49 (0.20; 0.78)	0.33 (-0.34; 1.00)	0.70 (0.51; 0.85)	0.44 (0.14; 0.70)	0.30 (-0.35; 1.00)	0.67 (0.47; 0.83)	0.30 (-0.02; 0.61)	0.30 (-0.02; 0.61)
Inductive cough	0.93 (0.78; 0.99)	0.87 (0.69; 1.00)	0.52 (-0.24; 0.87)	0.90 (0.73; 0.98)	0.73 (0.5; 0.98)	0.47 (-0.27; 0.92)	0.87 (0.69; 0.96)	0.73 (0.49; 0.97)	0.50 (-0.13; 0.57)

¹ Inter-rater agreements are estimated with the median percentage agreement (Pa), the median kappa (κ), and Gwet's agreement coefficient type 1 (AC1).

Tableau 27. – Inter-veterinarian, -technician, and -producer agreement and their 95% CI of each bovine respiratory disease (BRD) clinical sign from 254 filmed veal calves for the 2-levels scoring combination: 0,1/2,3 (absence-mild/moderate-severe) for nasal discharge, ocular discharge, and ear droop or head tilt; or 0/1,2 (absence/presence) for abnormal respiration and induced cough.¹

IV- DISCUSSION

In this study, we estimated inter-rater agreement of BRD clinical signs in veal calves in individual housing (nasal discharge, ocular discharge, ear droop/head tilt, abnormal respiration, and induced cough) for different scoring combinations, with various and complementary statistical coefficients. The 2-scale combination 0,1/2,3 (absence, mild/moderate, severe) for nasal discharge, ocular discharge, and ear droop/head tilt and the 2-scale combination 0/1,2 (absence/presence) for abnormal respiration and induced cough showed the highest repeatability between raters. The inter-rater agreement varied from one BRD clinical sign to another. Induced cough and abnormal respiration had the highest and lowest inter-rater agreement, respectively. Only inter-rater agreements of induced cough and ear droop/head tilt were at least moderate (>0.6) for all statistical coefficients and consequently were judged clinically acceptable. Generally, veterinarians were more consistent in scoring the different BRD clinical signs than technicians and producers, except for ear droop/head tilt, where producers more consistently scored this sign.

In this study, the 150 videos were scored blindly and individually online without a time limit. The influence of other raters was therefore nil, as was the fatigue and stress usually associated with a time-limited session. The assessment could also be considered close to the real world, where only a few calves are examined at once. Each video could be watched several times but had to be validated once to assess the next one. The score could not be modified afterward, thus avoiding the influence of other videos. On the other hand, a limitation of the videos was that raters could not adjust their view angle or limit calf movement to assess the presence and severity of the targeted BRD clinical sign. Additionally, each BRD clinical sign was isolated, which prevented influence from other BRD signs when evaluating a given aspect. Moreover, the raters were informed beforehand what was to be evaluated in the video influencing the mind preset before taking the test. Ideally, the 18 raters should have examined each calf, on site, in a short time period. That seemed technically impossible in veal calf production (limited space, biosecurity, availability of raters, etc.). Simultaneous on-farm scoring could also impact individual scoring due to possible interactions between raters. The recording of videos was, therefore, considered the best alternative to standardize the object to be scored (i.e., the same images and video for each scorer).

On each video, the presence and the severity of each BRD clinical sign included in this study was defined based on the two clinical scoring systems developed in pre-weaned dairy calves: the

systematic scoring system of WI (McGuirk, 2008) and CA (Love et al., 2014). In contrast with these two scoring systems, spontaneous cough could not be systematically assessed because the videos were short. Although often used by operators for BRD detection, this sign is difficult to standardize since spontaneous cough observation cannot be dissociated from the period of observation. For example, a calf coughing every 15 min could be observed for a 1 min period with no coughing, whereas a calf coughing every 2h would be unlikely to be observed coughing. Therefore, standardization of assessment of the upper airway by pinching the trachea was considered a more reliable and systematic alternative. However, we could not study the variability of stimulating the trachea across operators that could influence our results in real life. Additionally, rectal temperature was measured in our study but its variability between raters was not studied despite it has already been reported (Naylor et al., 2012).

The combination scale of severity differs between both WI and CA. Subjectively, the WI score subdivides each of its clinical signs into 4 levels, while CA subdivides them into 2 levels. In our study, we studied the most reliable scale for each BRD clinical sign. We found that a 2-scale combination of clinical signs (0,1/2,3 (absence, mild/moderate, severe) for nasal discharge, ocular discharge and ear droop/head tilt, and 0/1,2 (absence/presence) for abnormal respiration and induced cough) was the most reliable scale to assess each BRD clinical sign. It was anticipated that the inclusion of fewer levels of clinical predictors will simplify the assessment of severity and will improve the reliability of each BRD clinical sign, minimizing the possibility of disagreement. Indeed, the multiple level scale could lead to ambiguous overlap by inexperienced individuals, making it difficult to appropriately classify BRD clinical signs in calves.

Additionally, the simplification from multiple levels to 2 levels has already been reported in dairy cows for improving the lameness scoring system (Schlageter-Tello et al., 2014) as well as in dairy calves' BRD clinical signs with the CA scoring system (Love et al. (2014). Although the 2-scale combination was the same for induced cough (0/1,2; presence/absence), abnormal respiration (0/1,2; presence/absence) and ear droop/head tilt (0,1/2,3; absence, mild/moderate, severe) in the CA, the 2-scale combination for nasal discharge and ocular discharge was different with a division of absence/presence (0/1,2,3) vs absence, mild/moderate, severe (0,1/2,3) in our study. However, the statistical methods used in the previous study was different, which could explain the discrepancy with our results.

In accordance with our hypothesis, the inter-rater agreement was different depending on the BRD clinical sign. This finding was in agreement with Buczinski et al. (2016a). We found that, whatever the combination, induced cough was the sign with the highest inter-rater agreement. By definition, coughing is a reflex characterized by a forceful and noisy expulsion of air through the glottis (Woolums, 2015). This BRD clinical sign is, therefore, more likely to be detected by the specific movement of the neck and the characteristic sound, visible and audible in videos. Additionally, its severity assessment followed a strict definition without implying interpretation (absence *vs* induced single cough *vs* induced repeated coughs), making its scoring more objective.

Abnormal respiration was the BRD clinical sign with the lowest inter-rater agreement. In contrast with the induced cough, abnormal respiration indicates an inappropriate degree of breathing efforts based on an assessment of the respiratory rate, rhythm, and character (Woolums, 2015). In our study, we decided to assess abnormal respiration as it was reported in the CA scoring system, i.e. without a specific definition including frequency rate or abnormal dynamic description (Love et al., 2014). Because its assessment is based on different criteria, detecting abnormal respiration needs a minimum of education and definition, which could explain why abnormal respiration was the sign with the lowest inter-rater agreement. However, it is also possible that abnormal respiration is more difficult to assess by video because of its dynamic characteristic and the contrast between luminosity at farms and the color of the calves, thus decreasing the perception of movement, and thus explaining the low inter-rater agreement observed in our study. Because ambient temperature was recently reported as an important confounder in the BRD detection (Maier et al., 2019), the more objective BRD clinical sign “tachypnea without dyspnea” would be of interest in a future study.

The assessment of the other BRD clinical signs (nasal discharge, ocular discharge, ear droop/head tilt) was more precisely defined than abnormal respiration, but it needed more interpretation than induced cough (i.e., it was not only a movement or a noise). This interpretation could explain the variable reliability between raters of these BRD clinical signs. However, ear droop/head tilt showed the highest inter-rater agreement after induced cough. The typically high prevalence of *Mycoplasma bovis* responsible for otitis is commonly observed in veal farms (Radaelli et al., 2008). This characteristic could have made raters more aware of this clinical sign. This result could be different in small herds or other farms where *Mycoplasma bovis* is less prevalent.

The reliability of individual BRD clinical signs (presence and severity) was never compared between different professions working with veal calves such as veterinarians, technicians, and producers. For two statistical coefficients (median Pa and median κ), the veterinarians were more consistent in scoring BRD clinical signs than technicians and producers, except for ear droop/head tilt, where it was more consistently detected by producers. This finding could suggest a lack of standardization between professions in teaching the assessment of BRD clinical signs. This lack of standardization in cattle medicine was described previously by Sicho et al. (2019) who mentioned that veterinarians were infrequently involved in calf-care worker training and that most training occurred either worker-to-worker or manager-to-worker. This difference of reliability in scoring BRD clinical signs could also explain the difference reported in the detection of the disease between veterinarians and producers (Virtala et al., 1999, Olson et al., 2019). Consequently, our results promote the necessity of providing a clear and common definition and training of each BRD clinical sign to all professions (veterinarians, technicians, and producers) in order to improve BRD detection, identify appropriate interventions and reduce antimicrobial use.

In this study, we showed that inter-rater agreement varied according to BRD clinical signs. Thus, it is logical to think that if individual BRD clinical signs are not assessed in the same way by two different raters, this lack of reliability is also present when we use scoring systems. This low repeatability between raters could explain why current scoring systems lack accuracy (Buczinski, et al., 2015a; Love, et al. 2016). For instance, currently, two herds or batches with the same prevalence could have a different proportion of treatment because of the lack of scoring system reliability. Thus, if we want to detect and treat veal calves in individual housing adequately and universally, it is important to implement a scoring system that is used similarly by each operator with the goal of standardizing case definition. Additionally, because confidence interval surrounding any prediction for an individual is substantially wider than for a group, a reliable scoring system would be more accurate if it is applied on a population basis (Hayes et al., 2010).

Interestingly, considering all statistical coefficients, only inter-rater agreements of induced cough and ear droop/head tilt were at least moderate (>0.6) and consequently judged clinically acceptable when applied by raters with different backgrounds. Although a handling of calves will be needed, those two BRD clinical signs, in addition to the rectal temperature, could therefore be potentially used to design a reliable scoring system for monitoring and deciding a treatment group. Such

scoring system need, however, to be developed, validated, and compared with scoring system including all the BRD clinical signs to detect veal calves with bronchopneumonia that require a treatment and not only calves with BRD clinical signs.

The current study has its own limitations. Firstly, the raters included persons working with veal calves and experienced in BRD with minimal training in score assessment. It is, therefore, possible that inter-rater agreement would be different with novice raters or raters inexperienced with veal calves. Secondly, intra-rater agreement was not considered in this study, so it is difficult to know if the difference observed was within a group (veterinarians, technicians, or producers) or because of the inaccuracy of one rater. Additionally, since we generally found better agreement among veterinarians, it is unknown if a more intensive training period with technicians and producers would increase agreement between the latter two groups. Moreover, the inclusion of veterinarians more experimented in the group of veterinarians than technicians and producers could influence the results between groups. However, the videos collected during this study could be used to train veterinarians, technicians, and producers to detect BRD clinical signs in the future, as already performed in human medicine (Crawley et al., 2017).

V- CONCLUSION

We found that repeatability between raters was highest when BRD clinical signs were assessed with a 2-scale combination (0,1/2,3 (absence, mild/moderate, severe) for nasal discharge, ocular discharge, and ear droop/head tilt and 0/1,2 (absence/presence) for abnormal respiration and induced cough). Interestingly, among these BRD clinical signs, both induced cough and ear droop/head tilt were the most reliable among raters. In the future, these two BRD clinical signs and rectal temperature could be used to design a simple and repeatable screening scoring system. Future research could determine if this discrepancy could be improved through standardized training. Improving the consistency of rating other clinical signs could also help for better day-to-day BRD monitoring.

CONCLUSION ARTICLE 1: Afin de développer un SCRD simple et répétable, nous avons estimé dans cet ARTICLE 1, la répétabilité de chacun des signes cliniques de BPI par différents professionnels de veaux lourds susceptibles d'utiliser le futur SCRD. Comme nous en avons fait l'hypothèse, la répétabilité des signes cliniques de BPI est variable. La POSITION ANORMALE DES OREILLES (normal/uni ou bilatérale baisse des oreilles ou tête penchée) et la TOUX INDUITE (présence/absence) sont les deux signes les plus répétables à prioriser pour l'élaboration du nouveau SCRD. Des variabilités d'évaluation des signes cliniques sont présentes entre les producteurs, les techniciens et les médecins vétérinaires de veaux lourds, laissant suggérer la nécessité de standardiser et entraîner les professionnels sur la détection des signes cliniques. Détaillons à présent l'élaboration du nouveau SCRD incluant ces signes cliniques répétables dans le prochain article.

ARTICLE 2

Development and validation of a clinical respiratory disease scoring system for guiding treatment decisions in veal calves using a Bayesian framework.

Publié dans le *Journal of Dairy Science*, 2022 Oct 4;

J. Berman,^{1*} D. Francoz,¹ A. Abdallah,^{1,2} S. Dufour,³ and S. Buczinski.¹

¹ From the Département des sciences cliniques, Faculté de médecine vétérinaire, Université de Montréal, Saint-Hyacinthe, QC, Canada.

² From the Department of Animal Medicine, Faculty of Veterinary Medicine, Zagazig University, Zagazig, Egypt.

³ From the Département de Pathologie et Microbiologie, Faculté de médecine vétérinaire, Université de Montréal, St-Hyacinthe, QC, Canada.

Abbreviations: BCI = Bayesian credible interval; BRD = Bovine Respiratory Disease; CaCRSC = Californian score; CRSC = clinical respiratory scoring chart; DNS = deep nasopharyngeal swab; Hap = haptoglobin concentration; HNPV = batch-level negative predictive value; HPPV = batch-level positive predictive value; HSe = batch sensitivity; HSp = Batch specificity; LR- = negative likelihood ratio; LR+ = positive likelihood ratio; MCT = misclassification cost-term; NPV = negative predictive value; PPV = positive predictive value; QcCaCRSC = Quebec score; Se = Sensitivity ; Sp = Specificity ; TRIPOD = Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis; TUS = thoracic ultrasonography; VcCRS new clinical scoring ; VcCRSC= new clinical scoring chart; WiCRSC = Wisconsin score.

ABSTRACT

Objectives: Active infectious bovine respiratory disease (BRD) is an infection of the airways that needs to be diagnosed correctly so appropriate treatment can be initiated. The simplest and most practical test to detect active BRD in dairy calves raised for veal is the detection and interpretation of clinical signs by producers or technicians. However, the clinical scoring system currently available in veal calves lacks sensitivity and specificity contributing to economic losses and high use of antimicrobials. An accurate and reliable

batch-level test to detect active BRD is essential to tailor antimicrobial use and reduce economic losses in veal calves. The objective of this study was therefore to develop and validate a new veal calf respiratory clinical scoring system (VcCRS) including reliable clinical signs (cough, ear droop or head tilt) and increased rectal temperature to detect active BRD in batches of veal calves housed individually and to describe its accuracy for identifying batches of veal calves to treat.

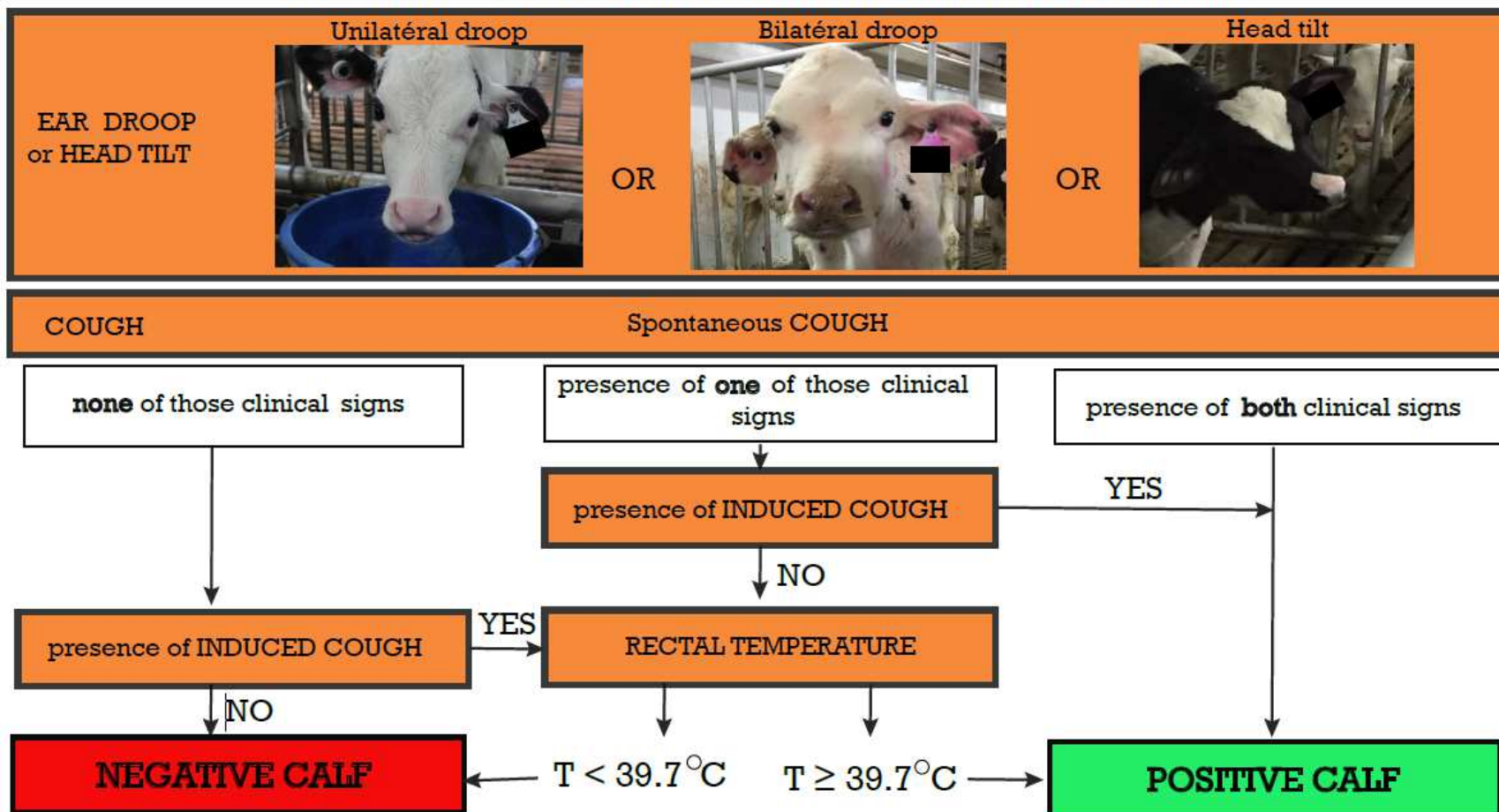
Methods: During 2017-2018, clinical examination, thoracic ultrasonography (TUS) and a haptoglobin concentration (Hap) were prospectively performed on 800 veal calves housed individually in Québec, Canada. Deep nasopharyngeal swab (DNS) was performed on 250 veal calves. A Bayesian latent class model accounting for imperfect accuracy of TUS and Hap was used to obtain weights for the clinical signs and developed the VcCRS. The VcCRS was then validated externally in 3 separate datasets. Finally, the applicability of VcCRS at batch-level was determined.

Results: We found that calves with 2 of the following findings—cough, unilateral/bilateral ear droop or head tilt, or increased rectal temperature $\geq 39.7^{\circ}\text{C}$ —were considered positive and had a 31% chance of having active BRD. Without at least 2 of these 2 findings, a calf had a 100% chance of not having active BRD. At the batch level, we found that a batch with ≥ 3 positive calves among 10 calves sampled 2 weeks after arrival at the fattening unit had a 94% chance of having an active BRD prevalence $\geq 10\%$. A batch with < 3 positive calves had a 95% chance of not having an active BRD prevalence $\geq 10\%$.

Conclusion: In this study, we developed a simple individual and batch-level score that is reliable across examiners and performs effectively in the detection of active BRD in veal calves. The implementation of this VcCRS in the veal calf industry would promote the elaboration of a protocol tailoring antimicrobial use.

Key words: Cattle; bronchopneumonia; diagnostic; latent class model.

VEAL CALF CLINICAL RESPIRATORY DISEASE SCORING SYSTEM



I- INTRODUCTION

Infectious bovine respiratory disease (**BRD**) is a disease of the respiratory tract caused by viruses and bacteria (Woolums, 2015). Most of the time, the infection starts in the upper respiratory tract before descending in the lower respiratory tract and causing inflammatory and lesions to the lung parenchyma (defined in this article as active BRD) (Pancieria and Confer, 2010, Zeineldin et al., 2019). Active BRD needs to be treated with anti-inflammatories, antimicrobials, or both treatment approaches (Woolums, 2015, Buczinski and Pardon, 2020). Once the inflammation and infection are resolved, it is common to observe lung scar tissue (defined here as inactive BRD) (Ollivett et al., 2015). Because dairy calves intended for veal production are commonly commingled during collection, transportation, and housing in a fattening unit, active BRD is highly prevalent in this production system (50% of calves with lung lesions at slaughter) (Leruste et al., 2012) and could be responsible for an important part of mortality (16 to 50% of deaths caused by BRD (Lava et al., 2016a, Winder et al., 2016), lower carcass weight, and lower carcass quality (Van der Mei and Van den Ingh, 1987, Pardon et al., 2013). To prevent and control active BRD in veal calves, antimicrobials are commonly used, representing up to 73% of antimicrobials used during the production cycle (Lava et al., 2016b). These antimicrobials can be given as individual treatment, but the vast majority (> 95%) are given as metaphylactic therapy, i.e., the simultaneous antimicrobial therapy of clinically healthy calves and of calves that have clinical signs of active BRD in a shared compartment (Pardon et al., 2012a). Reducing individual treatment and, above all, optimal targeting of metaphylactic group treatments are key factors to substantially reduce antimicrobial use in the veal industry. Accurate and reliable individual and group-level diagnosis of active BRD is therefore essential to tailor antimicrobial use in this production.

In Canada, the province of Québec is a major player in veal calf production, producing around 80% of Canadian veal calves (Producteurs de Bovins du Québec, 2020). Considering the relatively high number of calves in a fattening unit in Québec (mean = 470 calves) (Producteurs de Bovins du Québec, 2020), the simplest and most practical test to detect active BRD in a veal fattening unit is the detection and interpretation of clinical signs by non-veterinarians (e.g., producers or technicians). However, due to the variety of infectious agents involved in BRD, clinical signs vary in intensity and duration, which can

make clinical diagnosis difficult (McGuirk and Peek, 2014). Additionally, clinical diagnosis in veal calves has been reported to be variable among different examiners (Buczinski et al., 2016, Berman et al., 2021). A clinical scoring system assigns values to each predictor, which are used to determine a total score, thus making it possible to assess disease more objectively than with unstructured clinical evaluation alone (Hayes et al., 2010, Love et al., 2014). A simple, objective, and reliable clinical respiratory scoring chart (**CRSC**) would therefore be a useful tool to improve and standardize active BRD identification in veal calves.

Currently, 3 CRSCs to detect BRD in pre-weaned dairy breed calves have been published. The earliest CRSC used a grading system of 0 to 3 for the following findings and clinical signs: elevated rectal temperature, nasal discharge, cough, ocular discharge, and ear droop or head tilt (**WiCRSC**; McGuirk (2008)). Although this clinical scoring system used weights and decision rules, it did not use quantitative methods to assign these weights. Moreover, WiCRSC was reported to be unreliable between examiners with minimal training (Buczinski et al., 2016) and had only moderate diagnostic accuracy (screening sensitivity (**Se**) and Specificity (**Sp**) of 46-62% and 74-91%, respectively) (Buczinski et al., 2015b, Love et al., 2016). Another CRSC was developed to circumvent these issues in pre-weaned dairy calves (**CaCRSC**; Love et al. (2014)). It was later adapted and evaluated while accounting for the imperfect reference standard definition in pre-weaned dairy calves in Québec (**QcCaCRSC**; Buczinski et al. (2018b)). The CaCRSC and QcCaCRSC used the same clinical signs as WiCRSC but dichotomized them (presence/absence). They also included an additional sign: breathing quality. Moreover, the weights assigned were determined quantitatively. However, the performances of CaCRSC and QcCaCRSC were reported to be similar to that of WiCRSC, with Se and Sp of 47% and 87%, respectively, for CaCRSC (Love et al., 2016), and Se varying from 67% to 83% and Sp of 69% to 83% for QcCaCRSC (Buczinski et al., 2018b). Different reasons could explain these moderate diagnostic performances. Firstly, WiCRSC and CaCRSC's accuracy measures were validated using different imperfect reference tests for comparison, which could increase interpretation variability as a function of both the lesions and the agents involved. Secondly, those scores included clinical signs that do not have high inter-rater reliability

(Buczinski et al., 2016, Berman et al., 2021), which could increase variability across examiners beyond clinical variability and, thus, across studies.

We recently reported the reliability of the respiratory clinical signs (i.e., nasal discharge, cough, ocular discharge, ear droop or head tilt, and respiration) commonly used to detect respiratory disease in veal calves by different types of persons involved in veal calves' health monitoring (producers, technicians, and veterinarians) (Berman et al., 2021). We showed that induced cough (presence or absence) and ear droop or head tilt (absence or presence of slight unilateral ear droop, bilateral ear droop, or head tilt) were the most reliable clinical signs versus all other clinical signs that were assessed (Cohen's kappa \geq 0.6; (Berman et al., 2021)). In presence of active BRD, the inflammation of the trachea causes spontaneous or easily induced cough. Increased coughing frequency when using continuous monitoring of calves sounds algorithm was 99.2% specific and 50.3% sensitive for disease detection in dairy calves (Vandermeulen et al., 2016). Ear droop or head tilt occurred in case of pain, depression or if an otitis is conjointly present. In the latter case, pneumonia is commonly associated with otitis media (Francoz et al., 2004, Gosselin et al., 2012). Adding those clinical signs with the objective measure of rectal temperature investigating fever caused by the infectious process, in order to elaborate a new score would reduce variability due to examiners and therefore, likely increase assessment reliability.

The objective of the current study was therefore to develop and validate a new clinical scoring (**VcCRS**) and chart (**VcCRSC**) including reliable clinical signs to detect active BRD in veal calves housed individually and to describe its accuracy for identifying batches of veal calves to treat metaphylactically.

II- MATERIALS AND METHODS

We proceeded in 3 steps: (1) developing a new diagnostic score (VcCRS), (2) validating the developed score at the individual level using external validation, and (3) modeling the application of the score at the batch level. The **TRIPOD** guidelines (Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis) were used to facilitate reporting on the design, conduct, and results of the current study (Moons et al.,

2015). The study protocol was approved by the Comité d'éthique de l'utilisation des animaux de l'Université de Montréal (17_rech-1898).

II-1- Step 1: Developing a new clinical score (VcCRS): attributing weights to each clinical sign using a quantitative method and choosing a cut-off to detect active BRD (800 veal calves)

II-1-1- Study population

During a prospective cross-sectional study performed from October 1, 2017, to December 20, 2018, a total of 800 veal calves housed individually were randomly recruited from 80 batches (10 calves per batch) of 51 commercial veal calf fattening units (farm) in Québec (Délimax, Veaux Lourds Ltée, Saint-Hyacinthe, Québec, Canada, and Les Aliments Prolacto Inc., Villeroy, Québec, Canada). All fattening units performed an all-in/all-out management. Since the number of calves present in every unit was known before the visit and each box on the farms was numbered, 10 boxes' numbers were randomly chosen prior to the visit (RANDOM function in Excel (Microsoft, Redmond, WA)). Whenever a box was empty due to early mortality, the next random number was selected.

The calves were mostly obtained from multiple local dairy farms after commingling through auction markets (Buczinski et al., 2021). Data were collected within 2 weeks after arrival at the fattening unit. We intervened before the main peak incidence of active BRD in veal calves, generally occurring around 3 weeks after arrival (Miller et al., 1980, Pardon et al., 2012b). At that time, calves were about 3 to 4 weeks old and were housed in individual duckboard pens. Multiple batches of calves (i.e., groups of calves that arrived at the farm to be fattened together) were considered per farm, and the selected batches were distributed across all 4 seasons. Data collection was done in the morning (between 9:00 AM and 11:00 AM), between meals.

II-1-2- Collection of data

Sex (male or female), breed (Holstein, Jersey, Red Holstein, Ayrshire), and treatment received before sampling (antimicrobial, anti-inflammatory, or both treatments) were recorded. On the same day, each calf underwent a physical examination where the main clinical signs hypothesized to be of value for the scoring system were measured. For that

purpose, we selected only the most reliable clinical signs (cough, ear droop or head tilt) determined previously in Berman et al. (2021) and increased rectal temperature as predictors. Moreover, for each calf, thoracic ultrasonography (**TUS**) and a haptoglobin concentration (**Hap**) were conducted successively, as described in the next section. These latter 2 tests were conducted to estimate the true disease status (active or non-active BRD). Finally, a sample of 250 veal calves from the last 25 sampled batches underwent a deep nasopharyngeal swab (**DNS**) to describe the bacteria present in the studied population.

Clinical signs: A complete physical examination was performed on each calf by the same experienced operators (A.A., J.B., or S.B.). The presence of abnormalities upon the physical examination (e.g., diarrhea, navel infection, arthritis, mass, skin or eye disorders) were recorded. The presence of respiratory clinical signs (i.e., nasal discharge, ocular discharge, induced or spontaneous cough, abnormal respiration, and ear droop or head tilt) was recorded based on the 2-level reliable combination reported previously by Berman et al. (2021) and summarized in **Supplementary Methods**. Increased rectal temperature was considered when rectal temperature was $\geq 39.7^{\circ}\text{C}$, as reported in feedlots (Timsit et al., 2011a, Timsit et al., 2011b). Because the physical examination was performed first, the operators were blinded to both reference test results (TUS and Hap).

Deep nasopharyngeal swab: A DNS was taken as previously reported from the right or left nostril (Godinho et al., 2007) and placed in transport media (BBL Port-A-Cul Tubes, Becton, Dickinson and Company, Sparks, MD) for routine bacterial and mycoplasma cultures. All laboratory analyses were performed in a diagnostic laboratory accredited by the American Association of Veterinary Laboratory Diagnosticians (**AAVLD**). Details are available in **Supplementary Methods**.

Reference tests: In the absence of a gold standard to measure active BRD, the true status of active BRD was considered as a latent variable and was investigated by the following 2 reference tests:

- *Thoracic ultrasonography*: Bilateral TUS was performed immediately after the physical examination by the same experienced operators (A.A., J.B., or S.B.), using a 7.5 MHz linear probe (Imago, ECM, Angoulême, France), with the method described by Ollivett and Buczinski (2016) (details in

Supplementary Methods). Presence of active BRD by TUS was considered if the maximal depth of consolidation observed was ≥ 3 cm (TUS positive) (Berman et al., 2019). The operators were blinded to the Hap results, but not to the physical examinations.

- *Haptoglobin concentration*: Blood samples from jugular veins were collected for each calf (redtop tube) immediately after both physical examination and TUS. Within 2 h after collection, samples were centrifuged at $1,500 \times g$ for 15 min, and the serum was stored in aliquots at -20°C until analysis by the University of Guelph Animal Health Laboratory (Guelph, ON, Canada). Before analysis, serum samples were thawed at room temperature. Serum Hap content was measured in duplicate by the hemoglobin binding capacity method (Skinner et al., 1991), using an automated analyzer (Cobas 6000 c 501, Roche Diagnostics, Indianapolis, IN). The lower limit of quantification was $0.3 \mu\text{mol/L}$ (0.03 g/L), the inter-assay CV was 6.5%, and the intra-assay CV was 0.9%. In the absence of an established cut-off in veal calves, presence of active BRD by Hap concentration was considered if the concentration of Hap was $\geq 2.5 \mu\text{mol/L}$ (0.25 g/L), based on the distribution of our data and as reported by Timsit et al. (2009) in feedlots (Hap-positive) (details in **Supplementary Methods**).

II-1-3- VcCRS development

The calf was the unit of interest. Analyses were performed using the following software: SAS, v9.4, Cary, NC, and OpenBUGS, version 3.2.3 rev 1012, MRC, UK.

Assignment of a weight to each clinical sign: Score weights were attributed individually (univariable) for each selected clinical sign (cough, ear droop or head tilt) or increased rectal temperature (3 univariate models) to avoid underestimating the impact of individual signs due to their interaction with other signs. Briefly, we quantified the association between 1 selected clinical sign or increased rectal temperature and the real probability of active BRD using Bayesian univariable mixed logistic regression models where the

dependent variable was the real status of active BRD (a latent unmeasured variable) and the predictor was the selected clinical sign or increased rectal temperature (a measured observation). Using these models, we determined the weight (w) to attribute to each individual dichotomous predictor (j) simply by using the coefficient reported in the model for this predictor. For the 3 different predictors, the true unmeasured probability of active BRD in the i^{th} calf from the j^{th} batch of the k^{th} farm (PrBRD_{ijk}^+) was expressed by the following formula:

$$\text{logit}(\text{PrBRD}_{ijk}^+) = \beta_0 + w * \text{Selected predictor}_{ijk} + \mu_{\text{Batch } jk} + \mu_{\text{Farm } k}$$

Thus, farm (μ_{Farm}) and batch (μ_{Batch}) random intercepts accounted for the data structure, with calves clustered within batches, and batches within farms.

Because of the absence of a gold standard to measure the target condition, we used a latent class logistic regression analysis (**LCM-regression; Figure 5a** and details in **Supplementary Methods**) to associate the observed clinical sign (predictor) with the real status of active BRD (latent variable), following a previously reported framework (McInturff et al., 2004, Buczinski et al., 2018b). The probability that a calf had active BRD (PrBRD_i^+) was estimated using the observed TUS result for that calf and TUS's sensitivity (Se) and specificity (Sp). Briefly, the probability of the i^{th} calf for testing TUS-positive (Pr_{TUS^+}) was described as a function of the probability of the calf being truly affected by the target condition (i.e., active BRD) (PrBRD_i^+) as:

$$\text{Pr}_{\text{TUS}^+}_{ijk} = \text{PrBRD}_{ijk}^+ \times Se_{\text{TUS}} + (1 - \text{PrBRD}_{ijk}^+) \times (1 - Sp_{\text{TUS}})$$

For the i^{th} calf, the latent class variable (active BRD; Y_{ijk}) was assumed to be a Bernoulli event defined by the probability of having active BRD (PrBRD^+):

$$Y_{ijk} \sim \text{Bernoulli}(\text{PrBRD}_{ijk}^+)$$

Choice of priors: In this Bayesian framework, we had to describe the prior available knowledge (as prior distributions) on the intercept (β_0), the clinical sign's coefficient (w), the mean and variance of the batch ($\mu_{\text{Batch } jk}$) and farm ($\mu_{\text{Farm } k}$) random intercepts, and finally, TUS Se and Sp .

Priors on w were determined considering a diffuse range of odds ratios between 0.01 and 100 (back-transformed for use in a logit scale as a range of coefficients between -2 and 4.6), corresponding to the normal distribution $w \sim \text{Normal}(1.4, 1.7)$ and intercept $\varepsilon_j \sim \text{Normal}(1.4, 1.7)$:

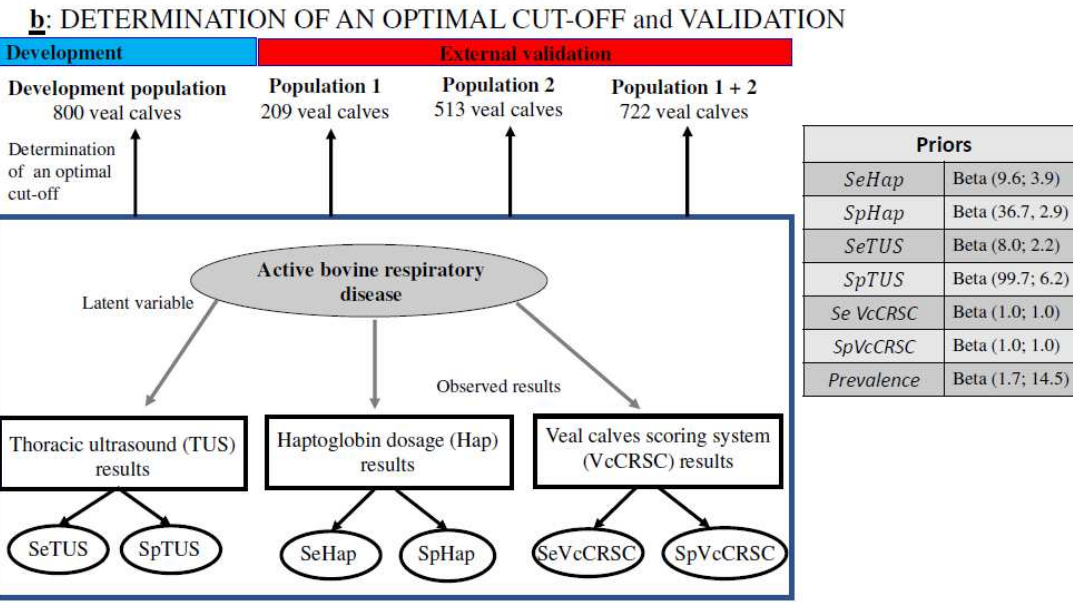
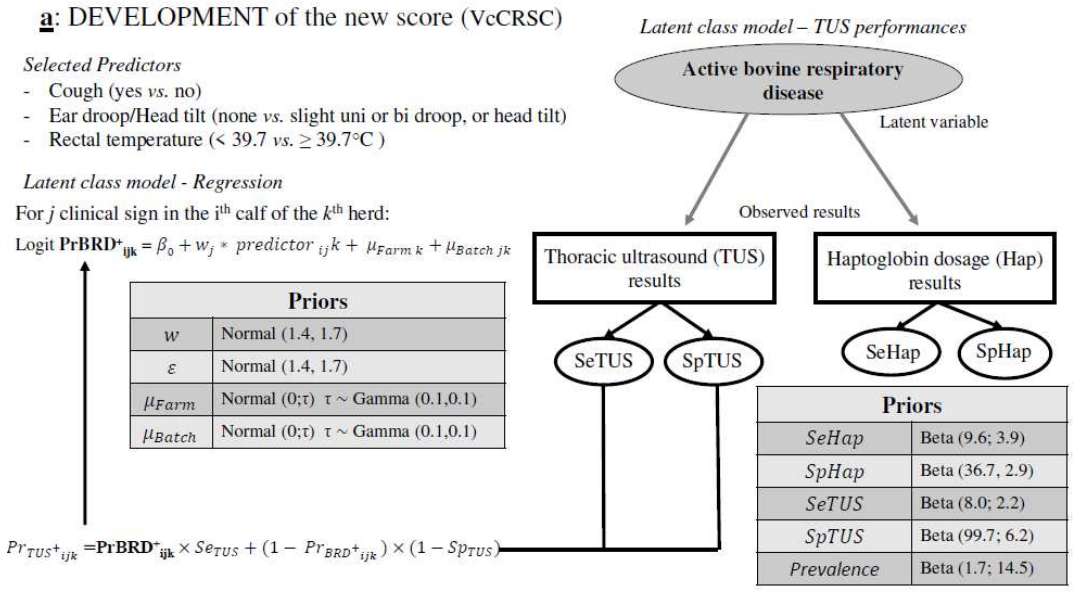
$\mu_{Farm} \sim \text{Normal}(0; \tau)$ and $\tau \sim \text{Gamma}(0.1, 0.1)$

$\mu_{Batch} \sim \text{Normal}(0; \tau)$ and $\tau \sim \text{Gamma}(0.1, 0.1)$

The choice of Gamma (0.1, 0.1) for precision specification was considered as reasonably non-informative (Gelman, 2006).

The choice of Gamma (0.1, 0.1) for precision specification was considered as reasonably non-informative (Gelman, 2006).

In LCM-TUS performances, we assumed that TUS and Hap were conditionally independent because both Hap and TUS assess different biological processes (Berman et al., 2019). Another assumption of the model was that the TUS accuracy was the same for all calves included in the study (Buczinski et al., 2018b). Beta distributions (β) were used as informative priors for all parameters of interest to optimize precision (Se and Sp of TUS, Se and Sp of Hap, and active BRD prevalence) (Dunson, 2001, Gustafson, 2005). The choice of priors is detailed in **Supplementary Methods** and shown in **Figure 5a**.



PrBRD⁺: Probability that calf i has active bovine respiratory disease; w_j : weight coefficient for predictor j ; μ_{Farm} : accounted for data structure with calves clustered within a specific farm; μ_{Batch} : accounted for data structure with calves clustered within specific batches; SeTUS: sensitivity of thoracic ultrasonography; SpTUS: specificity of thoracic ultrasonography; SeHap: sensitivity of haptoglobin; SpHap: specificity of haptoglobin; SeVcCRSC: sensitivity of VcCRSC; SpVcCRSC: specificity of VcCRSC; Prevalence: prevalence of active BRD.

Figure 5. – Diagram of the latent class models (LCM): LCM-Regression and LCM-TUS performances (a); LCM-Validation (b); used for determining and validating the clinical scoring system chart (VcCRS) to detect active bovine respiratory disease (BRD) in veal calves. Priors of each model are shown.

II-1-4- Sample size calculation

The sample size of 800 calves was obtained *a priori* to optimize the accurate estimation of active BRD status by physical examination, TUS, and Hap (i.e., with a 95% confidence interval varying by ± 2.5 -12.5% of Se and Sp for each test) and considering an expected active BRD prevalence of 20%.

II-1-5- From the logistic analysis to the VcCRS

The coefficients obtained from the regression analysis were used to define the specific sign score weight (W). The scores' weights were obtained after rounding logistic regression coefficient parameters multiplied by 10 (Moons et al., 2002, Toll et al., 2008). At the end of the 3 different univariable analyses for the 3 selected predictors, a specific score with the different scores weight was built as follows:

$$\text{Specific score} = W_{\text{Cough}} * \text{Cough (presence-absence)} + W_{\text{Ear droop/head tilt}} * \text{Ear droop/head tilt (slight uni- or bilateral ear droop or head tilt)} + W_{\text{temperature}} * \text{Temperature} (\geq 39.7^{\circ}\text{C})$$

II-1-6- Assessment of model sensitivity to priors

The probability that a calf had active BRD (PrBRD_i^+) was estimated with the posterior distribution of Se and Sp of Hap instead of the Se and Sp of TUS (**Table 30; sensitivity analysis scenario**).

II-1-7- Software

The model was based on a total of 20,000 iterations using a 5,000-iteration burn-in. Three different chains with different initial values were run for each model. Rapid mixing and stationary distribution were sought as signs of good convergence to the posterior distribution. The convergence of the models was checked using visual trace plots and Gelman-Rubin statistic plots. Autocorrelation was detected using autocorrelation plots, and thinning was performed when required. The posterior distributions of each parameter were reported as medians and the corresponding 95% Bayesian credible interval (**BCI**).

II-1-8- Determination of the optimal cut-off

To investigate the optimal cut-off for VcCRS to detect active BRD, the posterior distributions of Se and Sp of VcCRS were computed using another Bayesian LCM (**LCM-**

validation; Figure 5b and details in **Supplementary Methods**) to compare the 3 tests, VcCRS, TUS, and Hap, in the development population across a range of possible cut-off values (0, 6, 9, 10, 15, 16, 19, and 25). The three tests were considered conditionally independent based on previous study and because all tests, VcCRS, Hap and TUS assess different biological processes (Berman et al., 2019). The priors were the same as used in LCM-TUS performances for Se and Sp of TUS, Se and Sp of Hap, and prevalence of active BRD. Non-informative priors were used for Se and Sp of VcCRS corresponding to a β (1.0; 1.0) distribution. Priors are presented in **Figure 5b**.

A misclassification cost-term (**MCT**) analysis was conducted as described by Greiner (1996). Briefly, MCT analysis is a powerful tool to illustrate the robustness of the optimal cut-off because it considers not only the Se and Sp of VcCRS but also the prevalence of the active BRD (Prev) and the cost ratio between false negatives (FN) and false positives (FP). The MCT can be plotted for different cost ratios of FN to FP (r), making it possible to develop a cut-off that accounts for different cost ratios associated with test misclassification. The exact plausible ranges for the relative costs of r are presently unknown for veal calves. We used wide ranges that were obtained in a previous study where 4 different experts were asked to determine this value in feedlot calves (Buczinski et al., 2015a). In the absence of specific studies on veal calves reporting this ratio, we assumed plausible ratios of $r = 1:1$, $r = 3:1$, and $r = 8:1$, indicating that the cost of an FN case is generally higher than that of an FP case (Buczinski et al., 2018b). The MCT was calculated for each specific cut-off using the following formula:

$$\text{MCT} = (1 - \text{Prev}) * (1 - \text{Sp}_{VcCRS}) + r * \text{Prev} * (1 - \text{Se}_{VcCRS})$$

The minimum value of MCT was considered as the value that minimizes the misclassification costs (Greiner, 1996, Buczinski et al., 2018b). To determine the variability of the optimal cut-off with the prevalence, the MCT analysis was performed for different batches' active BRD prevalence scenarios: low (Prev = 5%), medium (Prev = 10%), and high prevalence (Prev = 25%).

II-2- Step 2: External validation of the developed score and estimation of its performance at the individual level using different populations (209, 313, and 722 veal calves)

The VcCRS was validated using 2 separate data sets from previous studies (population 1 (Berman et al., 2019), and population 2 (Morin, 2020)) that measured the same predictors (cough, ear droop or head tilt, and increased rectal temperature) but were sampled previously by other investigators (temporal validation) (Moons et al., 2015). The target condition (active BRD) was measured identically by a Bayesian latent class model in population 1 (TUS and Hap) (Berman et al., 2019), but only with TUS in population 2 (Morin, 2020). Details of the data sets are shown in **Table 28**.

	Development population	Population 1^a	Population 2^b
Calves (number)	800	209	513
Age	3-4 weeks	3-4 weeks	3-4 weeks
Male, No. (%)	756 (94.5)	194 (92.8)	Non available
Breed Holstein	768	188	Non available
Batches (number)	80	1	4
Farms	51	1	2
Sampling dates	October 2017-December 2018	July and October 2015	February to May 2019
Predictors	Nasal discharge	Nasal discharge	Nasal discharge
	Ocular discharge	Ocular discharge	Ocular discharge
	Ear droop/head tilt	Ear droop/head tilt	Ear droop/head tilt
	Abnormal respiration	Abnormal respiration	Cough
	Cough	Cough	
Imperfect reference tests	Thoracic ultrasonography	Thoracic ultrasonography	Thoracic ultrasonography
	Haptoglobin dosage	Haptoglobin dosage	

^a Population from Berman, J., D. Francoz, S. Dufour, and S. Buczinski. 2019. Bayesian estimation of sensitivity and specificity of systematic thoracic ultrasound exam for diagnosis of bovine respiratory disease in pre-weaned calves. *Prev Vet Med* 162:38-45.

^b Population from Morin MP, Dubuc J, Francoz D, Buczinski S. 2020. Randomized clinical trial of intranasal vaccination with Once® PMH IN. BOVILIS® Coronavirus Vaccine placebo on prevalence of lung consolidation diarrhea and health parameters in veal calves. Research report.

Tableau 28. – Demographic and clinical features of the validation populations.

Predictors (cough, increased rectal temperature, and ear droop/head tilt) were measured identically as in the score development population but sampled previously by other investigators (temporal validation). The target condition (active bovine respiratory disease) was estimated by the same reference tests (thoracic ultrasonography and serum haptoglobin concentration) as in the developing population in population 1, but only by thoracic ultrasonography in population 2.

The Se, Sp, positive predictive value (**PPV**), negative predictive value (**NPV**) (i.e., the probability of a calf with or without active BRD receiving a positive or negative test result, respectively), positive likelihood ratio (**LR+**) (i.e., percentage of calves with active BRD having a positive test result (Se), divided by the percentage of calves without active BRD but deemed positive based on the test result (i.e., false positive fraction = $1 - Sp$)), and negative likelihood ratio (**LR-**) (i.e., percentage of calves with active BRD classified as negative based on the test result (i.e., false negative fraction = $1 - Se$), divided by the percentage of calves without active BRD with a negative test result (Sp)) were estimated by LCM analysis (**LCM-validation**; in **Figure 5b**, and details in **Supplementary Methods**) to compare 2 or 3 imperfect tests (TUS, Hap (only population 1), and VcCRS) in 1 or 2 populations. Thus, three LCM were used: 1) a LCM where TUS and Hap were compared to the VcCRS in population 1; 2) a LCM where TUS was compared to VcCRS in population 2; and 3) a LCM where TUS was compared to VcCRS in two distinct populations (1 and 2). The model's priors are shown in **Figure 5b**.

II-3- Step 3: Simulation: Comparing and choosing interpretation for batch-level treatment decision (800 veal calves from 80 batches)

We use the term “batch” or “herd” as a general term for any cluster or aggregate ≥ 100 veal (Québec, 2020). To improve the applicability of the VcCRS in veal production, we estimated VcCRS's batch-level positive predictive value (**HPPV**, i.e., the probability of a positive batch (i.e., with \geq active BRD prevalence threshold) receiving a positive batch-level test result)) and batch-level negative predictive value (**HNPV**, i.e., the probability of a negative batch (i.e., with $<$ active BRD prevalence threshold) receiving a negative batch-level test result)). The active BRD prevalence threshold can vary according to several factors including the calves (price paid in the auction market, quality, origin), the anticipated meat marketing context, specific farm risk factors (ventilation, caretakers), season or the agent involved. We, therefore, decided to report HPPV and HNPV for a range of relevant active BRD prevalence thresholds in veal calves (i.e., 5%, 10%, 15%, 20%, 25%, and 30%) according to personal communication with the veal industry in Québec and a study on a feedlot (Baptiste and Kyvsgaard, 2017). Both HPPV+ and HNPV- were estimated by the following formulas (Christensen and Gardner, 2000):

$$\text{HPPV} = \frac{\text{HTP} * \text{HSe}}{\text{HTP} * \text{HSe} + (1 - \text{HTP}) * (1 - \text{HSp})}$$

$$\text{HNPV} = \frac{(1 - \text{HTP}) * \text{HSp}}{(1 - \text{HTP}) * \text{HSp} + \text{HTP} * (1 - \text{HSe})}$$

Batch-level predictive values depend on the true proportion of positive batches (HTP) (i.e., the pre-test probability that a batch is positive (\geq active BRD prevalence threshold)). Because the true batch prevalence of active BRD is unknown in veal calves, HPV was estimated for a range of HTPs: 5%, 10%, 15%, and 20%. Batch-level predictive values also depend on the batch-level Se of VcCRS (**HSe**; i.e., the probability that a positive batch (i.e., with \geq active BRD prevalence threshold) yields a positive batch-level test result) and the batch-level Sp of VcCRS (**HSp**, i.e., the probability that a negative batch (i.e., with $<$ active BRD prevalence threshold) yields a negative batch-level test result), estimated by the following formulas (Christensen and Gardner, 2000):

$$\text{HSe} = \sum_0^{k-1} \binom{n}{k-1} AP^{k-1} (1 - AP)^{n-(k-1)}$$

$$\text{HSp} = \sum_0^{k-1} \binom{n}{k-1} Sp^{n-(k-1)} (1 - Sp)^{k-1}$$

where AP corresponds to the apparent prevalence, defined as AP = prevalence threshold * Se + (1 - prevalence threshold) * (1 - Sp). The HSe and HSp of VcCRS are dependent on the Se and Sp of the individual test, the number of animals tested (n), the prevalence threshold in infected batches, and the batch cut-off value (k, e.g., 1, 2, or 3 positive test results) used to classify the batch as positive.

We first determined the minimal number of veal calves (n) to sample (10, 20, or 30) in a batch (≥ 100 calves) that optimized HPPV and HNPV, considering the most probable value of HTP at 5% based on the batch prevalence of lung consolidation on dairy farms in Québec (Buczinski et al., 2018a). We hypothesized an active BRD prevalence threshold of 15%, which represents the threshold where the number of veal calves needing treatment is minimal according to Baptiste and Kyvsgaard (2017).

After the minimal number of veal calves was fixed (10, 20 or 30), we determined the batch cut-off value (k) that optimized HPPV and HNPV for the different scenarios of prevalence thresholds and HTPs.

Additional details on the materials and methods, including the statistical models used and the datasets, are available as supplementary materials.

III- RESULTS

III-1- Step1: Development of the VcCRS

III-1-1- Study population

Descriptive data of the 800 veal calves recruited in this study are shown in **Table 29**. Most of the veal calves were male Holsteins. A total of 115 calves (14%) was treated by anti-inflammatory, antimicrobials, or both approaches before collecting data. Calves' treatments before sampling had no impact on the results (assessed by comparing proportion of TUS and Hap positive and proportion of bacteria isolation in treated and untreated calves, data not shown). A total of 46 calves (5%) had another disease (diarrhea, navel infection, arthritis, ringworm, mandibular mass, eye disorders) detected at physical examination. Deep nasopharyngeal results showed that most calves (n = 135; 54%) were carrying *Mycoplasma bovis*. *Pasteurella multocida* and *Mannheimia haemolytica* were isolated in 54 (21.4%) and 3 (1.2%) samples.

Categorical variables		Number of calves	Proportion of calves (%)
Batches	Milk fed (white)	610	76.3
	Grain fed (Red)	190	23.8
Season	Fall	320	40
	Winter	90	11.3
	Spring	240	30
	Summer	150	18.8
Breed	Holstein	769	96
	Jersey	5	0.63
	Red Holstein	18	2.3
	Ayrshire	8	1
Sex	Female	44	5.5
	Male	756	94.5
Treatment before sampling*	Yes	115	14.4
	No	683	85.6
Co-morbidity	Diarrhea	36	4.5
	Navel infection	4	0.5
	Arthritis	1	0.13
	Ringworm	1	0.13
	Mandibular mass	1	0.13
	Eye disorders	2	0.25
	None	754	94.3
Bacteria isolated from a deep naso-pharyngeal sample (250 calves)	<i>Mycoplasma bovis</i>	135	54
	<i>Pasteurella multocida</i>	54	21.4
	<i>Mannheimia spp</i> **	28	11.2
	<i>Mannheimia varigena</i>	10	4
	<i>Mycoplasma spp.</i> ***	9	3.6
	<i>Mannheimia hemolytica</i>	3	1.2
	<i>Trueperella pyogenes</i>	1	0.4
	<i>Bibersteinia trehalosi</i>	1	0.4
	> 1 bacterium	41	16.4
	> 2 bacteria	2	0.8
	Negative	57	22.8
	Contaminated	3	1.2
Missing	2	0.8	

* Antimicrobials, anti-inflammatory, or both; ** Other than *Mannheimia hemolytica*

*** Other than *Mycoplasma bovis*

Tableau 29. – Descriptive data (milk/grain fed batches, season, breed, sex, treatment before sampling, co-morbidity and bacteria isolated from a deep naso-pharyngeal sample) of the random 800 veal calves included in the study. Data collection was performed 2 weeks after arrival at the fattening unit. Calves were around 3 to 4 weeks old.

III-1-2- Descriptive statistics

Data collection was possible on each of the 800 recruited calves.

Clinical signs (predictors): The proportion of clinical signs in the 800 veal calves was: nasal discharge (n = 77; 9.6%); cough (n = 147; 18.4%); ocular discharge (n = 142; 17.8%); ear droop or head tilt (n = 35; 4.4%); abnormal respiration (n = 32; 4.0%). A total of 39 calves had increased rectal temperature (n = 39; 4.9%). The frequencies of calves with 1 of the 3 predictors included in the VcCRS development (i.e., ear droop or head tilt, cough, and increased rectal temperature), TUS positive, and Hap positive are detailed in **Figure 6**.

Reference standard tests: TUS was performed on each calf successfully. A total of 110 calves (13.8%) had lung consolidations ≥ 3 cm and were deemed positive on TUS. Five samples were not analyzed for determining Hap concentration (insufficient quantity or missing identification). Consequently, five calves had Hap missing data. Among these five calves, two were TUS positive. A total of 133 calves (16.7%) had Hap ≥ 2.5 $\mu\text{mol/L}$ (0.25 g/L) and was deemed positive on Hap. A total of 28 calves were TUS and Hap positives, 105 calves were TUS negative but Hap positive, 80 calves were TUS positive but Hap negative, and 582 calves were TUS and Hap negatives.

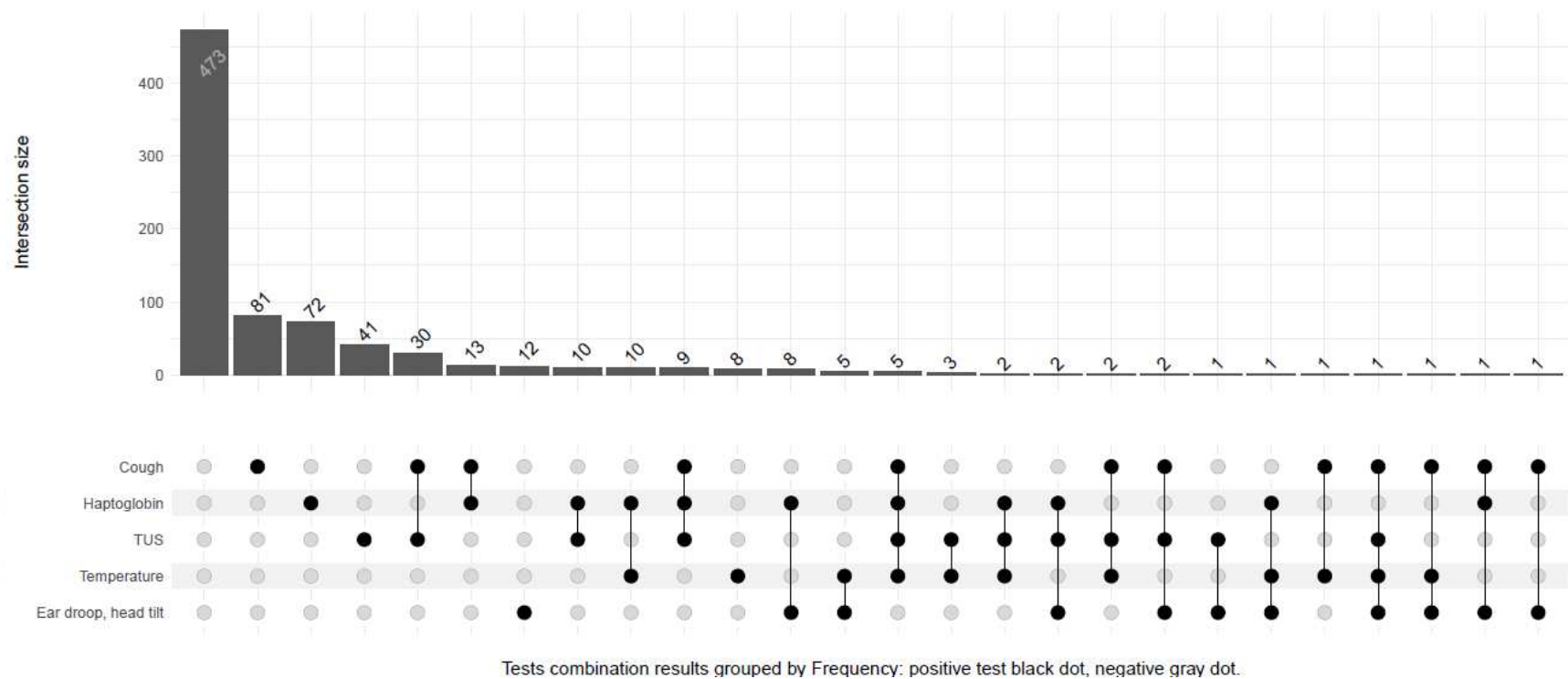


Figure 6. – Upset plot showing the total set size and overlaps between the presence of the 3 predictors included in the developed score (ear droop/head tilt, cough, and increased rectal temperature), TUS results, and haptoglobin results in 795 veal calves. The number of positive tests is indicated on the y-axis. The shaded circles connected by solid lines in the lower panel show the intersecting positive test datasets.

III-1-3- Final VcCRS

The posterior distribution Se and Sp of TUS were estimated at 76% (95%BCI: 42; 96%) and 90% (95%BCI: 87; 95%), respectively. However, the convergence of LCM-regression was possible when lower limit of Se of TUS was truncated to 60%, compatible with what we know on Se of TUS in the literature (Rabeling et al., 1998, Ollivett et al., 2015, Berman et al., 2019). The posterior distributions of the weight coefficients for cough, ear droop or head tilt, and increased rectal temperature are detailed in **Table 30**.

Main model					
	<i>Priors</i>	<i>Posteriors distributions</i>			
Se of TUS	β (8.0; 2.2)	0.76 (0.42; 0.96) (0.60;0.96)*			β (5.7, 2.5)
Sp of TUS	β (99.7; 6.2)	0.90 (0.87; 0.95)			β (316.6, 36.1)
Regression parameters					
		Temperature	Ear droop/head tilt	Cough	Weight
β temperature	N (1.4, 1.7)	0.59 (-0.55; 1.7)			6
β ears droop/head tilt	N (1.4, 1.7)		0.90 (-0.25; 1.98)		9
β cough	N (1.4, 1.7)			0.96 (0.14;1.73)	10
Intercept	N (1.4, 1.7)	-2.2 (-2.9; -1.7)	-2.2 (-2.9; -1.7)	-2.3 (-3.0; -1.8)	-
Sensitivity analysis scenario					
	<i>Priors</i>	<i>Posteriors distributions</i>			
Se of Hap	β (9.6; 3.9)	0.62 (0.34; 0.87) (0.60; 0.87)			β (5.9, 4.0)
Sp of Hap	β (36.7; 2.9)	0.87 (0.85; 0.91)			β (764.5, 115.1)
Regression parameters					
		Temperature	Ear droop/head tilt	Cough	Weight
β temperature	N (1.4, 1.7)	0.43 (-0.58; 1.4)			4
β ears droop/head tilt	N (1.4, 1.7)		0.78 (-0.23; 1.77)		8
β cough	N (1.4, 1.7)			0.79 (0.17;1.44)	8
Intercept	N (1.4, 1.7)	-1.7 (-2.4; -1.2)	-1.7 (-2.35; -1.24)	-1.8 (-2.5; -1.35)	-

* The distribution of Se was truncated at 0.60 to promote convergence of Bayesian latent class model

Tableau 30. – Posterior median distributions and 95% Bayesian credible intervals (BCI) regression coefficients and their weights for increased rectal temperature, ear droop/head tilt, and cough using the sensitivity (Se) and specificity (Sp) of thoracic ultrasonography (TUS) (Main Model) or the Se and Sp of haptoglobin concentration (Hap) (Sensitivity analysis scenario).

The posterior distribution of Se and Sp of Hap were estimated at 62% (95%BCI: 34; 87%) and 87% (95%BCI: 85; 91%), respectively. The weight coefficients' posterior distributions were similar from those of the main model (overlaps of 95%BCI) (**Table 30; sensitivity analysis scenario**).

Finally, we obtained the following VcCRS:

$$\text{Specific score} = 10 * \text{Cough} + 9 * \text{Ear} + 6 * \text{Fever}$$

Where Cough takes a value of 1 if induced or spontaneous cough is present and 0 otherwise, Ear takes a value of 1 if a slight uni- or bilateral droop or head tilt is present and 0 otherwise, and finally, Fever takes a value of 1 if rectal temperature $\geq 39.7^{\circ}\text{C}$ and 0 otherwise.

III-1-4- Optimal cut-off

The MCT values for all the possible cut-off values of the VcCRS and for $r = 1:1$, $r = 3:1$, and $r = 8:1$, are shown in **Figure 7**. Whatever the value of r , the MCT values were minimized for a cut-off of 15 in our study population with a prevalence of 5%. The cut-off of 15 was also optimal for $r = 1:1$ and $r = 3:1$ for a prevalence of 10%. For a prevalence of 25%, the cut-off of 15 remained acceptable for $r = 1:1$ but should be considered lower at 6 for a $r = 3:1$ and $r = 8:1$ (**Figure 7**). We noticed from a practical perspective that using a cut-off of 15, only 2 predictors are sufficient to optimize the detection of active BRD. Thus, a calf with 2 predictors among these 3 following predictors (increased rectal temperature, ear droop or head tilt, and cough) could be considered as VcCRS positive, that minimizing MCT in most scenarios.

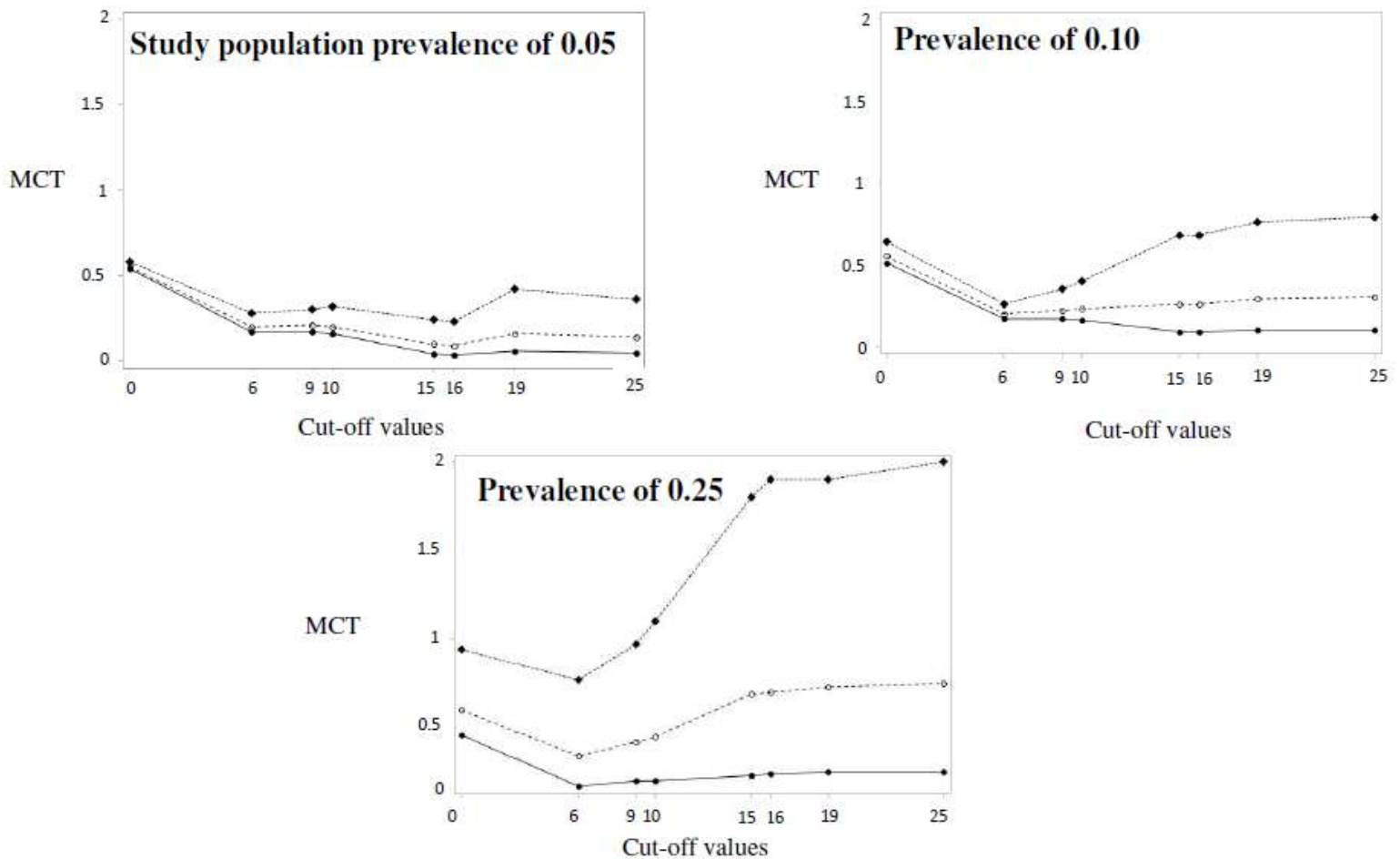


Figure 7. – Misclassification cost-term (MCT) of the cut-offs used to detect active bovine respiratory disease in a population of 800 veal calves. A Bayesian latent class model was used to estimate the sensitivity of the clinical respiratory scoring system (CRS) (SeVcCRS), its specificity (SpVcCRS), and its prevalence (π). MCT was calculated with the following formula: $MCT = (1-\pi) * (1-SpVcCRS) + r*\pi * (1-SeVcCRS)$. We assumed the plausible ranges for the relative costs of r (the false negative to false-positive ratio) $r = 1:1$ (continuous line, filled circle), $r = 3:1$ (dashed line, empty circle), and $r = 8:1$ (dashed line, black diamond). The minimum value of MCT can be considered as the value that minimizes the costs.

III-2- Step2: External validation of the VcCRS

The accuracy parameters (Se, Sp, PPV, NPV, LR+, and LR-) of VcCRS in the developing population and in both external populations are shown in **Table 31**. The performances remained similar in both population 1 and population 2. When we merged both external populations data sets (population 1 + population 2), the Se and Sp of VcCRS were 31% (95%BCI: 14; 70%) and 100% (95%BCI: 99; 100%). These latter values were further defined as the performances of VcCRS at the individual level.

	Development population	Population 1^a	Population 2^b	Population 1 and 2[*]
Se_VcCRSC	0.30 (0.12; 0.72)	0.51 (0.22; 0.87)	0.33 (0.07; 0.93)	0.31 (0.14; 0.70)
Sp_VcCRSC	0.99 (0.98; 1.0)	1.0 (0.97; 1.0)	0.99 (0.98; 1.0)	1.0 (0.99; 1.0)
PPV_VcCRSC	0.53 (0.25; 0.85)	0.86 (0.39; 0.99)	0.58 (0.16; 0.96)	Pop1: 0.89 (0.64; 0.99) Pop 2: 0.65 (0.26; 0.97)
NPV_VcCRSC	0.97 (0.92; 0.99)	0.97 (0.93; 1.0)	0.98 (0.94; 1.0)	Pop1: 0.94 (0.87; 0.99) Pop 2: 0.98 (0.95; 1.0)
LR+_VcCRSC	27.1 (7.8, 110)	104.8 (14.2; 3026)	59.1 (7.6; 813)	74.3 (18.3; 853)
LR-_VcCRSC	0.70 (0.28; 0.89)	0.49 (0.14; 0.79)	0.68 (0.07; 0.94)	0.69 (0.30; 0.87)
Se_TUS	0.80 (0.54; 0.96)	0.88 (0.68; 0.98)	0.79 (0.50; 0.96)	0.86 (0.64; 0.97)
Sp_TUS	0.90 (0.87; 0.93)	0.92 (0.89; 0.96)	0.95 (0.93; 0.98)	0.95 (0.93; 0.98)
Se_Hap	0.64 (0.40; 0.85)	0.72 (0.55; 0.86)	-	-
Sp_Hap	0.87 (0.84; 0.90)	0.95 (0.92; 0.98)	-	-
Prevalence	0.04 (0.02; 0.10)	0.08 (0.04; 0.15)	0.03 (0.006; 0.07)	Pop1: 0.09 (0.04; 0.16) Pop 2: 0.03 (0.007; 0.06)

^a Population from Berman, J., D. Francoz, S. Dufour, and S. Buczinski. 2019. Bayesian estimation of sensitivity and specificity of systematic thoracic ultrasound exam for diagnosis of bovine respiratory disease in pre-weaned calves. *Prev Vet Med* 162:38-45.

^b Population from Morin MP, Dubuc J, Francoz D, Buczinski S. 2020. Randomized clinical trial of intranasal vaccination with Once® PMH IN. BOVILIS® Coronavirus Vaccine placebo on prevalence of lung consolidation diarrhea and health parameters in veal calves. Research report.

* We used for this analysis a LCM where two independent populations (1 and 2) were used, but for which we assumed similar sensitivity and specificity across populations.

Tableau 31. – Posterior median and 95% credible intervals based on 4 Bayesian latent class models (LCM) for the prevalence of active bovine respiratory disease, sensitivity (Se) and specificity (Sp), positive (PPV) and negative (NPV) predictive values, positive (LR+) and negative (LR-) likelihood ratios of the clinical respiratory score (VcCRS), and Se and Sp of thoracic ultrasound (TUS) and concentration of haptoglobin (Hap) used to diagnose active bovine respiratory disease.

III-3- Step 3: Application at batch level

The determination of the minimal number of calves to sample (*n*) is shown in **Figure 8**.

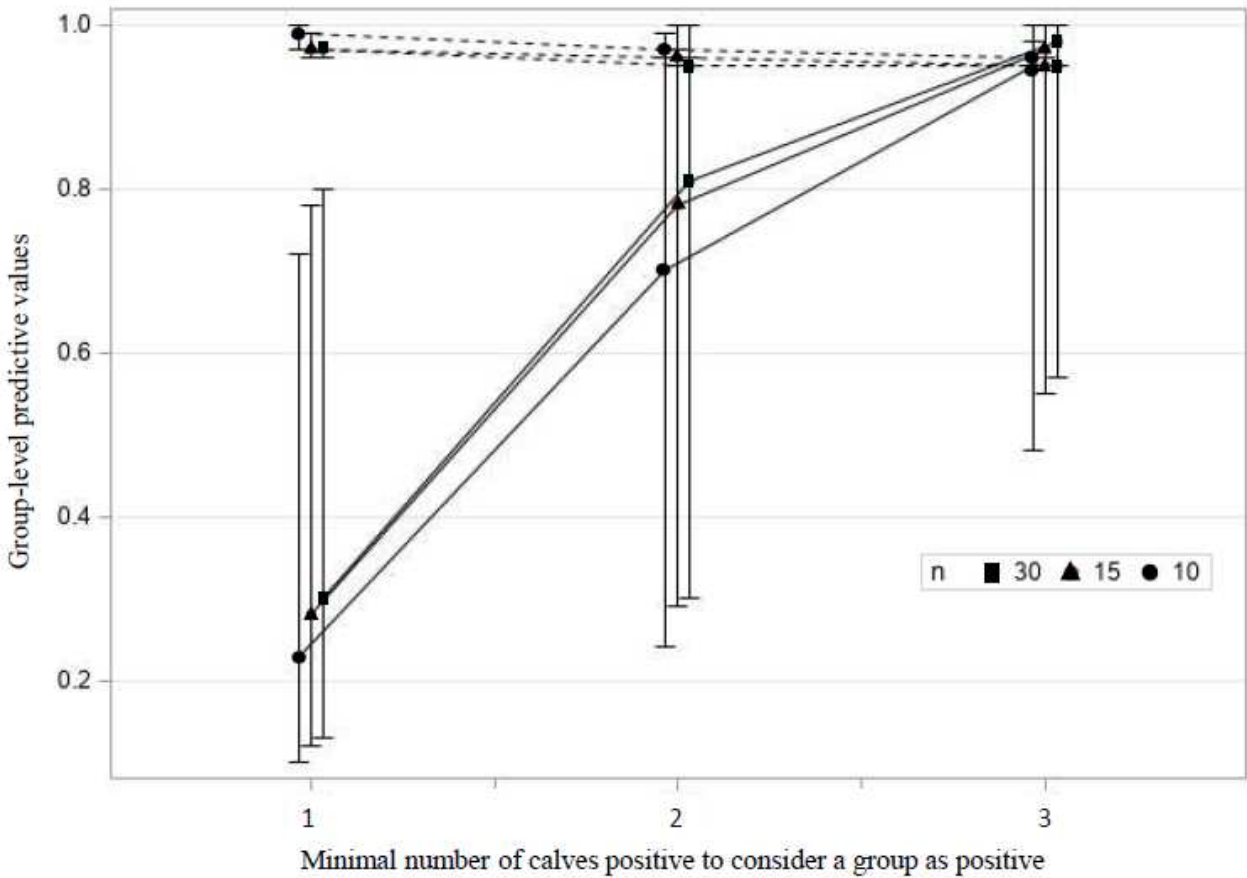


Figure 8. – Batch-level positive (continued line) and negative (dashed line) predictive values for a bovine respiratory disease prevalence threshold of 15%, considering the true proportion of positive batches of 5%, for different batch cut-off values (k) and numbers of sampled calves (n).

For an HTP of 0.05 and an active BRD prevalence threshold of 15%, there was no difference in HPPV and HNPV between sampling 30, 20, or 10 calves, for a k of 3. Because it is practical to sample fewer calves as possible in a fattening unit, the optimal number n was therefore considered at 10 random calves to screen in the batch to be assessed.

The different values of herd predictive values for the different ranges of HTP, prevalence threshold, and k are presented in Figure 5. We noticed that whatever prevalence threshold

and HTP, the optimal k was 3 except for a prevalence threshold of 5% where the optimal k was 4, meaning that 3 or more positive calves out of 10 assessed was a practical way to determine batch level positivity (**Graphical Abstract**).

IV- DISCUSSION

The objective of this study was to develop and validate a new CRSC to detect active BRD in veal calves housed individually, and to describe its accuracy for identifying individuals and groups of veal calves to treat. We found that calves with 2 of the following findings—cough, unilateral/bilateral ear droop or head tilt, or rectal temperature $\geq 39.7^{\circ}\text{C}$ —were considered as positive, with an individual Se and Sp of 31% (95%BCI: 14; 70%) and 100% (95%BCI: 99; 100%), respectively, for active BRD detection. At the batch level, we found that ≥ 3 positive calves among 10 calves sampled 2 weeks after arrival at the fattening unit allowed us to detect batches with active BRD prevalence $\geq 10\%$ with positive and negative batch-level predictive values $\geq 94\%$ and $\geq 95\%$, respectively.

Interestingly, our score differs from the previously developed CRSCs in pre-weaned dairy breed calves by the choice of clinical signs included (McGuirk, 2008, Love et al., 2014, Buczinski et al., 2018b). Previously, all reported respiratory clinical signs (i.e., nasal discharge, ocular discharge, cough, ear droop or head tilt, abnormal respiration) and increased rectal temperature were grouped for the development of a CRSC (McGuirk, 2008, Love et al., 2014, Buczinski et al., 2018b). In contrast, we first selected the clinical signs to include in the CRSC according to their reliability when used by different persons involved in veal calves' health monitoring (producers, technicians, and veterinarians) (Berman et al., 2021). This a priori selection has the advantages of reducing most operator variability from the CRSC used and being applied in a population-based approach, where the first-line diagnosis is generally not performed by a veterinarian.

Additionally, this selection of robust clinical signs limited the number of predictors to assess (only 3 predictors vs. 5 or 6 in WiCRSC and (Vc)CaCRSC, respectively), which could have impacted the Se of VcCRS (Se of 31%). Moreover, this selection implied that the 3 predictors included in VcCRS are less specific of lower respiratory tract disorder than other clinical sign such as dyspnea. However, comparison of Se and Sp of VcCRS and VcCRS + dyspnea did not show any difference on Se and Sp (Data shown in

Supplementary methods). Therefore, this selection simplifies the VcCRS's use in the context of veal calves in a large fattening unit. Even simpler, we showed that the presence of only 2 findings among cough, ear droop or head tilt, and increased rectal temperature was necessary to consider a calf positive. After this initial assessment, induced cough and/or increased rectal temperature are performed a second time, which limits the handling of calves and the potential biosecurity issues (e.g., transmission of disease). Interestingly, we showed that randomly sampling 10 calves from a batch of calves (e.g., batches varying from 100 to 800 calves) was sufficient to apply our score at the batch level, limiting time consumption. The VcCRS is therefore an easy, simple, and quick score to use on the group level in the veal calf industry.

Estimating VcCRS performances at both the individual and the group level contrasts with previously reported CRSCs. Using VcCRS at the group level had the advantage to know if a group of calves need to be treated or not with antimicrobials and would help to rationalize the metaphylactic treatment approach (Baptiste and Kyvsgaard, 2017). We showed that 3 calves with a positive individual VcCRS score among 10 randomly sampled in a group allowed us to detect batches with active BRD prevalence $\geq 10\%$, with batch-level predictive values > 0.90 . Concretely, if we considered the prevalence of risk as defined at 21% in Leruste et al. (2012), being VcCRS-positive at the batch level (i.e., 3 individual VcCRS positive calves among 10 sampled) would imply that the batch has a 99% chance of having active BRD prevalence $\geq 21\%$; otherwise, the batch has a 95% chance of not having active BRD prevalence $\geq 21\%$. It is, however, difficult to define a universally accepted active BRD prevalence threshold in veal calves, since it may depend on various factors, including the calves (price paid in the auction market, quality, origin), the anticipated meat marketing context, specific farm risk factors (ventilation, caretakers), season or the agent involved. That is why we reported the VcCRS performances at the batch level for a range of scenarios of active BRD prevalence in this study (**Figure 9**). With this approach, one could refer to this figure to adjust their intervention in a specific group under specific circumstances. At a predefined active BRD prevalence threshold, our score could be used in the future to guide accurate metaphylactic group treatments and, ultimately, reduce antimicrobial use (Baptiste and Kyvsgaard, 2017).

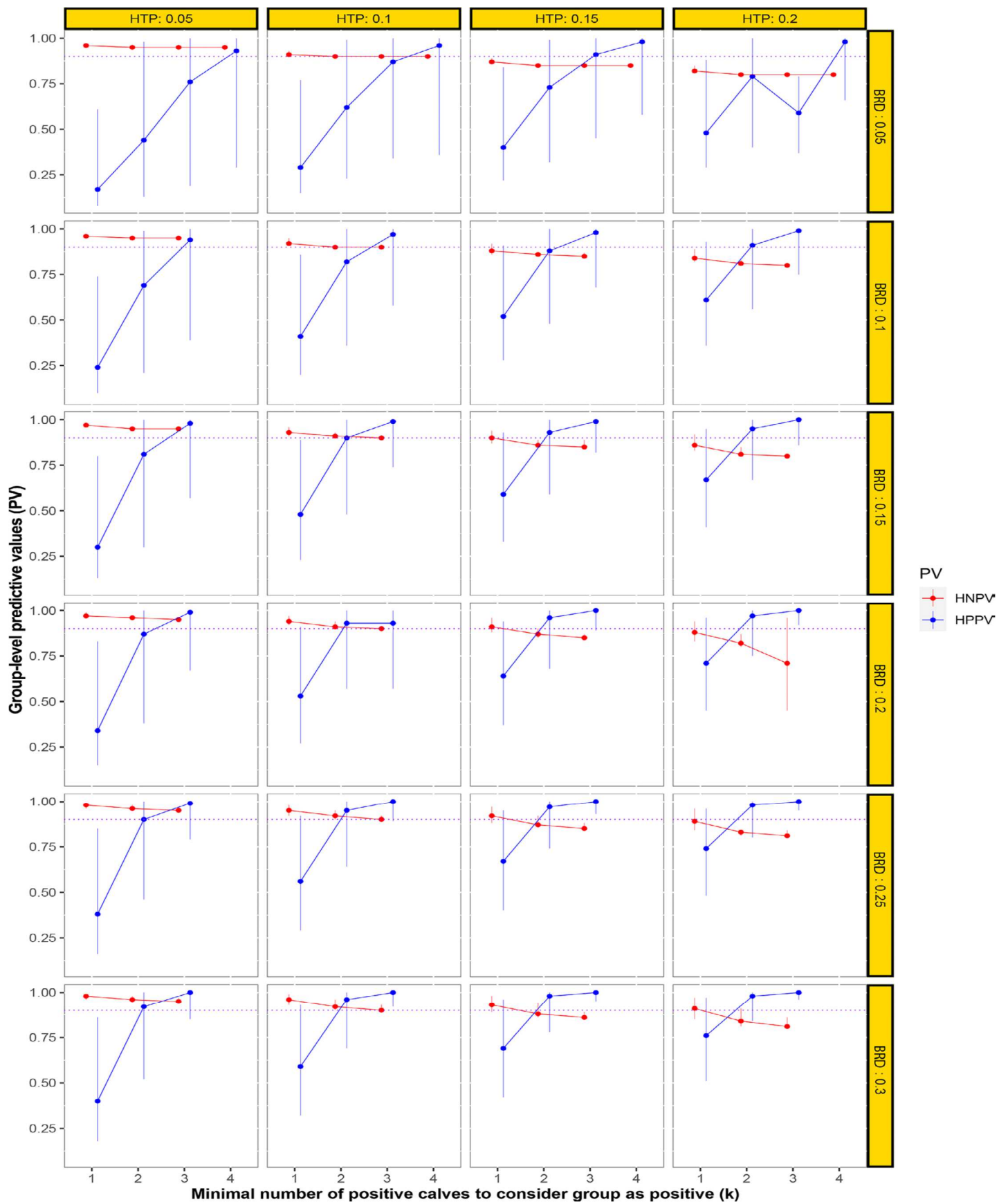


Figure 9. – Batch level positive (HPPV) and negative (HNPV) predictive values for a minimal number of positive calves to consider batch as positive (k), true proportion of positive batches (HTP) and active bovine respiratory disease (BRD) prevalence when the clinical respiratory chart score (VcCRSC) is applied on 10 veal calves in a batch. The dashed line represents the 0.9 value (previous page).

At the individual level, the Se of VcCRS was 31% and, therefore, lower than the Se of previous CRSCs (range: 46-83%) (Buczinski et al., 2015b, Love et al., 2016, Buczinski et al., 2018b). However, the individual Sp of VcCRS was almost perfect and higher than in previous studies (range: 74-91%) (Buczinski et al., 2015b, Love et al., 2016, Buczinski et al., 2018b). The almost perfect Sp of VcCRS had the advantage of limiting false positive calves and, therefore, useless individual antimicrobial treatment (Theurer et al., 2015). The selection of reliable clinical signs and the better definition of respiratory disease as active BRD could explain the Sp superiority of VcCRS versus previous clinical scores in individual pre-weaned dairy breed calves (Buczinski et al., 2015b, Love et al., 2016, Buczinski et al., 2018b). Although the individual performances seem low, with Se of 31%, the goal of this study was to implement a diagnostic tool to target treatment not at the individual level but at the group level (metaphylactic). The high Sp of VcCRS permits us to have a high HPPV at the group level while avoiding false positive batches and, therefore, the unnecessary treatment of the entire batch.

Our study differs from others that developed CRSCs in its designs following a robust statistical approach, as recommended by medical guidelines (Moons et al., 2015). Firstly, we externally validated our score, in contrast with previous CRSCs in calves (McGuirk, 2008, Love et al., 2014). This validation increases the robustness, credibility, applicability, and generalization of VcCRS across veal calf settings. Secondly, WiCRSC and CaCRSC used the combination of multiple tests (a composite reference test) to elaborate their scores and, therefore, used an inaccurate definition of respiratory disease to define cases (McGuirk, 2008, Love et al., 2014). In this study, we used Bayesian LCM, as recommended by OIE (World Organization for Animal Health) in the absence of a gold standard to limit this classification bias and ensure better accuracy estimation (Cheung et al., 2021).

Spectrum bias was also limited by selecting prospectively random calves from commercial veal facilities that represented wide spectrums of active BRD severity (ranging from healthy calves to calves with mild and moderate active BRD to calves with severe active BRD) (Buczinski and O'Connor, 2016). In contrast, the case-control designs used in McGuirk (2008) and Love et al. (2014) may have included cases at the ends of the spectrum of disease (either healthy calves or calves with severe active BRD), resulting in an overestimation of their scores' coefficients.

This study has some limitations, however. Firstly, the posterior median active BRD prevalence was low (5% vs. 20% expected for sample size calculation). This low prevalence is likely because we sampled calves before the main peak of active BRD incidence to promote early detection and limit the proportion of treated calves. Consequently, sensitivity 95%BCI estimates were wide. This uncertainty prevented us from adequately estimating the regression coefficients and forced us to truncate the lower limit of Se of the 95%BCI of TUS to promote convergence of our statistical model. However, this low active BRD prevalence implied large variation in the individual CRSC Se (median value of 31%, but 95%BCI from 14 to 70%) but not in accurately estimated Sp (median of 100%, with 95%BCI from 99 to 100%), which is reported to have the most impact on active BRD diagnosis (Theurer et al., 2015). Secondly, we chose by convenience a temporal external validation, i.e., we used participant data collected by the same investigators, using the same predictors and target condition definitions and measurements, but sampled from an earlier period (Moons et al., 2015). Transportability from one setting to another could have been increased by using a database collected by other investigators in another country (broad validation), in other types of patients (dairy calves), or considering the probable influence of infectious agents (high prevalence of *M. bovis* in this study). However, our study represents the first step in VcCRS's development, and other external validations could be performed in the future to tailor its use in other specific contexts.

V- CONCLUSION

In this study, we developed a simple batch-level score that is probably reliable across examiners and performs effectively in the detection of active BRD in veal calves.

According to the requirement of the industry, the active BRD prevalence threshold could be defined. The positive VcCRS group could then be treated metaphylactically or investigated. The implementation of this VcCRSC in the veal calf industry would promote the elaboration of a protocol tailoring antimicrobial use. A prospective cohort study comparing antimicrobial use and production outcomes of groups applying VcCRSC (exposed) and groups without (non-exposed) would be the next step to validate this CRSC and justify its large-scale implementation in veal calves.

SUPPLEMENTARY METHODS

I- Physical examination

Respiratory examination was performed and interpreted according Berman (2021). Briefly, nasal discharge was considered abnormal when bilateral, cloudy, mucopurulent or excessive mucus discharge was present; ocular discharge was considered abnormal when bilateral heavy or moderate ocular discharge was present; ear droop/head tilt was considered abnormal when slight unilateral droop, head tilt or bilateral droop was present; cough was considered abnormal when induced or spontaneous cough was present; and respiration was considered abnormal when tachypnea or dyspnea was present.

II- Deep nasopharyngeal swab

All conventional and mycoplasma bacteriologic cultures were performed at the Laboratoire d'Épidémiologie-surveillance animale du Québec of the Ministère de l'Agriculture, des Pêcheries et de l'Alimentation du Québec. Deep nasopharyngeal swabs were refrigerated, and conventional culture and mycoplasma culture was done within 4–6 hours of sampling.

III- Thoracic ultrasonography

A solution of 70% isopropyl alcohol was sprayed on the area of interest to improve image quality, and the probe was directly applied on the thorax without clipping. The following areas were scanned: the mid to ventral portion of the right and left lungs caudal to the heart (the caudal sites; from the 11th to the 3rd intercostal space) and the parenchyma cranial to the heart on the right side (the cranial site; from the 3rd to the 1st intercostal space). The maximal depth of consolidation on TUS (in cm) for each site (left caudal, right caudal,

cranial) was recorded and calculated by manual count, using the 1 cm² grid of the ultrasound unit (Ollivett and Buczinski, 2016, Berman et al., 2019).

IV- Latent class model analyses (LCM)

IV-1- LCM-TUS performances

Estimation of Se and Sp of the TUS to detect active bovine respiratory disease (BRD) in the 800 recruited veal calves (two tests, one population).

Assumption: We assumed that both TUS and Hap were conditionally independent, and that the accuracy of TUS was the same for all calves included in the study.

Latent class model elaboration: two independent tests, in one population: The data, $x = (x_{11}, x_{10}, x_{01}, x_{00})$, corresponded to the cross-classified test results for the n calves tested from the population where the vector x_{ab} represented the number of calves that tested positive (a or $b = 1$) or negative (a or $b = 0$) in both tests. Let $p = (p_{11}, p_{10}, p_{01}, p_{00})$ represented the respective probabilities from vector x . The multinomial sampling distribution was given by:

$$x \sim \text{multinomial}((p_{11}; p_{10}; p_{01}; p_{00}), n)$$

where the multinomial cell probabilities were given by:

$$\begin{aligned} p_{11} &= p(\text{TUS+}; \text{Hap+}) = \pi * (\text{SeTUS} * \text{SeHap}) + (1-\pi)((1 - \text{SpTUS})(1- \text{SpHap})) \\ p_{10} &= p(\text{TUS+}; \text{Hap-}) = \pi * (\text{SeTUS} * (1-\text{SeHap})) + (1-\pi)((1 - \text{SpTUS}) * \text{SpHap}) \\ p_{01} &= p(\text{TUS-}; \text{Hap+}) = \pi * ((1-\text{SeTUS}) * \text{SeHap}) + (1-\pi)(\text{SpTUS} * (1-\text{SpHap})) \\ p_{00} &= p(\text{TUS-}; \text{Hap-}) = \pi * ((1-\text{SeTUS}) * (1-\text{SeHap})) + (1-\pi)(\text{SpTUS} * \text{SpHap}) \end{aligned}$$

TUS+: Calves with lung consolidation depth ≥ 3 cm
TUS-: Calves with lung consolidation depth < 3 cm
Hap+: Calves with a concentration of Hap ≥ 0.25 g/l
Hap-: Calves with a concentration of Hap < 0.25 g/l
 π : Prevalence of active BRD

SeTUS: Sensitivity of TUS
SeHap: Sensitivity of Hap
SpTUS: Specificity of TUS
SpHap: Specificity of Hap

Priors: The unknown parameters consisted of SeTUS, SeHap, SpTUS, SpHap, and the Prev in one population, for a total of 5 unknown parameters. The degree of freedom was equal to $2^{\text{number of test}} - 1 = 2^2 - 1 = 3$ (Dendukuri et al., 2009). To optimize accuracy of posterior distribution of TUS and Hap, accurate priors were given for all parameters (Dunson, 2001, Gustafson, 2005).

The most probable value of active BRD prevalence and the range containing the true value with 95% confidence were obtained from the literature on veal calves (Pardon et al., 2012b). Given the inaccuracy in definitions of active BRD, relatively diffuse informative priors were used, with a mode at 0.05 and 95% certainty that prevalence was ≤ 0.25 , corresponding to a β (1.7; 14.5) distribution. We used informative prior distributions for Hap and TUS test accuracy parameters based on available literature (Timsit et al., 2009, Berman et al., 2019). The most probable value for Hap Se was determined to be 0.75, while we were 95% sure that it was ≥ 0.50 , corresponding to a β (9.6; 3.9) distribution. For Hap Sp, a mode of 0.95 was hypothesized, with 95% certainty that $Sp \geq 0.85$, corresponding to a β (36.7; 2.9) distribution (Se and Sp calculated by Abdallah et al 2016 for Timsit et al., 2009). The most probable value for TUS Se was determined to be 0.85, while we were 95% sure that it was ≥ 0.55 , corresponding to a β (8.0; 2.2) distribution. For TUS Sp, a mode of 0.95 was hypothesized, with 95% certainty that $Sp \geq 0.90$, corresponding to a β (99.7; 6.2) distribution (Berman et al., 2019).

WinBUGS code

```

model{
    #=====DATA=====#
list(t12=c(110, 67, 378, 245))
    #=====INITS=====#
Init 1: list(Se_tus = 0.9, Sp_tus = 0.9, Se_hap = 0.9, Sp_hap = 0.9, pi = 0.25)
Init 2: list(Se_tus = 0.8, Sp_tus = 0.8, Se_hap = 0.8, Sp_hap = 0.8, pi= 0.1)
Init 3: list(Se_tus = 0.7, Sp_tus = 0.7, Se_hap = 0.7, Sp_hap = 0.7, pi=0.05)
    #=====LIKELIHOOD=====#
# TWO INDEPENDENT TESTS IN ONE POPULATION#
t12[1:4] ~ dmulti(p12[1:4], n)
n <- sum(t12[1:4])
p12[1] <- pi*(Se_tus*Se_hap) + (1-pi)*((1-Sp_tus)*(1-Sp_hap))
p12[2] <- pi*(Se_tus*(1-Se_hap)) + (1- pi)*((1-Sp_tus)*Sp_hap)
p12[3] <- pi*((1-Se_tus)*Se_hap) + (1- pi)*(Sp_tus*(1-Sp_hap))
p12[4] <- pi*((1-Se_tus)*(1-Se_hap)) + (1- pi)*(Sp_tus*Sp_hap)
    #==PRIORS FOR PREVALENCE===#
pi ~ dbeta(1.7, 14.5) ##Informative with mode at 5% and 95% sure that lower than 0.25
    #=====PRIORS FOR SENSITIVITY AND SPECIFICITY=====#

```

```

#HAP#
Se_hap ~ dbeta(9.6,3.9) ## Elicited prior: mode= 0.75; 95% ≥ 0.50 Se = 0.76 (0.53; 0.92)
Sp_hap ~ dbeta(36.7, 2.9) ## Elicited prior: mode= 0.95; 95% ≥ 0.85 Sp = 0.95(0.88; 0.98)

#TUS#
Se_tus ~ dbeta(8.0,2.2) ## Elicited prior: mode= 0.85; 95% ≥ 0.55 Se = 0.89 (0.55, 1.0)
Sp_tus ~ dbeta(99.7,6.2) ## Elicited prior: mode= 0.95; 95% ≥ 0.90 Sp = 0.95 (0.92, 0.98)
}

```

IV-2- LCM-regression

LCM-regression is a logistic regression to estimate individual weight coefficients of the following clinical signs: cough, ears droop/head tilt, and increased rectal temperature.

WinBUGS code

```

Data ##data for each clinical sign are available in separate files

Model {
for (i in 1:800)
{ Y[i] ~ dbern(s[i])

## Y (lung consolidation) follow a Bernoulli event of probability s(i) =probability to have a positive test
s[i] <-ps[i]*Se_tus+(1-ps[i])*(1-Sp_tus)

## The probability of being consolidated s depend on the probability of being truly sick (ps) and test accuracy
(Se and Sp of TUS) for the ith calf

#====RECTAL TEMPERATURE====#
logit(ps[i]) <- temp[i]*b1 + b7 + u1[batch[i]] + u2[herd[i]]

#=====COUGH=====#
logit(ps[i]) <- cough[i]*b2 + b7 + u1[batch[i]] + u2[herd[i]]

#===EAR DROOP/HEAD TILT===#
logit(ps[i]) <- ear[i]*b3 + b7 + u1[batch[i]] + u2[herd[i]]
}

for (k in 1:80) { u1[k] ~ dnorm(0,tau1) } ## u1 is the batch effect for the ith calf
for (j in 1:47) { u2[j] ~ dnorm(0,tau2) } ## u2 is the herd effect for the ith calf

#=====PRIORS FOR COEFFICIENT AND RANDOM INTERCEPTS=====#

b1 ~ dnorm (1.4, 1.7) #coefficient for rectal temperature informative
b2~ dnorm (1.4, 1.7) #coefficient for cough informative
b3 ~ dnorm (1.4, 1.7) #coefficient for ear droop/head tilt informative
b7 ~ dnorm (1.4, 1.7) #intercept
tau1 ~ dgamma (0.1,0.1)
tau2 ~ dgamma (0.1,0.1)

```

#=====PRIORS ON SENSITIVITY AND SPECIFICITY=====#

```

Se_tus ~ dbeta (5.7, 2.5)I(0.60,) #truncature at 0.60
Sp_tus ~ dbeta (316.6, 36.1)
}

```

IV-3- LCM-Validation

Estimation of Se and Sp of the VcCRSC, TUS and Hap (in population 1) to detect active BRD. The developing population (800 calves) was used to define the optimal cut-off (three tests, one population). The population 1, population 2, and population 1+2 were used to validate VcCRSC externally.

Assumption: We assumed that VcCRSC, TUS and Hap were conditionally independents based on the results of our previous study (Berman et al., 2019). We, also, assumed, that the accuracy of TUS was the same for all calves included in the study.

Latent class model elaboration: three independent tests, in one population: The data, $x = (x_{110}, x_{111}, x_{100}, x_{101}, x_{010}, x_{011}, x_{001}, x_{000})$, corresponded to the cross-classified test results for the n calves tested from the population where the vector x_{ab} represented the number of calves that tested positive (a or $b = 1$) or negative (a or $b = 0$) in both tests. Let $p = (p_{111}, p_{110}, p_{101}, p_{100}, p_{011}, p_{010}, p_{001}, p_{000})$ represented the respective probabilities from vector x . The multinomial sampling distribution was given by:

$$x \sim \text{multinomial} ((p_{111}; p_{110}; p_{101}; p_{100}; p_{011}; p_{010}; p_{001}; p_{000}), n)$$

where the multinomial cell probabilities were given by:

$$\begin{aligned}
p_{111} &= p(\text{TUS+; Hap+; CRSC+}) = \pi(\text{SeTUS} * \text{SeHap} * \text{SeCRSC}) + (1-\pi)((1-\text{SpTUS})(1-\text{SpHap})(1-\text{SpCRSC})) \\
p_{110} &= p(\text{TUS+; Hap+; CRSC-}) = \pi(\text{SeTUS} * \text{SeHap} * (1-\text{SeCRSC})) + (1-\pi)((1-\text{SpTUS})(1-\text{SpHap})\text{SpCRSC}) \\
p_{101} &= p(\text{TUS+; Hap-; CRSC+}) = \pi(\text{SeTUS} * (1-\text{SeHap}) * \text{SeCRSC}) + (1-\pi)((1-\text{SpTUS}) * \text{SpHap} * (1-\text{SpCRSC})) \\
p_{100} &= p(\text{TUS+; Hap-; CRSC-}) = \pi(\text{SeTUS} * (1-\text{SeHap}) * (1-\text{SeCRSC})) + (1-\pi)((1-\text{SpTUS}) * \text{SpHap} * \text{SpCRSC}) \\
p_{011} &= p(\text{TUS-; Hap+; CRSC+}) = \pi((1-\text{SeTUS}) * \text{SeHap} * \text{SeCRSC}) + (1-\pi)(\text{SpTUS} * (1-\text{SpHap}) * (1-\text{SpCRSC})) \\
p_{010} &= p(\text{TUS-; Hap+; CRSC-}) = \pi((1-\text{SeTUS}) * \text{SeHap} * (1-\text{SeCRSC})) + (1-\pi)(\text{SpTUS} * (1-\text{SpHap}) * \text{SpCRSC}) \\
p_{001} &= p(\text{TUS-; Hap-; CRSC+}) = \pi((1-\text{SeTUS}) * (1-\text{SeHap}) * \text{SeCRSC}) + (1-\pi)(\text{SpTUS} * \text{SpHap} * (1-\text{SpCRSC})) \\
p_{000} &= p(\text{TUS-; Hap-; CRSC-}) = \pi((1-\text{SeTUS}) * (1-\text{SeHap}) * (1-\text{SeCRSC})) + (1-\pi)(\text{SpTUS} * \text{SpHap} * \text{SpCRSC})
\end{aligned}$$

TUS+: Calves with lung consolidation depth ≥ 3 cm
TUS-: Calves with lung consolidation depth < 3 cm
Hap+: Calves with a concentration of Hap ≥ 0.25 g/l
Hap-: Calves with a concentration of Hap < 0.25 g/l

SeTUS: Sensitivity of TUS
SeHap: Sensitivity of Hap
SpTUS: Specificity of TUS
SpHap: Specificity of Hap

CRSC+: Calves with a score \geq cut-off*

CRSC-: Calves with a score $<$ cut-off*

pi: Prevalence of active BRD

*CRSC+ and CRSC- were defined as having a score \geq and $<$ 15, respectively for validating VcCRSC in the population 1, population 2, and population 1 + 2.

WinBUGS code

```
model{
  #=====DATA=====#
  Population 1: List (t123=c(1, 5, 0, 0, 4, 17, 6, 176))
  Population 2 : List(t12=c(3, 3, 30, 477),)
  Population 1 et 2 : List(t12_pop1=c(6, 0, 21, 182), t12_pop2=c(3, 3, 30, 477))

  #=====INITS=====#

  Population 1:
  Init 1: list(Se_tus = 0.9, Sp_tus = 0.9, Se_hap = 0.9, Sp_hap = 0.9, prev = 0.25)
  Init 2: list(Se_tus = 0.8, Sp_tus = 0.8, Se_hap = 0.8, Sp_hap = 0.8, prev= 0.1)
  Init 3: list(Se_tus = 0.7, Sp_tus = 0.7, Se_hap = 0.7, Sp_hap = 0.7, prev=0.05)

  Population 2:
  Init 1: list(Se_tus = 0.9, Sp_tus = 0.9, prev = 0.25)
  Init 2: list(Se_tus = 0.8, Sp_tus = 0.8, prev= 0.1)
  Init 3: list(Se_tus = 0.7, Sp_tus = 0.7, prev=0.05)

  Population 1 et 2:
  Init 1 : list(Se_tus = 0.9, Sp_tus = 0.9, pi_pop1 = 0.25, pi_pop2 = 0.15)
  Init 2 : list(Se_tus = 0.8, Sp_tus = 0.8, pi_pop1= 0.1 pi_pop2= 0.1)
  Init 3 : list(Se_tus = 0.7, Sp_tus = 0.7, pi_pop1=0.05 pi_pop2=0.05)

  model
  {
  #===== LIKELIHOOD =====#

  Population 1
  #===== CONSIDER THE THREE TESTS INDEPENDENT IN ONE POPULATION =====#

  n <- sum(t123[1:8])
  t123[1:8] ~ dmulti(p123[1:8], n)

  p123[1]<- pi*Se_sc*Se_tus*Se_hap + (1-pi)*(1-Sp_sc)*(1-Sp_tus)*(1-Sp_hap)
  p123[2]<- pi*Se_sc*Se_tus*(1-Se_hap) + (1-pi)*(1-Sp_sc)*(1-Sp_tus)*Sp_hap
```



```

p123[3]<- pi*Se_sc*(1-Se_tus)*Se_hap + (1-pi)*(1-Sp_sc)*Sp_tus*(1-Sp_hap)
p123[4]<- pi*Se_sc*(1-Se_tus)*(1-Se_hap) + (1-pi)*(1-Sp_sc)*Sp_tus*Sp_hap
p123[5]<- pi*(1-Se_sc)*Se_tus*Se_hap + (1-pi)*Sp_sc*(1-Sp_tus)*(1-Sp_hap)
p123[6]<- pi*(1-Se_sc)*Se_tus*(1-Se_hap) + (1-pi)*Sp_sc*(1-Sp_tus)*Sp_hap
p123[7]<- pi*(1-Se_sc)*(1-Se_tus)*Se_hap + (1-pi)*Sp_sc*Sp_tus*(1-Sp_hap)
p123[8]<- pi*(1-Se_sc)*(1-Se_tus)*(1-Se_hap) + (1-pi)*Sp_sc*Sp_tus*Sp_hap

```

Population 2

```

#==== CONSIDER THE TWO TESTS INDEPENDANT IN ONE POPULATION====#

```

```

n <- sum(t12[1:4])

```

```

t12[1:4] ~ dmulti(p12[1:4], n)

```

```

p12[1]<- pi*Se_sc*Se_tus + (1-pi)*(1-Sp_sc)*(1-Sp_tus)
p12[2]<- pi*Se_sc*(1-Se_tus) + (1-pi)*(1-Sp_sc)*Sp_tus
p12[3]<- pi*(1-Se_sc)*Se_tus + (1-pi)*Sp_sc*(1-Sp_tus)
p12[4]<- pi*(1-Se_sc)*(1-Se_tus) + (1-pi)*Sp_sc*Sp_tus

```

Population 1 and 2 :

```

#==== CONSIDER THE TWO TESTS INDEPENDANT IN TWO POPULATIONS====#

```

```

# = POP 1 =#

```

```

n_pop1 <- sum(t12_pop1[1:4])
t12_pop1[1:4] ~ dmulti(p12_pop1[1:4], n_pop1)

```

```

p12_pop1[1]<- pi_pop1*Se_sc*Se_tus + (1-pi_pop1)*(1-Sp_sc)*(1-Sp_tus) # (+,+)
p12_pop1[2]<- pi_pop1*Se_sc*(1-Se_tus) + (1-pi_pop1)*(1-Sp_sc)*Sp_tus # (+,-)
p12_pop1[3]<- pi_pop1*(1-Se_sc)*Se_tus + (1-pi_pop1)*Sp_sc*(1-Sp_tus) # (-,+)
p12_pop1[4]<- pi_pop1*(1-Se_sc)*(1-Se_tus) + (1-pi_pop1)*Sp_sc*Sp_tus # (-,-)

```

```

# = POP 2 =#

```

```

n_pop2 <- sum(t12_pop2[1:4])
t12_pop2[1:4] ~ dmulti(p12_pop2[1:4], n_pop2)

```

```

p12_pop2[1]<- pi_pop2*Se_sc*Se_tus + (1-pi_pop2)*(1-Sp_sc)*(1-Sp_tus) # (+,+)
p12_pop2[2]<- pi_pop2*Se_sc*(1-Se_tus) + (1-pi_pop2)*(1-Sp_sc)*Sp_tus # (+,-)
p12_pop2[3]<- pi_pop2*(1-Se_sc)*Se_tus + (1-pi_pop2)*Sp_sc*(1-Sp_tus) # (-,+)
p12_pop2[4]<- pi_pop2*(1-Se_sc)*(1-Se_tus) + (1-pi_pop2)*Sp_sc*Sp_tus # (-,-)

```

```

#=====PRIORS FOR PREVALENCE=====#

```

```

prev ~ dbeta(1.7, 14.5) ## Informative with mode at 5% and 95% sure that lower than 0.25
pi_pop1 ~ dbeta(1.7, 14.5) ## Informative with mode at 5% and 95% sure that lower than 0.25
pi_pop2 ~ dbeta(1.7, 14.5) ## Informative with mode at 5% and 95% sure that lower than 0.25

```

```

#=====PRIORS FOR SENSITIVITY AND SPECIFICITY=====#

```

```

#HAP#

```

```

Se_hap ~ dbeta(9.6,3.9) ## Elicited prior: mode= 0.75; 95% ≥ 0.50 Se = 0.76 (0.53; 0.92)
Sp_hap ~ dbeta(36.7, 2.9) ## Elicited prior: mode= 0.95; 95% ≥ 0.85 Sp = 0.95(0.88; 0.98)

```

```

#TUS#

```

```

Se_tus ~ dbeta(8.0,2.2) ## Elicited prior: mode= 0.85; 95% ≥ 0.55 Se = 0.89 (0.55, 1.0)
Sp_tus ~ dbeta(99.7,6.2) ## Elicited prior: mode= 0.95; 95% ≥ 0.90 Sp = 0.95 (0.92, 0.98)

```

```
#VcCRSC#
```

```
Se_crsc ~ dbeta(1, 1) ## Non-informative
```

```
Sp_crsc ~ dbeta(1, 1) ## Non-informative
```

```
#####LIKEHOOD RATIO#####
```

```
LR_pos <- Se_sc/(1-Sp_sc)
```

```
LR_neg <- (1-Se_sc)/Sp_sc
```

```
#####PREDICTIVES VALUES #####
```

```
PPV <- (Se_sc*pi)/((Se_sc*pi)+(1-Sp_sc)*(1-pi))
```

```
NPV <- (Sp_sc*(1-pi))/((Sp_sc*(1-pi)+(1-Se_sc)*pi))
```

```
#####MISCLASSIFICATION TERM#####
```

```
mct_1 <- (1-pi)*(1-Sp_sc)+r_1*pi*(1-Se_sc)
```

```
mct_3 <- (1-pi)*(1-Sp_sc)+r_3*pi*(1-Se_sc)
```

```
mct_8 <- (1-pi)*(1-Sp_sc)+r_8*pi*(1-Se_sc)
```

```
r_1 <- 1
```

```
r_3 <- 3
```

```
r_8 <- 8
```

```
}
```

V- Comparison of different alternative score

Clinical scores	Clinical signs	Specific score	External validation	
Score 1	Only reliable clinical signs in (Berman et al., 2021)	= 10*Cough + 9*Ear + 6*Temperature	Se = 0.31 (0.14; 0.70)	Sp = 1.0 (0.99; 1.0)
Score 2	Clinical signs form the Californian score (Love et al., 2014)	= 10*Cough + 9*Ear + 6*Temperature+2*Eye +1*Nasal+9*Dyspnea	Se = 0.37 (0.14; 0.82)	Sp = 0.99 (0.97; 1.0)
Score 3	Reliable clinical signs + dyspnea	= 10*Cough + 9*Ear + 6*Temperature+ 9*Dyspnea	Se = 0.36 (0.15; 0.75)	Sp = 0.99 (0.97; 1.0)

CONCLUSION ARTICLE 2 : Dans cet ARTICLE 2, nous avons DÉVELOPPÉ et VALIDÉ un nouveau SCRD, facile d'utilisation et performant à l'échelle POPULATIONNELLE pour permettre d'estimer la prévalence de BPI dans un lot et ainsi mieux raisonner l'utilisation des antibiotiques en groupe dans les élevages de veaux lourds. Son utilisation à l'échelle individuelle est cependant peu performante pour détecter les veaux malades (faible Se). L'amélioration d'un test d'imagerie médicale permettrait d'améliorer la détection individuelle. Voyons à présent dans le prochain chapitre, laquelle de l'échographie et de la radiographie thoracique serait le test plus performant pour envisager son utilisation sur le terrain.

Chapitre 3 – Comparaison entre l'échographie thoracique et la radiographie thoracique chez les génisses de remplacement

ARTICLE 3

Comparison between thoracic ultrasonography and thoracic radiography for the detection of thoracic lesions in dairy calves using a two-stage Bayesian method

Publié dans le *Preventive Veterinary Medicine* 2020 Nov1; 184:105-153

Berman J^a, Masseau I^a, Fecteau G^a, Buczinski S^a, Francoz D^a

^a From the Département des sciences cliniques, Faculté de médecine vétérinaire, Université de Montréal, Saint-Hyacinthe, QC, Canada, J2S 2M2 .

Abbreviation: ACVR = American College of Veterinary Radiology; BCI = Bayesian credible interval; BP = Bronchopneumonia; CT = Computed tomography; CT+ = Presence of lung consolidation, and/or abscesses, and/or pneumothorax and/or pleural effusion on computed tomography; CT- = Absence of lung consolidation, abscesses, pneumothorax or pleural effusion on computed tomography; Covn = the covariance of the tests for negative cases; Covp = the covariance of the tests for positive cases; DIC = Deviance Information Criterion; LR+ = Positive Likelihood Ratio; LR- = Negative Likelihood Ratio; NPV = Negative Predictive Value; pD = Number of parameters effectively estimated; PPV = Positive Predictive Value; PPV_{JE} = Positive Predictive Value of the simultaneous testing of thoracic ultrasonography and thoracic radiography; Prev = Prevalence of thoracic lesions; Se = Sensitivity; Se_{CT} = Sensitivity of computed tomography; Se_{JE} = The joint sensitivity of the simultaneous testing (denoted by the subscript JE, for Joint Either); Se_p = Sensitivity of using TR and TUS in parallel; Se_s = Sensitivity of using TR and TUS in series; Se_{TR} = Sensitivity of thoracic radiography; Se_{TUS} = Sensitivity of thoracic ultrasonography; Sp = Specificity; Sp_{CT} = Specificity of computed tomography; Sp_{JE} = The joint specificity of the simultaneous testing (denoted by the subscript JE, for Joint Either); Sp_p =

Specificity of using TR and TUS in parallel; Sp_s = Specificity of using TR and TUS in series; Sp_{TR} = Specificity of thoracic radiography; Sp_{TUS} = Specificity of thoracic ultrasonography; TR = Thoracic Radiography; TR+= Presence of bronchial/interstitial/alveolar pattern, and/or pneumothorax and/or pleural effusion on thoracic radiography; TR-= Absence of bronchial/interstitial/alveolar pattern, pneumothorax or pleural effusion on thoracic radiography; TUS = Thoracic Ultrasonography; TUS+ = Presence of consolidation (≥ 1 cm), and/or evidence of a pneumothorax, and/or abscess and/or relevant (≥ 1 cm) pleural effusion on TUS; TUS - = Absence of consolidation (≥ 1 cm), and/or evidence of a pneumothorax, and/or abscess and/or relevant (≥ 1 cm) pleural effusion on TUS.

ABSTRACT

Objectives: Infectious bronchopneumonia is a lower respiratory tract disease with major economic consequences in dairy calves. Thoracic radiography (**TR**) and thoracic ultrasonography (**TUS**) are two imaging diagnostic procedures available in bovine medicine for identifying thoracic lesions. However, no study has investigated whether one of these tests is superior to the other or if they provide comparable results for the detection of thoracic lesions in calves. The objective of this study was therefore to estimate and to compare the performances of TUS and TR for the detection of thoracic lesions in dairy calves.

Methods: A prospective cross-sectional study was performed in a hospital setting. A total of 50 calves (≥ 7 days old; ≤ 100 kg; standing; $pCO_2 \geq 53$ mmHg; any reason of presentation) were enrolled. Every calf underwent TUS and TR. Only calves with thoracic lesions on TUS and/or TR were controlled by thoracic computed tomography (**CT**) (the gold standard). Calves without lesions were not controlled by CT. A two-stage Bayesian framework was used. The sensitivities (**Se**) and specificities (**Sp**) of both tests individually and used in series or parallel were estimated.

Results: The Se and Sp of TUS were 0.81 (95% **BCI** (Bayesian Credible Interval): 0.65; 0.92) and 0.90 (95% BCI: 0.81; 0.96), respectively. The Se and Sp of TR were 0.86 (95% BCI: 0.62; 0.99) and 0.89 (95% BCI: 0.67; 0.99), respectively. This study did not reveal any differences between both tests. Using TUS and TR in series was more specific than using both tests in parallel. The performances of TUS alone were not different from the performances of both tests in series or in parallel.

Conclusion: In conclusion, TUS and TR were equivalent in detecting thoracic lesions in this study. Using TUS alone allowed an accurate detection of thoracic lesions in dairy calves. Further studies enrolling a larger sample (> 400 calves) and allowing adequate power to be achieved would be necessary to confirm these results.

Key words: Pneumonia; diagnostic imaging; cattle; latent class; lungs; consolidation.

I- INTRODUCTION

Bronchopneumonia (**BP**) refers to an infection of the lower respiratory tract and is an important cause of death in dairy calves (Windeyer et al., 2014). Infection during the pre-weaning period affects survivability and reduces dairy performance later in life (Teixeira et al., 2017b, Dunn et al., 2018). Accurate detection of BP is therefore necessary for implementation of an optimal treatment from the onset of the disease and aiming at reducing mortality and long-term economic losses. However, the diagnosis of BP in dairy calves was found to be inaccurate when solely based on physical examination findings and abnormalities on auscultation (Buczinski et al., 2014). One promising avenue is the objective identification of thoracic lesions using diagnostic imaging.

Diagnostic imaging procedures, such as thoracic radiography (**TR**) and thoracic ultrasonography (**TUS**), have already been shown to provide useful information regarding the presence of thoracic lesions in calves (Jones et al., 1998; Ollivett and Buczinski, 2016). However, each of these tests has its advantages and limitations. Indeed, TR may provide an overview of the entire lung on a single projection in a relatively young animal; however, its interpretation is limited by the superimposition of anatomical structures, including the shoulder musculature, when TR is performed with the animal standing. In contrast, TUS is easily performed, with a good inter-operator agreement for the detection of lung consolidations (Buczinski et al, 2018b), but is limited to the assessment of lesions close to the thoracic wall (Reef et al., 1991; Babkine and Blond, 2009).

One previous report showed that TUS was less sensitive than TR in newborn dairy calves for detecting respiratory disorders (Jung and Bostedt, 2004). However, sensitivity was determined based on TR alone, known to be an imperfect test, rather than a gold standard such as computed tomography (**CT**), therefore limiting their ability to assess a potential

superiority of either TUS or TR for the detection of thoracic lesions in calves. In contrast, TUS was reported to be more sensitive than TR in detecting specific thoracic lesions in adult cattle, in other species, and in human medicine, by limiting false negatives caused by superimposition (Reef et al., 1991; Braun et al., 1997; Ye et al., 2015; Partlow et al., 2017; Ward et al., 2018). Thus, it is currently unclear whether one of the tests is superior to the other in the detection of thoracic lesions or in the diagnosis of BP in dairy calves. A need for clarification is also necessary to determine whether both tests are complementary and thus must be performed serially, or whether they are equivalent and used according to individual preference.

The main objective of this study was therefore to estimate and compare the performances of TUS and TR in the detection of thoracic lesions (as a proxy for BP) in hospitalized dairy calves.

II- MATERIALS AND METHODS

Design, conduct, and results of the current study are reported in accordance with STARD (Standards for Reporting of Diagnostic Accuracy Studies) 2015 guidelines (Cohen et al., 2016). The study protocol was approved by the Comité d'éthique de l'utilisation des animaux de l'Université de Montréal (#20160111).

II-1- Population

Dairy calves admitted to the Centre Hospitalier Universitaire Vétérinaire (CHUV) of the Université de Montréal from March 22 to May 6, 2016, from June 13 to November 4, 2016, and from January 16 to October 4, 2017, were prospectively considered for this study. Calves were considered irrespective of the cause of presentation. The calves represented a random sample of the population of hospitalized calves. Dairy calves at least 7 days old and weighed less than 100 kg were eligible. The maximal weight was set at 100 kg based on CT gantry limitations. Calves had to be able to stand up on their own and to demonstrate a $p\text{CO}_2 \leq 53$ mmHg on an arterial blood gas analysis (Wilson et al., 2015). Critical calves ($p\text{CO}_2 > 53$ mmHg, in decubitus) and/or neonates (< 7 days old) were not included due to risks associated with the procedures. Calves were excluded from the study if the consent form was not signed by the owner in a timely fashion, or if the calf was considered unsuitable for diagnostic imaging (e.g., biosecurity).

II-2- Data collection

Each diagnostic test was conducted in the chronological order shown in **Figure 10**. Diagnostic tests were completed within 24 hours of presentation to avoid or minimize any changes in the evolution of the disease that could affect the results.

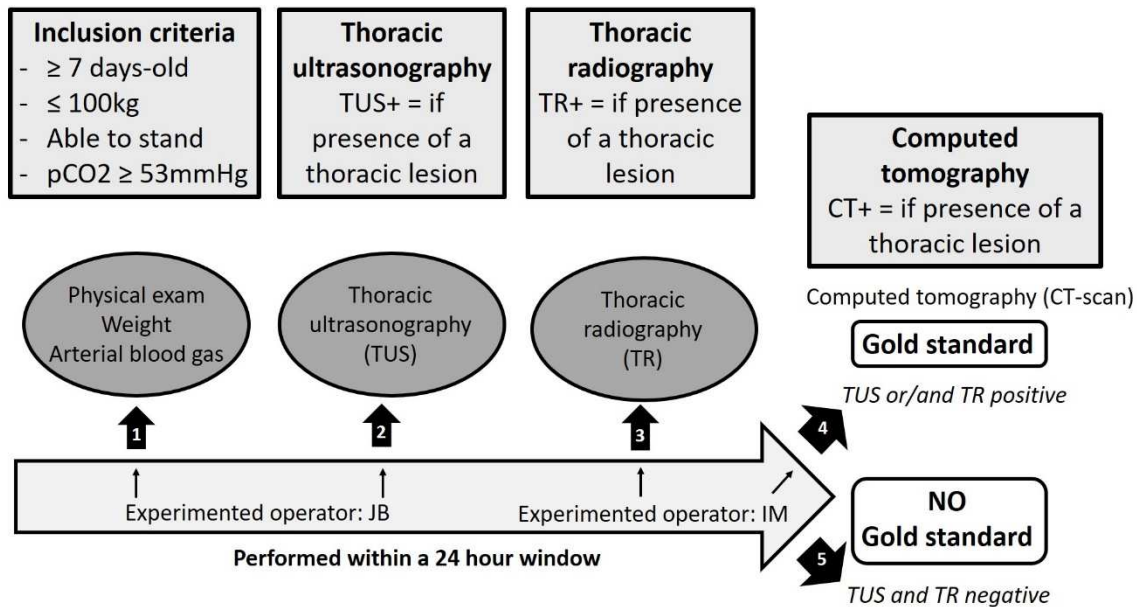


Figure 10. – Data collection and time frame for the 50 calves enrolled in the study. Each test was conducted in chronological order from 1 to 4 within 24 hours. Steps 1 and 2 were performed by one experiment operator (JB). Tests in Steps 3 and 4 were analyzed by another experienced operator (IM) who was unaware of the results from the previous tests.

II-2-2- Thoracic ultrasonography

Thoracic ultrasonography was performed with a portable ultrasound unit equipped with a 7.5 MHz linear transducer (Sonoscape S6V, Shenzhen, China). Each calf was clipped from the 11th to the 3rd intercostal space on the left and right sides, respectively. Acoustic coupling gel was then applied to provide good conductivity of ultrasound waves and to obtain high-quality images. The lungs were then scanned over the entire clipped area in a standardized fashion (Ollivett, 2014). The real-time sonographic appearance of the pleura and lungs was evaluated for the presence of lung consolidations (defined as hypoechoic lung tissue resembling the sonographic appearance of hepatic parenchyma), pneumothorax (recorded as positive when the sliding movement of the pleura was no

longer visible), pleural effusion (recorded as positive when the pleurae were separated from one another by anechoic to echogenic fluid, according to its particular or cellular content), or cavitory lesion (defined as a well-defined circular region inside the lung parenchyma) (Babkine and Blond, 2009). Other lesions (pleura irregularity, B-line, and comet-tail) were not analyzed because of the difficulty in interpreting these lesions in calves (Ollivett et al., 2014; Buczinski et al., 2018b). When present, the maximal depth of lung consolidations (cm), abscesses (cm), or effusions (cm) were measured using the veterinary measurement software kit available on the ultrasound unit. The depth and localization (intercostal space) of each lesion was recorded on a scoring sheet, as previously described (Buczinski et al., 2013). Standardized TUS was performed by the same experimental observer (JB), who was blind to the TR and CT results.

II-2-1-General examination and arterial blood sampling

A complete physical examination and an arterial blood gas analysis were performed immediately upon presentation. An amount of 0.5 to 1 mL of arterial 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 (Bleul, 2009). The analysis was promptly performed with a blood gas analyzer (ABL 80 FLEX, Radiometer America, Inc., Brea, CA, USA). The pCO₂ value was immediately measured to determine whether the value obtained met the inclusion criterion.

II-2-3- Thoracic radiography

Each calf underwent standing left to right latero-lateral and recumbent ventro-dorsal radiographs. All radiographs were performed with 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). According to the size of the calf, one to three radiographs were obtained for each position in order to cover the entire thorax. The 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). The TR studies were reviewed twice: immediately after image acquisition and 10 months after the enrolment of the last calf in the study. Immediately after image acquisition, the thoracic studies were evaluated by an American

College Veterinarian Radiologist (ACVR) board-certified specialist (IM), unaware of the TUS results, to determine whether thoracic lesions were present or absent. When a lesion was deemed present on either TUS or TR, the calf was submitted to a thoracic CT within 24 hours of the completion of the first examinations (**Figure 10**). The second radiographic review was performed by the same radiologist for more detailed descriptions of the lesions. Each radiographic image and each CT image were blindly reviewed. Each thoracic study (including the one to three projections per subject) was assessed for the presence of bronchial patterns, unstructured or structured interstitial patterns, and alveolar patterns. A bronchial pattern was defined as small to large caliber airways with prominent and/or thickened walls. An unstructured interstitial pattern consisted of an area of increased opacity of the lung field with ill-defined margins associated with blurred vascular borders. When the area (< 3 cm) of increased soft tissue opacity had relatively well-defined borders, it was referred to as a nodule and therefore considered a structured interstitial pattern. An alveolar pattern was defined as an area of increased soft tissue opacity with indistinct pulmonary vessels, with or without air bronchograms or border effacement between the lung lesion and the heart, cranial mediastinum, or diaphragm. Pleural effusion was diagnosed based on the retraction of lung lobe margins, pleural fissure lines, and fluid opacity between the thoracic wall and the lung margins. The presence of air between the lung lobes and the thoracic wall causing retraction of the lung margins was identified as a pneumothorax (Masseau et al., 2008; Fowler et al., 2017).

II-2-4- Thoracic CT-Scan

Thoracic CT studies were performed using a 16-detector row CT scanner (GE LightSpeed, GE Healthcare, Chicago, IL, USA). Calves were sedated with a combination of 0.1 mg/kg of diazepam (Valium, Roche, Nutley, NJ, USA), 0.05 mg/kg of butorphanol (Torbugesic, Pfizer, New York, NY, USA), and 0.05 mg/kg of xylazine (Rompun, Bayer, Mississauga, ON, Canada). Each calf was scanned in sternal recumbency from approximately 10 cm cranial to the thoracic inlet extending to the mid-level of the left kidney using the following parameters: 0.5 mm collimation, a 0.5 mm reconstruction interval, a 1 s tube rotation time, 120 kV, 250 mA, a pitch factor of 0.828, and a 0-degree tilt. High-resolution (1.25 mm slice thickness) transverse images were reconstructed using high- and low-spatial-frequency algorithms and reviewed using lung and soft tissue window settings. Sagittal and

dorsal reconstructed images of the thorax were also available. All diagnostic images were transferred to a PACS and viewed on a single workstation using AGFA Impax 6 (Agfa, Toronto, ON, Canada). One specialist (ACVR Diplomate) (IM) analyzed the images 10 months after the enrolment of the last calf. The identification of the images (name of the owner, identification number and name of the calf) were hidden. The radiologist was therefore unaware of the TUS and TR findings. The CT images were assessed for the presence of pleural effusion (defined as an accumulation of fluids within the pleural space), pneumothorax (defined as the presence of air within the pleural space), and abscesses (the presence of a nodular lesion with a thick wall that is > 3 mm and a horizontal interface separating gravity dependent on soft tissue/fluid attenuation from a non-dependent gas). Consolidation was defined as an area of lung tissue characterized by increased attenuation with indistinct pulmonary vascular or airway wall markings (Ohlerth et al., 2014). The TR and CT studies in each calf were evaluated independently of each other.

II-3- Definitions

The presence or absence of thoracic lesions on CT, TUS, or TR was classified according to a dichotomous diagnostic algorithm. CT was considered positive (**CT+**; the gold standard) if one or more of the following lesions were present: CT consolidation, abscesses, pneumothorax, or pleural effusion. TUS was considered positive (**TUS+**) when the calf had one or more lesions of TUS consolidations (≥ 1 cm), evidence of a pneumothorax, abscesses, or relevant (≥ 1 cm) pleural effusion, or a combination of any of these conditions. Only lesions with a depth of ≥ 1 cm were considered relevant for the data analysis (Buczinski et al., 2014, Buczinski et al., 2015a). TR was considered positive (**TR+**) if the calf showed evidence of one or more lesions consisting of a bronchial/interstitial/alveolar pattern, pneumothorax, or pleural effusion. When no lesions meeting the criteria mentioned above were identified, the test was automatically considered negative and designated **CT-**, **TR-**, or **TUS-**.

II-4- Statistical analyses

For the statistical analysis, the calf was the unit of interest. Descriptive data and contingency tables were constructed. In the absence of a gold standard for TUS- and TR-calves, a two-stage Bayesian method was used as reported previously (Liu et al., 2014).

The first stage considered the results obtained from two imperfect screening tests (TUS and TR). In the second stage, the gold standard test (CT) was performed only for positive samples from the simultaneous testing of the first stage (cases with TUS+ and/or TR+ tests), whereas the negative samples (cases with TUS- and TR- tests) were not verified by the gold standard test. The comparison between both tests was expressed in term of sensitivity (**Se**), specificity (**Sp**), positive predictive value (**PPV**), negative predictive value (**NPV**), positive likelihood ratio (**LR+**), and negative likelihood ratio (**LR-**).

II-4-1- Bayesian modeling for stage one

In the first stage, the analysis was performed by simultaneously applying the results of TUS and TR performed for each calf. Since both tests measured the same biological parameters (thoracic lesions), we assumed that they were conditionally dependent and as a reasonable default assumption (Wang et al., 2017). Distributions of covariance parameters were therefore defined as follows based on Dendukiri et al., 2001:

$$\text{Covp} \sim \text{dunif}(0; \min(\text{Se}_{\text{TR}}, \text{Se}_{\text{TUS}}) - \text{Se}_{\text{TR}} * \text{Se}_{\text{TUS}})$$

$$\text{Covn} \sim \text{dunif}(0; \min(\text{Sp}_{\text{TR}}, \text{Sp}_{\text{TUS}}) - \text{Sp}_{\text{TR}} * \text{Sp}_{\text{TUS}})$$

Where **Covp** and **Covn** correspond to the covariance of the tests for positive (p) or negative (n) cases; **Se_{TR}** and **Sp_{TR}** correspond to the Se and Sp of TR, respectively; and **Se_{TUS}** and **Sp_{TUS}** correspond to the Se and Sp of TUS, respectively.

The model for stage one consisted in a model associated with two dependent tests in one population (Branscum et al., 2005). Briefly, the data, $x = (x_{11}, x_{10}, x_{01}, x_{00})$, corresponded to the cross-classified test results for the n calves tested from the population where the vector x_{ij} represented the number of calves that tested positive (i or j=1) or negative (i or j=0) in both tests. Let $p = (p_{11}, p_{10}, p_{01}, p_{00})$ represented the respective probabilities from vector x. The multinomial sampling distribution was given by:

$$x \sim \text{multinomial}((p_{11}; p_{10}; p_{01}; p_{00}), n)$$

where the multinomial cell probabilities were given by:

$$p_{11} = p(\text{TUS+}; \text{TR+}) = \text{Prev}(\text{Se}_{\text{TUS}} * \text{Se}_{\text{TR}} + \text{Covp}) + (1 - \text{Prev})((1 - \text{Sp}_{\text{TUS}})(1 - \text{Sp}_{\text{TR}}) + \text{Covn})$$

$$p_{10} = p(\text{TUS+}; \text{TR-}) = \text{Prev}(\text{Se}_{\text{TUS}} * (1 - \text{Se}_{\text{TR}}) - \text{Covp}) + (1 - \text{Prev})((1 - \text{Sp}_{\text{TUS}}) * \text{Sp}_{\text{TR}} - \text{Covn})$$

$$p_{01} = p(\text{TUS-}; \text{TR+}) = \text{Prev}((1 - \text{Se}_{\text{TUS}}) * \text{Se}_{\text{TR}} - \text{Covp}) + (1 - \text{Prev})(\text{Sp}_{\text{TUS}} * (1 - \text{Sp}_{\text{TR}}) - \text{Covn})$$

$$p_{00} = p(\text{TUS-}; \text{TR-}) = \text{Prev}((1 - \text{Se}_{\text{TUS}}) * (1 - \text{Se}_{\text{TR}}) + \text{Covp}) + (1 - \text{Prev})(\text{Sp}_{\text{TUS}} * \text{Sp}_{\text{TR}} + \text{Covn})$$

where **Prev** is the prevalence of thoracic lesions. If either individual test was positive, this condition indicated a positive result of the joint testing. The joint sensitivity (**Se_{JE}**) and specificity (**Sp_{JE}**) of the simultaneous testing (denoted by the subscript JE, for Joint Either) were as follows (Liu et al.; 2014):

$$Se_{JE} = P(TUS+ \cup TR+ | BP+) = Se_{TUS} + Se_{TR} - Se_{TUS} * Se_{TR} - Covp$$

$$Sp_{JE} = P(TUS- \cap TR- | BP-) = Sp_{TUS} * Sp_{TR} + Covn$$

II-4-2- Bayesian modeling for stage two

In the second stage, CT results were available only when positive samples were found as determined by the simultaneous testing carried out during the first stage of the analysis (i.e., calves with either TR+ and/or TUS+ results). We assumed that CT was the gold standard and was independent of TUS and TR. The sensitivity (**Se_{CT}**) and specificity (**Sp_{CT}**) of CT was therefore fixed to 100%.

If considering the vector $a = (a_{11}, a_{10}, a_{0.})$ which denoted the observed numbers of calves for each combination in stage two where a_{11} corresponding to the number of calf positive in TUS and TR, a_{10} corresponding to the number of calf positive in TUS or TR, and $a_{0.}$ corresponding to the sum of the test negatives from the first stage. Let $pa = (pa_{11}, pa_{10}, pa_{0.})$ represented the respective probabilities from vector a . The multinomial sampling distribution was given by:

$$a \sim \text{multinomial}(n, (pa_{11}, pa_{10}, pa_{0.}))$$

where the multinomial cell probabilities were given by:

$$pa_{11} = Prev * Se_{JE} * Se_{CT} + (1-Prev)(1-Sp_{JE})(1-Sp_{CT}) = Prev * Se_{JE}$$

$$pa_{10} = Prev * Se_{JE} * (1-Se_{CT}) + (1-Prev)(1-Sp_{JE})Sp_{CT} = (1-Prev)(1-Sp_{JE})$$

$$pa_{0.} = 1 - pa_{11} - pa_{10}$$

II-4-3- Prior distributions

In our main model, the unknown parameters consisted of Se_{TUS} , Se_{TR} , Sp_{TUS} , Sp_{TR} , $Covp$, $Covn$, and the prevalence of thoracic lesions in one population, for a total of 7 unknown parameters. The degree of freedom was equal to 3 ($2^{\text{number of test}} - 1$) (Branscum et al., 2005). Since the degree of freedom was lower than the number of unknown parameters,

informative prior distributions on two parameters were needed to keep the model identifiable (Dendukuri et al., 2009). Beta distribution was used as priors for all parameters of interest. Among the unknown parameters, Se and Sp of TUS are the most accurately reported in the literature (Rabeling et al., 1998, Ollivett, 2014, Berman et al., 2019). Informative priors were therefore used for these parameters. The most probable value for TUS Se was set at 0.85, based on a 95% confidence it was ≥ 0.70 , corresponding to a β (23.9; 5.0) distribution. For TUS Sp, a mode of 0.95 was hypothesized, with 95% certainty Sp was ≥ 0.80 , corresponding to a β (42.6; 5.6) distribution. A vague prior distribution (β (1; 1)) was chosen for Se and Sp of TR and for the prevalence of thoracic lesions.

II-4-4- The sequential use of TR and TUS

To assess the sequential use of TR and TUS, the Se and Sp of using TR and TUS in parallel (**Se_p** and **Sp_p**) or in series (**Se_s** and **Sp_s**) were calculated. With a series interpretation, only animals that tested positive to both tests were considered positive. With a parallel interpretation, animals that tested positive to one test or both tests were considered positive. Using TR and TUS in series or in parallel was assessed as follows based on Dohoo (2012):

- from the parallel interpretation of the two tests:

$$Se_p = 1 - (1 - Se_{TR}) * (1 - Se_{TUS}) - Covp$$

$$Sp_p = Sp_{TR} * Sp_{TUS} + Covn$$

- from the series interpretation of the two tests:

$$Se_s = Se_{TR} * Se_{TUS} + Covp$$

$$Sp_s = 1 - (1 - Sp_{TR}) * (1 - Sp_{TUS}) - Covn$$

II-4-5- Comparison between TR and TUS

We used absolute differences to compare the Se and Sp of both tests with $Sediff = Se_{TUS} - Se_{TR}$ and $Spdiff = Sp_{TUS} - Sp_{TR}$. With this approach, a $Sediff$ (or $Spdiff$) > 0 means that $Se_{TUS} > Se_{TR}$ ($Sp_{TUS} > Sp_{TR}$) and $Sediff < 0$ means that $Se_{TUS} < Se_{TR}$ ($Sp_{TUS} < Sp_{TR}$).

II-4-6- Assessment of model sensitivity to assumptions

We made two assumptions in this study: (1) TUS and TR were conditionally dependent (assessing the same biological entity, thoracic lesions) and (2) CT was a gold standard. To assess model sensitivity to the first assumption, we ran a model considering both tests being independent (**model 2; Table 36**). For the second assumption, we considered CT as imperfect. We assumed that the Se and Sp of CT were at least superior to 0.5 to detect thoracic lesions based on his common use as a gold standard (Schwarz et al., 1999; Bourcier et al., 2014; Hu et al., 2014; Chavez et al., 2014; Fowler et al., 2017). The most probable value for CT Se and Sp were set at 0.70, with a 95% certainty it was ≥ 0.50 , corresponding to a β (13.3; 6.3) distribution (**model 3; Table 36**).

II-4-7- Assessment of model sensitivity to priors

To investigate the influence of the priors on the model, we considered our main model with relaxing informative priors on Se and Sp of TUS (Se_{TUS} : mode of 0.70; 95% lower bound 0.50 corresponding to a β (13.3; 6.3) distribution and Sp_{TUS} : mode of 0.80, 95% lower bound 0.50 corresponding to a β (7.6; 2.6) distribution) (**sensitivity analysis scenario; Table 36**).

II-4-8- Software

The Bayesian computations were implemented using public domain software (WinBUGS, version 1.4.3, MRC biostatistics Unit, Cambridge, UK). The WinBUGS program is shown **in Supplementary Methods**. Three chains were run with different initial values. The first 5,000 iterations of each chain were discarded as burn-in, whereas the next 105,000 were used to obtain posterior distributions (Brooks et Gelman, 1998). The convergence of the models was assessed by visual inspection of the history, density plots, and Brooks-Gelman-Rubin statistics. Autocorrelation plots were also verified to ensure proper behavior of the Markov chains. We used the DIC tool dialog box to estimate the deviance information criterion (**DIC**) and the number of parameters effectively estimated (**pD**) for the different models as fitting criteria. We used these two parameters as recommended in previous studies (Berkvens et al., 2006; Liu et al., 2014). The posterior distributions of each parameter were reported as medians and the corresponding 95% Bayesian credible interval (**BCI**). Using the Bayesian models, Se and Sp resulting from the parallel and series

interpretations, the positive and negative likelihood ratios, and the positive and negative predictive values were computed directly with WinBUGS.

II-4-9- Models assessment and fit

Model evaluation were performed as previously described (Liu et al., 2014). A model was concluded to fit the data well if (1) the value of pD was smaller or equal to the degree of freedom (equal to 3 ($2^{\text{number of test}} - 1$) in our study) (Berkvens et al., 2006), (2) most frequencies of the cells in the 2 x 2 tables comparing TUS and TR fell within their corresponding BCI, and (3) the 95% BCI of the estimated positive predictive value of the simultaneous testing included its actual value. The predictive value of the simultaneous testing (**PPV_{JE}**) was defined as follows based on Liu et al. (2014):

$$PPV_{JE} = Se_{JE} * prev / ((Se_{JE} * prev + (1 - Sp_{JE}) * (1 - prev))$$

We used the DIC to compare the different models. The model with the smallest DIC (a difference of at least 5 points) was the model that would best predict a replicated dataset with the same structure (Spiegelhalter et al., 2002).

III- RESULTS

III-1- Population

A total of 54 calves met the inclusion criteria during the study period. Three calves were then excluded following the owner's refusal of letting the animal participate in the study. An additional calf was excluded after receiving a diagnosis of bovine viral diarrhea virus infection, resulting in immediate isolation of the animal. The selection and classification of calves are presented in **Figure 11**. Among the 50 remaining calves, two calves were excluded based on the unavailability of CT images due to technical issues with the CT scanner, and one calf was excluded following the identification of a thoracic lesion on TR during the first review of the study, whereas TUS had been deemed to be exempt of lesions. The calf had already been discharged from the hospital at the time of review, preventing the performance of a CT study.

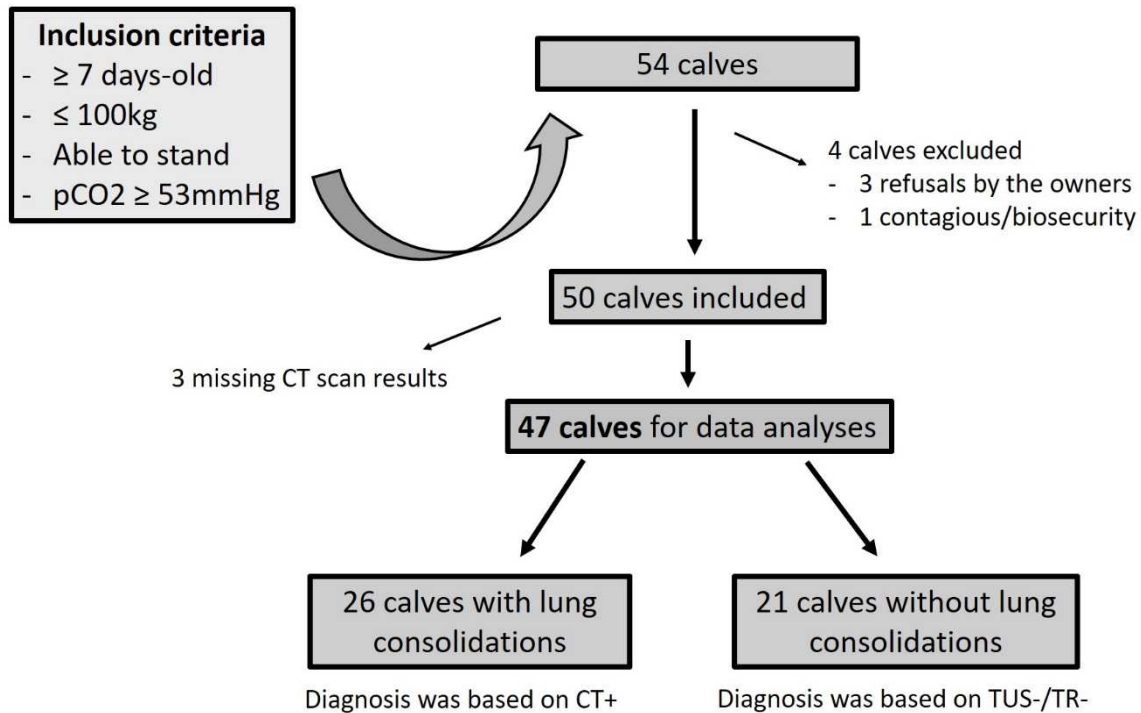


Figure 11. – During the study period, a total of 50 calves were enrolled in the study. They were then classified according to the results obtained by computed tomography (CT; deemed positive, CT+, if pulmonary lesions were present), thoracic ultrasonography (TUS; deemed negative, TUS -, in the absence of pulmonary lesions or when the lesions ≤ 1 cm), and thoracic radiography (TR; deemed negative, TR-, in the absence of thoracic lesions).

The descriptive data of the studied population of 47 calves are described in **Table 32**. No complications were reported using CT in sedated calves in this study.

<i>Qualitative variables</i>		<i>Numbers</i>	<i>Proportion (%)</i>
Breed	Holstein	37	79
	Jersey	5	11
	Red Holstein	3	6
	Ayrshire	2	4
Sex	Female	43	91
	Male	4	9
<i>Continuous variables</i>		<i>Mean</i>	<i>SD</i>
Age (days)		35.5	25
Weight (kg)		55.5	16.4

Tableau 32. – Descriptive data of the 47 calves included in the study (≥ 7 days old; $p\text{CO}_2 \leq 53$ mmHg; able to stand; ≤ 100 kg).

III-2- Thoracic lesions

The descriptive data of the different thoracic lesions found at CT, TUS, and TR are presented in **Table 33**. A total of 26 calves had lung consolidation on CT (CT+ = 26). One calf had a pneumothorax associated with lung consolidation on CT. Lung consolidations were also overrepresented on TUS (n = 24). Only two other types of lesions were identified on TUS (one cavitory lesion and one pneumothorax), both associated with lung consolidation (TUS+ = 24). All calves with thoracic lesions on TR (pneumothorax, interstitial pattern, or bronchial pattern) had an alveolar pattern (TR+ = 26).

Imaging exam	Number of calves	Comments
Thoracic ultrasonography		
Lung consolidation	24	
Pleural effusion	0	
Pneumothorax	1	With consolidations
Cavitory lesion	1	With consolidations
Absence of lesion	23	
Total TUS ^a	24	
Thoracic radiography		
Bronchial pattern	4	With an alveolar pattern
Interstitial pattern	13	With an alveolar pattern
Alveolar pattern	26	
Nodular pattern	1	With an alveolar pattern
Pleural effusion	0	
Pneumothorax	1	With an alveolar pattern
No lesion	21	
Total TR ^b	26	
Computed tomography		
Lung consolidation	26	
Pleural effusion	0	
Pneumothorax	1	With consolidations
Total CT ^c	26	

^a TUS +: having a thoracic lesion (consolidation (≥ 1 cm), pneumothorax, abscess, pleural effusion (≥ 1 cm)) in thoracic ultrasonography

^b CR+: having a thoracic lesion (bronchial/interstitial/alveolar pattern, pneumothorax or pleural effusion) in thoracic radiography

^c CT+: having a thoracic lesion (consolidation, abscesses, pneumothorax or pleural effusion) in thoracic computed tomography

Tableau 33. – Descriptive data of the thoracic lesions identified on thoracic ultrasonography (TUS), thoracic radiography (TR), and thoracic computed tomography (CT) in the 47 dairy calves included in the study.

III-3- Comparison between TUS and TR

The descriptive data summarizing the comparison between TUS and TR are presented in **Table 34**.

	First step			Second step			
	TR+	TR-	Total	lesions+	lesions-	Total	
TUS+	22	2	24	TUS+ and/or TR+	26	2	28
TUS-	4	19	23	TUS- and TR-	2.0 (0.10; 10.0)	16.0 (6.6; 21.5)	19
Total	26	21	47	Total	28.0 (22.0; 38.0)	19.0 (9.3; 25.2)	47

TUS +: having a thoracic lesion (consolidation (≥ 1 cm), pneumothorax, abscess, pleural effusion (≥ 1 cm)) in thoracic ultrasonography

TR+: having a thoracic lesion (bronchial/interstitial/alveolar pattern, pneumothorax or pleural effusion) in thoracic radiography

lesions+: having a thoracic lesion on the gold standard (CT) or estimate as having a lesion by the Bayesian framework

lesions-: Not having a thoracic lesion on the gold standard (CT) or estimate as not having a lesion by the Bayesian framework

Tableau 34. – Descriptive data summarizing the comparison between thoracic ultrasonography (TUS) and thoracic radiography (TR) in the 47 calves included in the study. The first step corresponds to the direct comparison between TUS and TR. The second step highlights the prior densities and median estimates (95% Bayesian credibility interval (BCI)) from the comparison of TUS and TR combined based on the presence or absence of computed tomography (CT) thoracic lesions.

The latent class model results are reported in **Table 35**. The prevalence of thoracic lesions was 0.60 (95%BCI: 0.46; 0.80). There was no difference in Se, Sp, PPV, and LR between both tests. The difference between Se and Sp of TUS and TR ($Se_{diff} = Se_{TUS} - Se_{TR}$ and $Sp_{diff} = Sp_{TUS} - Sp_{TR}$) were -0.04 (95%BCI: -0.20; 0.14) and 0.008 (95%BCI: -0.13; 0.23), respectively.

	Sensitivity (95% BCI)	Specificity (95% BCI)	PPV (95% BCI)	NPV (95% BCI)	LR+ (95% BCI)	LR- (95% BCI)
TUS	0.81 (0.65; 0.92)	0.90 (0.81; 0.96)	0.93 (0.84; 0.98)	0.77 (0.41; 0.91)	8.0 (4.1; 20.1)	0.21 (0.09; 0.39)
TR	0.86 (0.63; 0.99)	0.89 (0.67; 0.99)	0.92 (0.79; 1.0)	0.82 (0.36; 0.99)	7.9 (2.1; 135.3)	0.16 (0.01; 0.51)

Tableau 35. – Posterior densities and median estimates, including 95% Bayesian credibility interval (BCI) of test sensitivity, test specificity, predictive positive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+), and negative likelihood ratio (LR-), for thoracic ultrasonography (TUS) and thoracic radiography (TR) using a two-stage Bayesian method.

III-4- Sequential use of TR and TUS

The combined Se and Sp of both tests when used in parallel were 0.93 (95%BCI: 0.72; 1.00) and 0.84 (95%BCI: 0.64; 0.93), respectively. The combined Se and Sp of both tests when used in series were 0.75 (95%BCI: 0.56; 0.89) and 0.96 (95%BCI: 0.87; 1.0), respectively.

III-5- Assessment of the Bayesian model

The results from the two-stage Bayesian model using beta prior on Se and Sp of TUS fitted the data well, in accordance with the previously described criteria (**Supplementary Methods**). All the estimated values were close to their true values and fell within 95%BCI over their corresponding true values. A pD of approximately 1.9 indicated that the two-Bayesian strategy improved the identifiability of the models (Berkvens et al., 2006). The values indicated that the distributions of the replicated and actual data were similar. The different models used to analyze the influence of the assumptions and the priors are presented in **Table 36**. Since the 95%BCI overlaps, it was impossible to come up with a better model than the main model. Regardless of the model used, there was no difference between Se and Sp of both tests.

	Model 1		Model 2		Model 3		Sensitivity analysis scenario	
	Priors densities	Posterior densities	Priors densities	Posterior densities	Priors densities	Posterior densities	Priors densities	Posterior densities
Se _{TUS}	β (23.9; 5.0)	0.81 (0.65; 0.92)	β (23.9; 5.0)	0.85 (0.74; 0.94)	β (23.9; 5.0)	0.81 (0.64; 0.92)	β (13.3; 6.3)	0.64 (0.48; 0.83)
Se _{TR}	β (1.0; 1.0)	0.86 (0.63; 0.99)	β (1.0; 1.0)	0.93 (0.77; 1.0)	β (1.0; 1.0)	0.86 (0.62; 0.99)	β (1.0; 1.0)	0.83 (0.56; 0.96)
Sp _{TUS}	β (42.6; 5.6)	0.90 (0.81; 0.96)	β (42.6; 5.6)	0.91 (0.84; 0.96)	β (42.6; 5.6)	0.90 (0.81; 0.96)	β (7.6; 2.7)	0.68 (0.48; 0.96)
Sp _{TR}	β (1.0; 1.0)	0.89 (0.67; 0.99)	β (1.0; 1.0)	0.94 (0.81; 1.0)	β (1.0; 1.0)	0.89 (0.64; 0.99)	β (1.0; 1.0)	0.81 (0.12; 0.99)
Se _{CT}	Gold standard	1.0	Gold standard	1.0	β (13.3; 6.3)	0.84 (0.71; 0.93)	Gold standard	1.0
Sp _{CT}	Gold standard	1.0	Gold standard	1.0	β (13.3; 6.3)	0.65 (0.42; 0.84)	Gold standard	1.0
Prev	β (1.0; 1.0)	0.60 (0.46; 0.80)	β (1.0; 1.0)	0.56 (0.45; 0.66)	β (1.0; 1.0)	0.60 (0.45; 0.83)	β (1.0; 1.0)	0.76 (0.51; 0.96)
Covp	U (a; b)	0.05 (0.001; 0.17)	-	-	U (a; b)	0.05 (0.001; 0.17)	U (a; b)	0.15 (0.008; 0.21)
Covn	U (c; d)	0.03 (0.0008; 0.10)	-	-	U (c; d)	0.03 (0.0007; 0.10)	U (c; d)	0.04 (0.0009; 0.16)
DIC	-	22.4	-	22.8	-	26.4	-	20.8
pD	-	2.3	-	2.2	-	2.2	-	0.29

β : beta distribution; U: uniform distribution; Prev: the prevalence of thoracic lesions; Se_{TR}/Sp_{TR}: sensitivity/specificity of thoracic radiography; Se_{TUS}/Sp_{TUS}: sensitivity/specificity of thoracic ultrasonography. Covp/Covn: covariance of the tests for positive (p) or negative (n) bronchopneumonia cases. $a = ((\text{Se}_{\text{TR}}-1)*(1-\text{Se}_{\text{TUS}})/b = \min(\text{Se}_{\text{TR}}, \text{Se}_{\text{TUS}}) - \text{Se}_{\text{TR}}*\text{Se}_{\text{TUS}}$; $c = (\text{Sp}_{\text{TR}} - 1)*(1-\text{Sp}_{\text{TUS}})$; $b = \min(\text{Sp}_{\text{TR}}, \text{Sp}_{\text{TUS}}) - \text{Sp}_{\text{TR}}*\text{Sp}_{\text{TUS}}$; Model 1: Model considering TUS and TR conditionally dependant and CT as a gold standard; Model 2: Model considering TUS and TR independent and CT as a gold standard; Model 3: Model considering TUS and TR conditionally dependant and CT as an imperfect test with a vague prior distribution β (1.0; 1.0); Sensitivity analysis scenario: the model 1 with priors Se_{TUSmode}: 70%, 95% lower bound 50%/Sp_{TUSmode}: 80%, 95% lower bound 50%; DIC = deviance information criterion; pD = posterior mean of the deviance.

Tableau 36. – Posterior median and 95% credible intervals based on four plausible Bayesian latent class models for the prevalence of bronchopneumonia, the sensitivity (Se) and specificity (Sp) of thoracic radiography (TR), thoracic ultrasonography (TUS), and computed tomography (CT) used to diagnose BP in 47 dairy calves.

IV- DISCUSSION

In this study, we estimated the performances of TR and TUS to detect thoracic lesions (as a proxy for BP) in hospitalized dairy calves. We found no statistical difference between both tests in any estimated accuracy parameter (Se, Sp, PPV, NPV, Sediff, Spdiff, LR+, and LR-). Despite wide BCI of accuracy estimates in this study, the results showed that the median numerical value of sensitivities and specificities of TUS and TR were > 0.80 and that using TUS and TR in series was more specific than using both tests in parallel to detect thoracic lesions in dairy hospitalized calves. The accuracy of TUS alone was not different from the accuracy obtained by using both tests in series or in parallel.

In this study, we compared TUS and TR. Unlike previous studies (Reef et al., 1991; Braun et al., 1997; Jung and Bostedt, 2004), we found no difference between TUS and TR to detect thoracic lesions. This absence of difference could be real or could be secondary to a lack of power, preventing us from detecting any difference in Se, Sp, PPV, NPV, Sediff, Spdiff, LR+, and LR- between both tests (type II error) (Dohoo et al., 2012). The lack of power in this study explains the wide 95%BCI reported and the 95% BCI overlap between parameters. Indeed, no sample size calculation was estimated prior to the study since little information was available. Therefore, the number of 50 calves was decided arbitrarily based on previous reports in newborn calves ($n=56$) (Jung and Bostedt, 2004), in adult cattle ($n=3$) (Braun et al.; 1997), and one human study that found a difference between both diagnostic procedures with 52 pediatric participants (Ioro et al., 2015). To objectivize this lack of power, we performed a post hoc sample size calculation. A total of 430 calves would be needed to detect a difference in sensitivities of 0.81 vs 0.86 of TUS and TR, respectively, with a power of 80% (PROC POWER, SAS 9.4). This number is higher than anticipated. It is possible that either modality performs better in a different species or in a different subset population. Moreover, our population may have presented with a wider spectrum of disease ranging from normal lungs, to mild to severe consolidations in contrast to previous studies limited to severe cases of pneumonia. Based on our results, we believe it is essential to consider a large sample size when a comparison between TUS and TR is desired. Realistically, it would be unlikely or even impossible to obtain this number in a hospital setting and would raise major ethical questions relevant to the additional manipulations necessary for a project of this scale. Furthermore, it is unlikely that a

difference in sensitivity of 0.81 vs 0.86, although obtained with a sufficient power, would have any clinical impact on the selection of TR vs TUS to detect thoracic lesions in a dairy calf. In other words, considering the good performances of both procedures, carrying out such as study enrolling 430 calves and submitting a subset to CT may be unjustified.

Despite wide BCI of accuracy estimates in this study, sensitivities and specificities of TUS and TR to detect thoracic lesions in dairy calves were > 0.80 , and therefore comparable with previous reports showing a Se and a Sp of TUS to detect BP varying from 0.85 to 0.94 and from 0.95 to 1.0, respectively (Rabelling et al., 1998; Ollivett et al. 2014; Berman et al., 2019), and an excellent agreement ($\kappa > 0.6$) between TR and BP lesions found in necropsy or in CT (Jones et al., 1998; Fowler et al., 2017). In this study, we also reported the overall diagnostic ability of using both tests in addition with performances only. We did not find any difference in sensitivities of TUS and TR when both tests were used in parallel, in series, or alone. Indeed, in our study, TUS or TR alone was sufficiently accurate to detect thoracic lesions, and the performance of both tests did not improve our detection ability. In practice, TUS is easier to perform (need only one operator) and more easily accessible than TR by veterinarians (Ollivett and Buczinski, 2016; Buczinski et al., 2013). Thus, using TUS alone seems the most appropriate test for detecting thoracic lesions in dairy calves.

In this study, we compared TUS and TR based on CT as the gold standard, validating a more accurate and appropriate comparison for determining a superior test. The excellent ability of CT to detect thoracic lesions has led to its use as a gold standard in several studies carried out in companion animals and in humans comparing the diagnostic performances of TR and TUS (Schwarz et al., 1999; Bourcier et al., 2014; Hu et al., 2014; Chavez et al., 2014). The comparison in BP diagnostic accuracy between thoracic radiography and CT was also conducted in calves (Lubbers et al., 2007, Ohlerth et al., 2014). Apart from its frequent use in several studies for the detection of a broad spectrum of lesions and various species and its application to numerous lung diseases, the fact remains that the use of thoracic CT as the gold standard can be questioned. Indeed, Dohoo et al. 2012, defined a gold standard as a test with Se and Sp equal to 100%. Since such tests are rare, some authors rightly criticize the very essence of the term “gold standard”, since perfection is rarely

achieved. They further promote the term “reference standard test” for a test providing the highest Se/Sp among commonly available tests designed to detect specific conditions (Gold et al. 2010). Since, to our knowledge, no other ante-mortem test is superior to CT for the detection of thoracic lesions, we have defined CT as our gold standard in this study. This implies that CT would perfectly discriminate between calves with and without thoracic lesions and would therefore provide unbiased estimates of the diagnostic accuracy of TUS and TR. However, it is currently unknown whether CT has a Se and Sp of 100% for the detection of thoracic lesions in dairy calves in the absence of data comparing CT to macro/microscopic lesions. To assess this assumption, we ran a model where CT scan was considered as an imperfect test to detect thoracic lesions (**model 3; Table 36**). Performances of TUS and TR in this model were not different from those in our main model considering CT as a gold standard. The imperfection of CT had therefore no impact on our results. Despite this imperfection, we decided to present the inferences of the main model because it was in accordance with our study design and a simpler model.

Utilization of thoracic CT in calves has been limited to high-value animals in hospital settings due to its costs and availability (Lubbers et al., 2007; Ohlerth et al., 2014). Moreover, the calves must be sedated to perform this exam. In this study, because of economic considerations and negative side effects, we limited the need for CT by using a two-stage Bayesian framework for determining the accuracy of TUS and TR. Indeed, this approach could have estimated the status of negative calves in TUS and TR and therefore limited the costs, the need for sedation, and stress associated with their handling. Furthermore, this approach refers to the reduction alternatives mentioned in the three Rs in research Replacement, Reduction and Refinement (Balls, 1999). In our study, a two-step model was designed to reproduce the current clinical screening procedure in which animals negative for both tests were not investigated further with the classifier test. Of note, a one-stage model performed in two separate populations could have been used alternately (**Table 36**). In this model, the calves of the first population would have undergone all three diagnostics (i.e., TUS, TR and CT) while only TUS and TR would have been performed in the second population. The WinBUGS code for the one-stage model is shown in **Supplementary Methods**. Both models are mathematically equivalent (**Table 36**) and comparison between two-stage and one-stage models revealed similar results

(Supplementary Methods). The difference in DIC values between both models was < 2 . Since differences of 5 or more are commonly considered relevant in the literature, it is therefore difficult to conclude with high confidence on the superiority of one model over the other (Adrion and Mansmann, 2012). We, however, believe that based on our study design and the determination of which calf needed to undergo CT scans based on TR and TUS results, the two-stage model best fits our study design. Importantly, the two-stage Bayesian model better represents the diagnostic approach commonly used in hospital settings.

Our study assumed that TUS and TR were conditionally dependent. We considered these two tests to be conditional since they measure the same entity (thoracic lesions). Two tests can be positively correlated i.e., that the results change together, or negatively correlated i.e., that the results change in opposing directions. In this study, we assumed only a positive correlation between both tests. Indeed, since TUS and TR measure the same entity, it would be surprising to obtain systematically opposite results when comparing TUS and TR, in sick and healthy calves. For this reason, we deemed a negative correlation unlikely. However, the wide 95%BCI and the relatively close DIC and pD prevented us from choosing the best model. Moreover, as shown in Table 5, the priors impacted the posterior distributions, a probable consequence related to the small number of subjects in our population (Dunson et al., 2001). Nevertheless, no difference was present between TUS and TR, irrespective of the model. We decided to favour the priors used for the main model because they were based on what is reported in the literature and, therefore, more realistic.

In addition of the internal validity weakened by the lack of power, we recognise some limitations affecting the external validity in this study. First, the study population is a referral population. Most calves had received previous treatments that could have impacted the ability of both tests to detect thoracic lesions. Moreover, thoracic lesions in the form of lung consolidations were overrepresented in our study. The superiority of TUS or TR in the detection of other lesions such as pneumothorax, pleural effusion, or abscesses was therefore not possible to determine because of the small number of these other lesions. It is possible that TUS or TR is superior in detecting lesions other than lung consolidations in dairy calves, as reported in horses (Reef et al., 1991; Partlow et al., 2018). Another point

of discussion is that no link between the presence of thoracic lesions and BP was determined in this study. A further study is necessary to assess the performances of both tests in calves suspected of BP. Our study remains however, the first study in dairy calves to compare TUS and TR with a gold standard by using a two-stage Bayesian framework.

V- CONCLUSION

This study estimated Se and Sp of TUS and TR for the detection of thoracic lesions in hospitalized dairy calves. Since both tests were associated with similar performances when used alone or sequentially, we believe that TUS alone would be accurate enough for the detection of thoracic lesions in dairy calves. However, a study enrolling a higher number of individuals would be necessary to confirm these results. Comparison between TR and TUS in calves suspected of BP is also needed to assess the best utilisation of these tests in clinical cases.

First stage:						
	TR+	TR-				
TUS+	p ₁₁	p ₁₀				
TUS-	p ₀₁	p ₀₀				
<p>p₁₁ <- prev*(Se_tus*Se_tr+covp) + (1-prev)*((1-Sp_tus)*(1-Sp_tr)+covn)</p> <p>p₁₀ <- prev*(Se_tus*(1-Se_tr)-covp) + (1-prev)*((1-Sp_tus)*Sp_tr-covn)</p> <p>p₀₁ <- prev*((1-Se_tus)*Se_tr-covp) + (1-prev)*(Sp_tus*(1-Sp_tr)-covn)</p> <p>p₀₀ <- prev*((1-Se_tus)*(1-Se_tr)+covp) + (1-prev)*(Sp_tus*Sp_tr+covn)</p>						
Second stage:						
	CT+	CT-				
TUS+ or TR+	pa ₁₁	pa ₁₀				
TUS- and TR-	-	-	pa ₀₁			
<p>Assumption CT as a perfect test (Se_ct = sp_ct = 1)</p> <p>pa₁₁ <- prev* Se₁₁</p> <p>pa₁₀ <- (1-prev)*(1-Sp₁₀)</p> <p>pa₀₁ <- 1-pa₁₁-pa₁₀</p> <p>Assumption CT as an imperfect test (Se_ct= β (13.3; 6.3) and Sp_ct= β (13.3; 6.3))</p> <p>pa₁₁ <- prev* Se₁₁*Se_ct + (1-prev)*((1-Sp₁₁)*(1-Sp_ct))</p> <p>pa₁₀ <- prev* Se₁₁*((1-Se_ct) + (1-prev)*(1-Sp₁₀))*Sp_ct</p> <p>pa₀₁ <- 1-pa₁₁-pa₁₀</p> <p>Se₁₁ (TUS+ U TR+ BP+) <- Se_tus+Se_tr-Se_tus*Se_tr-covp</p> <p>Sp₁₁ (TUS- ∩ TR- BP-) <- Sp_tus*Sp_tr+covn</p>						
<p>Priors Se_tus ~ β (23.9; 5.0)</p> <p> Sp_tus ~ β (42.6; 5.6)</p>			<p>Priors Se_tus ~ β (23.9; 5.0)</p> <p> Sp_tus ~ β (42.6; 5.6)</p>			
<p>Assumption covariance between both test</p> <p>covp ~ duni(0, SeUB)</p> <p>covn ~ duni(0, SpUB)</p> <p>SeUB <- min(Se_tus, Se_tr) - Se_tus*Se_tr</p> <p>SpUB <- min(Sp_tus, Sp_tr) - Sp_tus*Sp_tr</p>			<p>Assumption covariance between both test</p> <p>covDp ~ duni(0, SeUB)</p> <p>covDn ~ duni(0, SpUB)</p> <p>SeUB <- min(Se_tus, Se_tr) - Se_tus*Se_tr</p> <p>SpUB <- min(Sp_tus, Sp_tr) - Sp_tus*Sp_tr</p>			
<p>Equivalence of equations</p>			a[1]	= p123[1]+ p123[3]+ p123[5]		
			a[2]	= p123[2] + p123[4] + p123[6]		
			a[3]	= p123[7] - 1 - p123[1] - p123[2] - p123[3] - p123[4] - p123[5] - p123[6]		

Tableau 37. – Comparison of the two-stage model the two-stage model used in this study with a one-stage model considering two populations: one population including calves where thoracic ultrasonography (TUS), thoracic radiography (TR), and thoracic computed tomography (CT) were performed, and another population including only calves where TUS and TR were performed.

SUPPLEMENTARY METHODS

I- WinBUGS code of the main model

WinBUGS code used for estimating diagnostic accuracy of thoracic ultrasonography (**TUS**) and thoracic radiography (**TR**) in 47 dairy hospitalized calves using a two-stage Bayesian latent class model.

Model {

#MODEL 1: CONSIDERING BOTH TESTS DEPENDENTS AND CT AS A PERFECT TEST

#FRIST-STAGE MODEL:

```
x[1:4] ~ dmulti(p[1:4], n)
p[1] <- prev*(Se_tus*Se_tr+covp) + (1-prev)*((1-Sp_tus)*(1-Sp_tr)+covn)
p[2] <- prev*(Se_tus*(1-Se_tr)-covp) + (1-prev)*((1-Sp_tus)*Sp_tr-covn)
p[3] <- prev*((1-Se_tus)*Se_tr-covp) + (1-prev)*(Sp_tus*(1-Sp_tr)-covn)
p[4] <- prev*((1-Se_tus)*(1-Se_tr)+covp) + (1-prev)*(Sp_tus*Sp_tr+covn)
```

```
covp ~ dunif(0, SeUB)
covn ~ dunif(0, SpUB)
SeUB <- min(Se_tus,Se_tr) - Se_tus*Se_tr
SpUB <- min(Sp_tus,Sp_tr) - Sp_tus*Sp_tr
rho <- covp / sqrt(Se_tus*(1-Se_tus)*Se_tr*(1-Se_tr))
rhon <- covn / sqrt(Sp_tus*(1-Sp_tus)*Sp_tr*(1-Sp_tr))
```

```
Seje <- Se_tus+Se_tr-Se_tus*Se_tr-covp
Spje <- Sp_tus*Sp_tr + covn
```

#SECOND-STAGE MODEL:

```
a[1:3]~dmulti(pa[1:3],n)
pa[1] <-prev*Seje
pa[2] <-(1-prev)*(1-Spje)
pa[3] <-1-pa[1]-pa[2]
```

PRIORS :

```
Se_tus~dbeta(23.9,5.0)
Se_tr~dbeta(1,1)
Sp_tus~dbeta(42.6,5.6)
Sp_tr~dbeta(1,1)
prev~dbeta(1,1)
```

#DIFFERENCE BETWEEN BOTH TESTS:

```
Sediff <- Se_tus - Se_tr
Spdiff <- Sp_tus - Sp_tr
Ses <- Se_tr*Se_tus + covp
Sps <- (1-((1-Sp_tr)*(1-Sp_tus)))-covn
Sep <- 1-((1-Se_tr)*(1-Se_tus))-covp
Spp <- Sp_tr*Sp_tus + covn
```

```

PPV_tus<- (prev*Se_tus)/((prev*Se_tus)+(1 - prev)*(1 - Sp_tus))
NVP_tus<- ((1 - prev)*Sp_tus)/(prev*(1 - Se_tus) + (1 - prev)*Sp_tus)
PPV_tr<- (prev*Se_tr)/((prev*Se_tr)+(1 - prev)*(1 - Sp_tr))
NVP_tr<- ((1 - prev)*Sp_tr)/(prev*(1 - Se_tr) + (1 - prev)*Sp_tr)
LRpos_tus <- Se_tus / (1-Sp_tus)
LRneg_tus <- (1-Se_tus) / Sp_tus
LRpos_tr <- Se_tr / (1-Sp_tr)
LRneg_tr <- (1-Se_tr) / Sp_tr

```

#MODEL CHECKING#

```

X11 <- n*prev*Se_tus
X10 <- n*(1-prev)*(1 - Sp_tus)
X01 <- n*prev*(1 - Se_tus)
X00 <- n*(1-prev)*Sp_tus
A11 <- n*prev*Seje
A10 <- n*(1-prev)*(1- Spje)
A01 <- n*prev*(1 - Seje)
A00 <- n*(1-prev)*Spje
A.1 <- A11 + A01
A.0 <- A10 + A00
PPVje<-Seje*prev/(Seje*prev+(1-Spje)*(1-prev))

```

#MODEL SENSITIVITY TO ASSUMPTIONS#

#MODEL 2 CONSIDERING BOTH TESTS INDEPENDENTS

```

x[1:4] ~ dmulti(p[1:4], n)
p[1] <- prev*Se_tus*Se_tr + (1-prev)*(1-Sp_tus)*(1-Sp_tr)
p[2] <- prev*Se_tus*(1-Se_tr) + (1-prev)*(1-Sp_tus)*Sp_tr
p[3] <- prev*(1-Se_tus)*Se_tr + (1-prev)*Sp_tus*(1-Sp_tr)
p[4] <- prev*(1-Se_tus)*(1-Se_tr) + (1-prev)*Sp_tus*Sp_tr

```

```

Seje<-Se_tus+Se_tr-Se_tus*Se_tr
Spje<-Sp_tus*Sp_tr

```

#MODEL 3 CONSIDERING CT AS AN IMPERFECT TEST

```

a[1:3]~dmulti(pa[1:3],n)
pa[1]<-prev*Seje*Se_ct + (1-prev)*(1- Spje)*(1-Sp_ct)
pa[2]<-prev*Seje*(1-Se_ct) + (1-prev)*(1- Spje)*Sp_ct
pa[3]<-1-pa[1]-pa[2]

```

PRIORS FOR CT:

```

Se_ct~dbeta(13.3,6.3)
Sp_ct~dbeta(13.3,6.3)

```

MODEL SENSITIVITY TO PRIORS

```

#Se_tus~dbeta(13.3,6.3)
#Sp_tus~dbeta(7.6,2.6)
}

```

Inits

```

list (prev=0.005,Se_tus=0.6, Sp_tus=0.8,Se_tr=0.6, Sp_tr=0.8)
list (prev=0.1,Se_tus=0.4,Sp_tus=0.6,Se_tr=0.4, Sp_tr=0.6)

```

data

```

list (n=47,x=c(22,2,4,19),a=c(26,2,19))

```

II- Models checking

Models checking the application of the two- stage Bayesian models for the comparison between thoracic ultrasonography (**TUS**) and thoracic radiography (**TR**).

Parameters	True value	Median (95% BCI)
PPV_{JE}	0.93	0.89 (0.78; 0.96)
X₁₁	22	22.8 (17.7; 28.0)
X₁₀	2	1.8 (0.61; 3.8)
X₀₀	18	17.1 (8.2; 22.8)
X₀₁	3	5.1 (2.0; 12.3)
A₁₁	26	25.6 (20.5; 30.5)
A₁₀	2	3.0 (1.2; 6.3)
DIC	22.4	
pD	2.3	

PPV_{JE} = positive predictive value of the simultaneous testing.

vector X = (X₁₁, X₁₀, X₀₁, X₀₀) represents the observed number of each combination of two test outcomes (1 or 0 for test positive or test negative).

A₁₁ and A₁₀ represent the outcomes of positive simultaneous testing (TUS or TR positive) when disease is present (A₁₁) or absent (A₁₀).

BCI = Bayesian credible intervals

DIC = Deviance Information Criterion

pD = Number of parameters effectively estimated

III- WinBUGS code of the alternative model

WinBUGS code used for estimating diagnostic accuracy of thoracic ultrasonography (TUS) and thoracic radiography (TR) in 47 dairy hospitalized calves using a one-stage Bayesian latent class model considering two populations: one population including calves where TUS, TR, and thoracic computed tomography (CT) were performed, and another population including only calves where TUS and TR were performed.

Model {

#MODEL 1: CONSIDERING BOTH TESTS DEPENDENTS AND CT AS A PERFECT TEST

n <- sum(t123[1:7])

t123[1:7] ~ dmulti(p123[1:7], n)

p123[1] <- prev*(Se_tus*Se_tr+covp)

p123[2] <- (1-prev)*((1-Sp_tus)*(1-Sp_tr)+covn)

p123[3] <- prev*(Se_tus*(1-Se_tr)-covp)

p123[4] <- (1-prev)*((1-Sp_tus)*Sp_tr-covn)

p123[5] <- (1-prev)*(Sp_tus*(1-Sp_tr)-covn)

p123[6] <- prev*((1-Se_tus)*Se_tr-covp)

p123[7] <- prev*((1-Se_tus)*(1-Se_tr)+covp) + (1-prev)*(Sp_tus*Sp_tr+covn)

```

covp ~ dunif(0, SeUB)
covn ~ dunif(0, SpUB)
SeUB<- min(Se_tus,Se_tr) - Se_tus*Se_tr
SpUB<- min(Sp_tus,Sp_tr) - Sp_tus*Sp_tr

# PRIORS
Se_tus~dbeta(23.9,5.0)
Se_tr~dbeta(1,1)
Sp_tus~dbeta(42.6,5.6)
Sp_tr~dbeta(1,1)
prev~dbeta(1,1)

#MODEL SENSITIVITY TO ASSUMPTIONS#

#MODEL 2 CONSIDERING BOTH TESTS INDEPENDENTS
p123[1] <- prev*Se_tus*Se_tr
p123[2] <- (1-prev)*(1-Sp_tus)*(1-Sp_tr)
p123[3] <- prev*Se_tus*(1-Se_tr)
p123[4] <- (1-prev)*(1-Sp_tus)*Sp_tr
p123[5] <- (1-prev)*Sp_tus*(1-Sp_tr)
p123[6] <- prev*(1-Se_tus)*Se_tr
p123[7] <- prev*(1-Se_tus)*(1-Se_tr) + (1-prev)*Sp_tus*Sp_tr

#MODEL 3 CONSIDERING CT AS AN IMPERFECT TEST
p123[1] <- prev*(Se_tus*Se_tr+covp)*Se_ct + (1-prev)*((1-Sp_tus)*(1-Sp_tr)+covn)*(1-Sp_ct)
p123[2] <- prev*(Se_tus*Se_tr+covp)*(1-Se_ct) + (1-prev)*((1-Sp_tus)*(1-Sp_tr)+covn)*Sp_ct
p123[3] <- prev*(Se_tus*(1-Se_tr)-covp)*Se_ct + (1-prev)*((1-Sp_tus)*Sp_tr-covn)*(1-Sp_ct) p123[4] <-
prev*(Se_tus*(1-Se_tr)-covp)*(1-Se_ct) + (1-prev)*((1-Sp_tus)*Sp_tr-covn)*Sp_ct
p123[5] <- prev*((1-Se_tus)*Se_tr-covp)*(1-Se_ct) + (1-prev)*(Sp_tus*(1-Sp_tr)-covn)*Sp_ct
p123[6] <- prev*((1-Se_tus)*Se_tr-covp)*Se_ct + (1-prev)*Sp_tus*(1-Sp_tr)*(1-Sp_ct)
p123[7] <- 1 - p123[1] - p123[2] - p123[3] - p123[4] - p123[5] - p123[6]

# PRIORS FOR CT
Se_ct~dbeta(13.3,6.3)
Sp_ct~dbeta(13.3,6.3)
}

Inits
list (prev=0.005,Se_tus=0.6, Sp_tus=0.8,Se_tr=0.6, Sp_tr=0.8)
list (prev=0.1,Se_tus=0.4,Sp_tus=0.6,Se_tr=0.4, Sp_tr=0.6)

data
list (t123=c(22, 0, 1, 1, 1, 3, 19))

#t123 is of the form:
c( (TUS +; TR +; CT +), (TUS +; TR +; CT -), (TUS +; TR -; CT +), (TUS +; TR -; CT -),
(TUS -; TR +; CT +), (TUS -; TR +; CT -), (TUS -; TR -)

```

I- Comparison of both models

Comparison of the posterior medians and 95% credible intervals of the two-stage model used in this study with a one-stage model considering two populations: one population including calves where thoracic ultrasonography (**TUS**), thoracic radiography (**TR**), and thoracic computed tomography (**CT**) were performed, and another population including only calves where TUS and TR were performed.

	Model 1		Model 2		Model 3	
	Posterior densities <i>Two-stage</i>	Posterior densities <i>One-stage</i>	Priors densities <i>Two-stage</i>	Posterior densities <i>One-stage</i>	Priors densities <i>Two-stage</i>	Posterior densities <i>One-stage</i>
Se_{TUS}	0.81 (0.65; 0.92)	0.82 (0.66; 0.91)	0.85 (0.74; 0.94)	0.85 (0.74; 0.93)	0.81 (0.64; 0.92)	0.82 (0.66; 0.92)
Se_{TR}	0.86 (0.63; 0.99)	0.88 (0.68; 0.98)	0.93 (0.77; 1.0)	0.93 (0.78; 0.99)	0.86 (0.62; 0.99)	0.87 (0.65; 0.99)
Sp_{TUS}	0.90 (0.81; 0.96)	0.90 (0.81; 0.96)	0.91 (0.84; 0.96)	0.91 (0.82; 0.96)	0.90 (0.81; 0.96)	0.90 (0.80; 0.96)
Sp_{TR}	0.89 (0.67; 0.99)	0.90 (0.70; 0.99)	0.94 (0.81; 1.0)	0.92 (0.77; 0.99)	0.89 (0.64; 0.99)	0.89 (0.61; 0.99)
Se_{CT}	1.0	1.0	1.0	1.0	0.84 (0.71; 0.93)	0.85 (0.72; 0.94)
Sp_{CT}	1.0	1.0	1.0	1.0	0.65 (0.42; 0.84)	0.66 (0.44; 0.85)
Prev	0.60 (0.46; 0.80)	0.61 (0.44; 0.81)	0.56 (0.45; 0.66)	0.56 (0.42; 0.70)	0.60 (0.45; 0.83)	0.59 (0.41; 0.83)
Cov_p	0.05 (0.001; 0.17)	0.05 (0.0002; 0.16)	-	-	0.05 (0.001; 0.17)	0.05 (0.001; 0.17)
Cov_n	0.03 (0.0008; 0.10)	0.02 (0.0006; 0.08)	-	-	0.03 (0.0007; 0.10)	0.02 (0.0006; 0.09)
DIC	22.4	21.7	22.8	20.4	26.4	27.5
pD	2.3	3.4	2.2	2.7	2.2	2.6

Prev: the prevalence of thoracic lesions; Se_{TR}/Sp_{TR}: sensitivity/specificity of thoracic radiography; Se_{TUS}/Sp_{TUS}: sensitivity/specificity of thoracic ultrasonography. Cov_p/Cov_n: covariance of the tests for positive (p) or negative (n) bronchopneumonia cases. Model 1: Model considering TUS and TR conditionally dependant and CT as a gold standard; Model 2: Model considering TUS and TR independent and CT as a gold standard; Model 3: Model considering TUS and TR conditionally dependant and CT as an imperfect test with a vague prior distribution β (1.0; 1.0); DIC = deviance information criterion; pD = posterior mean of the deviance

CONCLUSION ARTICLE 3 : Dans cet ARTICLE 3, nous avons comparé lequel des tests d'imagerie médicale entre l'échographie et la radiographie thoracique était meilleur pour DÉTECTER LES LÉSIONS PULMONAIRES de la BPI chez les veaux de race laitière. Nous avons montré qu'il n'y avait PAS DE DIFFÉRENCE de Se et Sp entre les deux tests. Cependant, il faut garder en tête que le nombre limité de veaux dans cette étude a pu limiter la faculté à détecter une différence. Maintenant que nous avons comparé la radiographie et l'échographie thoracique pour détecter des lésions pulmonaires de BPI, continuons en comparant les performances de ces deux tests pour détecter cette fois-ci le statut actif de la BPI nécessitant un traitement.

ARTICLE 4

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

Publié dans le *Journal of Veterinary Internal Medicine* 2021 Jul;35(4):2058-68.

Julie Berman ^a, Isabelle Masseur ^a, Gilles Fecteau ^a, Sébastien Buczinski ^a, David Francoz ^a

^a From the Département des sciences cliniques, Faculté de médecine vétérinaire, Université de Montréal, Saint-Hyacinthe, QC, Canada, J2S 2M2.

Abbreviation: ACVIM-FA = American College of Veterinary Internal Medicine – Food Animals; ACVR = American College of Veterinary Radiology; BP = Bronchopneumonia; CI= Confidence intervals; Kalpha = Krippendorff’s alpha; LR+ = Positive Likelihood Ratio; LR- = Negative Likelihood Ratio; Se = Sensitivity; Sp = Specificity; TR = Thoracic Radiography; TR+= Presence of bronchial/interstitial/alveolar pattern, and/or pneumothorax and/or pleural effusion on thoracic radiography; TR-= Absence of bronchial/interstitial/alveolar pattern, pneumothorax or pleural effusion on thoracic radiography; TUS = Thoracic Ultrasonography; TUS+ = Presence of consolidation (≥ 1 cm), and/or evidence of a pneumothorax, and/or abscess and/or relevant (≥ 1 cm) pleural effusion on TUS; TUS - = Absence of consolidation (≥ 1 cm), and/or evidence of a pneumothorax, and/or abscess and/or relevant (≥ 1 cm) pleural effusion on TUS; VMTH = Veterinary Medical Teaching Hospital.

ABSTRACT

Objectives: 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. The objectives were 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.

Methods: Fifty hospitalized dairy calves (≥ 7 days old; ≤ 100 kg; standing; $p\text{CO}_2 \geq 53$ mmHg; any reason of presentation) were sampled. 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 inter-expert agreement. The sensitivities (**Se**) and specificities (**Sp**) of TUS and TR alone and in series or parallel were compared (McNemar's test; $p < 0.05$).

Results: Inter-expert 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 analysed alone, in series or in parallel.

Conclusion: TUS 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.

Key words: Bovine respiratory disease; diagnostic imaging tests; cattle; panel diagnosis method; kappa.

I- INTRODUCTION

Infectious bronchopneumonia (**BP**) in cattle is an infection of the lower respiratory tract with a multifactorial etiology (Woolums, 2015). In the absence of established consensus (Buczinski and Pardon, 2020), several studies have relied on one 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 (e.g., abnormal neutrophil count, or presence of toxic neutrophils, bands, or both modalities) combined with an increase of acute phase protein concentration (e.g., fibrinogen, haptoglobin), and shows metabolic abnormalities reflecting impaired gas exchange (Reinhold et al., 2002, Woolums, 2015, Berman et al., 2019). Active BP is commonly associated with lung lesions characterized on thoracic radiography (**TR**) by focal or multifocal areas of alveolar and interstitial patterns (Tegtmeier and Arnbjerg,

2000), and on thoracic ultrasonography (**TUS**) by lung consolidation (Berman et al., 2019). In contrast, non-active BP implies that clinical signs and hematologic abnormalities have resolved with or without remaining TR or TUS lung lesions (Ollivett et al., 2015). Resolution of lung opacities might lag behind clinical and hematological resolution of infection or scars might persist long after the initial infection is cured (Surén et al., 2008). In dairy calves, failure to accurately diagnose active BP and prolonged use of antimicrobials in non-active BP are associated with major economic losses and overuse of antimicrobials (Windeyer et al., 2014, Teixeira et al., 2017b, Dunn et al., 2018). 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 (Buczinski and Pardon, 2020). Unfortunately, both clinical tools provide suboptimal performances in detecting active BP with sensitivities (**Se**) (i.e. the ability of a test to detect a calf with active BP) and specificities (**Sp**) (i.e. the ability of a test to detect a healthy calf) of 62 to 100% and 62 to 87% for physical examination (Buczinski et al., 2015a, Love et al., 2016, Maier et al., 2019), and 73% and 53% for thoracic auscultation, respectively (Buczinski et al., 2016b). Development of superior diagnostic strategies will improve the diagnosis of active BP.

Thoracic ultrasonography and TR detect thoracic lesions in dairy calves (Farrow, 1999, Ollivett and Buczinski, 2016). TUS is a non-invasive, inexpensive, and relatively easy to interpret imaging test (Ollivett and Buczinski, 2016). Previous on-farm studies have enabled TUS detection of active BP in pre-weaned dairy calves with Se and Sp of 89% and 95%, respectively (Berman et al., 2019). However, TUS is limited to detecting lesions beneath the chest wall (Babkine and Blond, 2009). In comparison, TR detects thoracic lesions in calves with Se and Sp of 86% and 89%, respectively (Berman et al., 2020). However, its ability to discriminate between active and non-active BP in calves remains unknown and expertise required for interpretation limits its use in practice (Farrow, 1999). In contrast to TUS, TR can detect lesions deep into the lung parenchyma and provides an overview of the entire lung, displayed on one or two 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 (Berman et al., 2020). These data are inconsistent with those reported for the detection of active BP in neonatal calves for which TUS was less sensitive than TR (Jung and Bostedt, 2004), while TUS was more sensitive in detecting the same condition in adult cattle (Reef et al., 1991, Braun et al., 1997). 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 (i.e., test with a Se and Sp of 100%) has been validated for the antemortem diagnosis of active BP in dairy calves (Buczinski and Pardon, 2020). To overcome this deficiency, results from a series of imperfect tests (i.e., test with a Se and Sp < 100%) can be combined to construct a reference standard outcome based on a panel diagnosis method (**PDM**) (i.e., a consensus procedure among experts concluding on whether each calf has active BP or not based upon a set of imperfect test results) (Rutjes et al., 2007). 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 non-active BP could more closely reflect the diagnostic process taking place in a clinical setting (Rutjes et al., 2007). 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 (Reitsma et al., 2009).

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 (i.e., 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.

II- 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 (Cohen et al., 2016). The study protocol was approved by the institutional ethical committee (#20160111).

II-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. (Berman et al., 2020) 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) (Jung and Bostedt, 2004) and one human study revealing differences between diagnostic procedures in 52 pediatric participants (Iorio et al., 2015). Fifty calves of various dairy breeds (i.e., 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 pCO₂ ≥ 53 mmHg or presented in decubitus were excluded to allow immediate treatment intervention (Wilson et al., 2015). Calves for which the consent form was not signed in time for enrolment or considered unsuitable for transport to the diagnostic imaging suite (e.g., biosecurity) were also excluded (**Figure 12**).

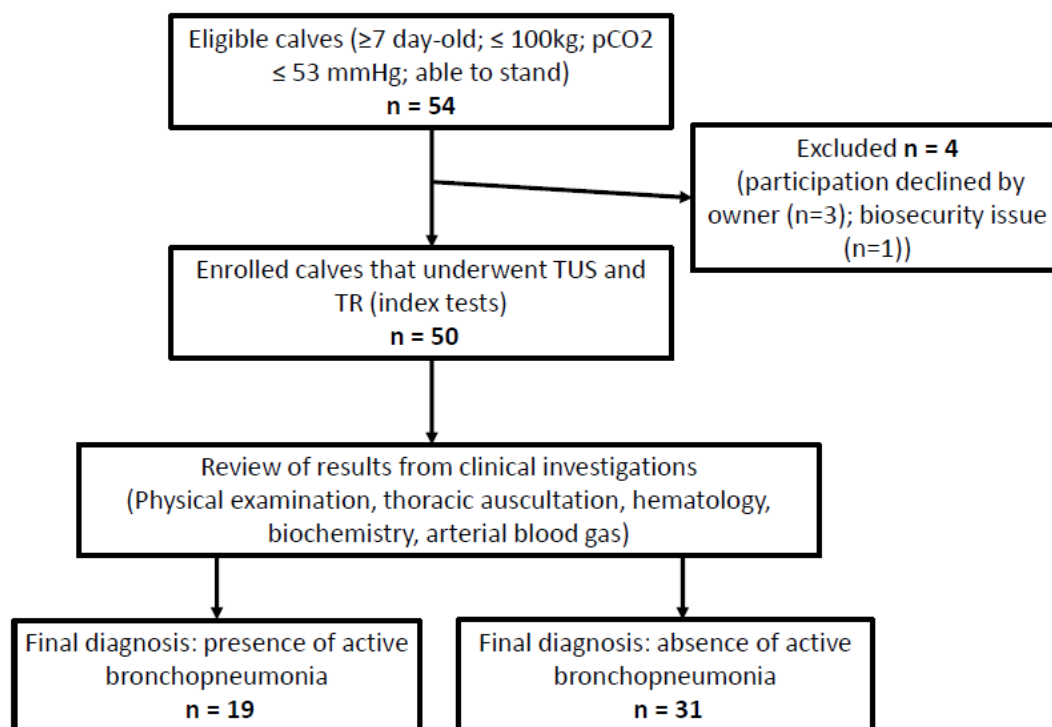


Figure 12. – Flowchart illustrating the overall flow of calves in the study.

II-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 one single 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, complete blood count (CBC) and serum biochemistry profile. Collection of data is detailed in **Table 38**.

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* An evaluation sheet was used to record the location of abnormal sounds, as previously described in Buczinski, et al. (2014).	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. (2009). The analysis took place less than 5 min after collection. PH, pO ₂ , pCO ₂ , SatO ₂ and HCO ₃ ⁻ were obtained using an automated analyzer (ABL 80 FLEX, Radiometer America, Inc., Brea, CA, USA).	pH, pO ₂ , pCO ₂ , SatO ₂ and HCO ₃ ⁻ values
Complete Blood Count	Twenty mL 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, NY, USA).	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.
Serum biochemistry profile	Twenty mL 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 Dx C 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 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.

Tableau 38. – Relevant information collected by one single ACVIM-FA board certified veterinarian on the 50 calves included in the study and used for the panel diagnosis method.

Component test results were reported on a single spreadsheet (MS Excel, Microsoft Corporation, Redmond, WA) facilitating blind review (**Supplementary Results**). Three experts consisting of board-certified specialists from the American College of Veterinary Internal Medicine – Food Animals (**ACVIM-FA**) (Expert 1, Expert 2, Expert 3), unaware of the TUS and TR results, independently reviewed the spreadsheet five months after the end of the study. Aiming for each calf to be evaluated by at least two 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 two 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.

II-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 subsequent to treatment initiation or disease progression.

II-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 (Ollivett and Buczinski, 2016). A portable ultrasound unit equipped with a 7.5 MHz linear transducer (Sonoscape S6V, Shenzhen, China) was used.

The real-time sonographic evaluation of the pleurae and lungs is described in **Figure 13**. 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 (Buczinski et al., 2014).

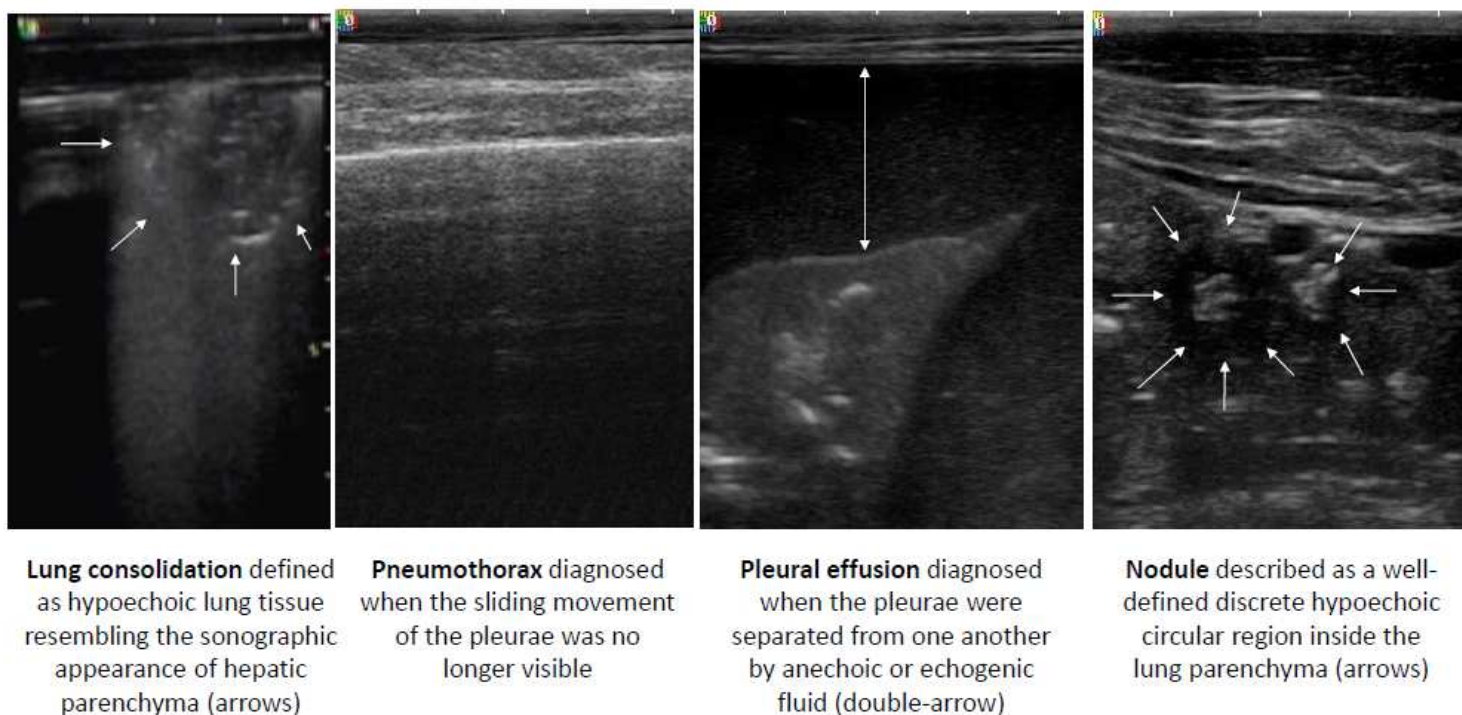


Figure 13. – Ultrasound images corresponding to the four lesions assessed in dairy calves based on Babkine and Blond (2009). Pleural irregularities, B-lines, and comet-tails were overlooked since their clinical relevance in calves is controversial (Ollivett et al., 2015).

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 (Buczinski et al., 2013). 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.

II-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 Vertex 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 (Berman et al., 2020).

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 ten months after enrolment of the last calf in the study. Radiographic parameters used to determine if TR was considered positive (TR+) are described in **Table 40**. Thoracic radiographic studies exempt of any of the listed thoracic lesions were considered negative (TR-).

II-4- Statistical analysis

All statistical analyses were performed using commercial software (SAS v 9.4, SAS, Cary, NC) and R statistical software (R Core Team 2020) (Proutskova, 2017). Additional details of statistical analyses are present in **Supplementary Methods**.

II-4-1- Assessment of the inter-expert agreement

To determine whether expert experience influenced the classification of calves with or without active BP, the inter-expert agreement was calculated using a Krippendorff's alpha (**Kalpha**; kripp.boot package <https://github.com/MikeGruz/kripp.boot>) (Zapf et al., 2016, Pardon et al., 2019). Briefly, the Kalpha represents the pairwise agreement between two experts, averaged over all expert pairs and number of calves, handling for missing values (Gwet, 2014). The confidence intervals (**CI**) were obtained using 5000 bootstrap replicates in the absence of validated parametric method to calculate them (Zapf et al., 2016). In the absence of specific benchmarks of agreement, Kalpha was interpreted according to Landis and Koch (1977).

II-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+**) (i.e., 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 (i.e., false positive fraction = $1 - Sp$)), and negative likelihood ratio (**LR-**) (i.e., percentage of calves with active BP classified as negative based on the test result (i.e., 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 (Dohoo et al., 2012, Timsit et al., 2018). Confidence intervals were calculated using exact binomial methods for Se and Sp, and using a natural logarithmic transform and the delta method for LR (Pepe, 2003b).

II-4-3- Comparison of TUS and TR

TUS 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. (2012) Considering an assumption of covariance between both tests and the paired design of the study, a two-sided McNemar's test was performed (**Supplementary Methods**).

Additionally, as recommended by Pepe (2003b) 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 (Pepe, 2003b). Relative probabilities were estimated as follows:

$$rSe (TUS, TR) = Se \text{ of TUS} / Se \text{ of TR}$$

$$rSp (TUS, TR) = Sp \text{ of TR} / Sp \text{ of TUS}$$

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 (2003b) (**Supplementary Methods**).

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% confidence intervals for the ratios of Se and Sp (Pepe, 2003b).

II-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 one 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. (2012). Confidence intervals were estimated using an exact binomial method, (Pepe, 2003b) (**Supplementary Methods**).

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 two-sided McNemar's test. A p -value < 0.05 was considered significant (**Supplementary Methods**).

III- RESULTS

III-1- Study sample

Fifty-four calves met the inclusion criteria. **Figure 12** 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 39**.

<i>Qualitative variables</i>		<i>Numbers</i>	<i>Proportion (%)</i>
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 ^e	3	6
Treatment prior to admission ^f	Yes	40	80
	No	10	20
<i>Continuous variables</i>		<i>Mean</i>	<i>SD</i>
Age (days)		30	24.5
Weight (kg)		53	16

^a Digestive: diarrhea, colic, regurgitation, anorexia, intestinal obstruction, bloat; ^b Locomotor: septic arthritis, contracture, laceration; ^c Umbilical: hernia, urachus; ^e Other: Fever of unknown origin.

^f Treatment prior to admission: antimicrobials and/or anti-inflammatory.

Tableau 39. – Descriptive data of the 50 calves recruited in the study (≥ 7 days-old; $pCO_2 \leq 53$ mmHg; ≤ 100 kg; able to stand).

III-2- Clinical diagnosis of active BP

For each calf, results of component tests and classification by experts are shown in **Supplementary Results**. There was no missing data. Diagnosis of active BP was consensual between two experts for 80% (n = 40) of the study sample. Ten remaining calves required assessment by a third expert (highlighted in dark grey). 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)).

III-3- Performances of TUS and TR

Thoracic lesions identified on TUS and TR are presented in **Table 40**.

	Thoracic lesions	Groups	
		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: (i) lung consolidation defined as hypoechoic lung tissue resembling the sonographic appearance of hepatic parenchyma; (ii) pneumothorax diagnosed when the sliding movement of the pleurae was no longer visible; (iii) pleural effusion diagnosed when the pleurae were separated from one another by anechoic or echogenic fluid; and (iv) nodule described as a well-defined discrete hypoechoic circular region inside the lung parenchyma (Babkine and Blond, 2009).

^b Thoracic radiography: (i) bronchial pattern defined as the presence of airways with prominent and/or thickened walls; (ii) interstitial pattern defined as the presence of an area of the lung field with increased opacity combined with blurred vascular borders; (iii) nodule defined as the presence of a focal area of increased soft tissue opacity (< 3 cm) with relatively well-defined borders; (iv) alveolar pattern defined as the presence of an area of increased soft tissue opacity silhouetting with pulmonary vessels located within the opacified area; (v) 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 (vi) pneumothorax diagnosed based on the presence of air between the lung lobes and the thoracic wall causing retraction of the lung margins (Berman et al., 2020).

Tableau 40. – 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.

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 non-active 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 three and two calves with active BP, respectively. **Table 41** shows descriptive data summarizing the comparison of TUS and TR in calves with and without active BP.

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

TUS+: presence of lung consolidation (depth \geq 1cm), pneumothorax, pleural effusion, or nodule (depth \geq 1cm) on thoracic ultrasonography

TUS-: absence of lung consolidation (depth \geq 1cm), pneumothorax, pleural effusion or nodule (depth \geq 1cm) 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

Tableau 41. – 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.

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 42**). 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 42**.

	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)

Tableau 42. – 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.

III-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; $p > 0.05$). 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; $p = 0.06$). 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.

III-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 a 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; $p < 0.05$). 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; $p > 0.05$).

IV- 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 (Schmidt and Factor, 2013, Buczinski and O'Connor, 2016c, Buczinski and Pardon, 2020). Secondly, the classification of active or non-active 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 (Schmidt and Factor, 2013, Buczinski and O'Connor, 2016c). Thirdly, in contrast to studies reporting observational comparisons of TUS and TR (Reef et al., 1991, Braun et al., 1997, Jung and Bostedt, 2004), 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 $n = 56$ (Jung and Bostedt, 2004), $n = 2$ (Braun et al., 1997), or $n = 1$ (Reef et al., 1991)). 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 (Berman et al., 2020) or latent class analysis (Berman et al., 2019).

In this study, calves were classified as having active or non-active 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 (Rutjes et al., 2007, Reitsma et al., 2009). Indeed, two limitations resulted from this imperfect gold standard. Firstly, a misclassification error led to an underestimation of Se and Sp of both tests (Schmidt and Factor, 2013, Buczinski and O'Connor, 2016c). However, this misclassification error was independent of TUS and TR results (non-differential 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, three experienced and qualified experts were selected, a number reportedly sufficient to provide reproducible diagnoses in human

medicine (Van Houten et al., 2019). Despite high levels of qualification, the inter-expert agreement ($K_{\text{alpha}} = 0.58$) showed moderate agreement highlighting the current limits and difficulties in classifying a calf as having active or non-active BP based on clinical parameters 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 (Berman et al., 2020). 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 (Berman et al., 2019). This result is also similar to the Se of TUS reported in chronic BP with an Se of 85% (Rabeling et al., 1998) and in subclinical BP (i.e. calves with lung lesions but without clinical signs) with an Se of 94% (Ollivett et al., 2015). This result is also similar to the Se found for detecting thoracic lesions in the same study population (Se of 81%) (Berman et al., 2020).

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 (Berman et al., 2020) and an Sp of TUS > 90% in detecting chronic, subclinical or active BP (Rabeling et al., 1998, Ollivett et al., 2015, Berman et al., 2019). 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 (Surén et al., 2008). 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 non-active 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% *versus* 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 (Babkine and Blond, 2009), 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 post-mortem examination, and consequently, do not have a clear clinical relevance (Ollivett et al., 2015). We also only considered lesions deeper than or equal to 1 cm as clinically relevant (Buczinski et al., 2014). 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 (Farrow, 1999, Fowler et al., 2017). 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 (e.g., 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 (e.g., 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 (Reef et al., 1991, Braun et al., 1997, Jung and Bostedt, 2004). The exact reason for this discrepancy is uncertain, although our study has a few dissimilarities with respect to these previous studies: study sample (newborns (Jung and Bostedt, 2004) and adults (Reef et al., 1991, Braun et al., 1997) *vs* calves) and clinical cases included (including various pneumopathies (Jung and Bostedt, 2004) or only

focusing on severe pleuropneumonia (Reef et al., 1991, Braun et al., 1997) 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 population 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 ten 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 length of the study. Unnecessary manipulations on a large number of calves, despite their non-invasiveness, 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 towards 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.

V- 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).

SUPPLEMENTARY METHODS

I- Software

All statistical analyses were performed using commercial software (SAS v 9.4, SAS, Cary, NC) and R statistical software (R Core Team 2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/> using kripp.boot package (Proutskova, 2017).

II- Inter-expert agreement

The Krippendorff's alpha (Kalpha) was calculated using kripp.boot package <https://github.com/MikeGruz/kripp.boot>. The confidence intervals (CI) were obtained using 5000 bootstrap replicates as described by Zapf et al. (2016). The benchmarks of Landis and Koch (1977) (poor agreement for values below 0.20; slight agreement for values between 0.21 and 0.40; moderate agreement for values between 0.41 and 0.60; good agreement for values between 0.61 and 0.80; and, very good agreement for values between 0.81 and 1.00) was used to interpret Kalpha.

III- Covariance of TUS and TR

Covariance between both tests was estimated as follows (Dohoo et al., 2012):

$$\text{Covp} = p(\text{TUS+} \cup \text{TR+}) - \text{Se of TUS} * \text{Se of TR}$$

$$\text{Covn} = p(\text{TUS-} \cup \text{TR-}) - \text{Sp of TUS} * \text{Sp of TR}$$

Where Covp is the covariance between TUS and TR in calves with active BP; \cup is the union (OR) term between 2 conditions and therefore $p(\text{TUS+} \cup \text{TR+})$ is the proportion of calves with active BP with a positive result on both tests; Covn is the covariance between TUS and TR in calves without active BP; and $p(\text{TUS-} \cup \text{TR-})$ is the proportion of calves without active BP with a negative result on both tests.

IV- McNemar's test

A two-sided McNemar's test was performed to compare the marginal probabilities, i.e., the difference between the proportion of calves positive on TUS but negative on TR and the proportion of calves negative on TUS but positive on TR. As recommended by Kim and Lee (2017), the difference between Se of TUS and Se of TR was estimated by assessing the difference between marginal probabilities in calves with active BP. The difference between Sp of TUS and Sp of TR was estimated by assessing the difference between marginal probabilities in calves without active BP. A p -value < 0.05 was considered significant to reject the null hypothesis H_0 : Se of TUS = Se of TR or Sp of TUS = Sp of TR.

V- Relative probabilities of TUS and TR

Relative probabilities were estimated as follows (Pepe, 2003b):

$$rSe(TUS, TR) = Se \text{ of TUS} / Se \text{ of TR}$$

$$rSp(TUS, TR) = Sp \text{ of TR} / Sp \text{ of TUS}$$

The difference between both tests was assessed by calculating the joint 95% confidence intervals of both $rSe(TUS, TR)$ and $rSp(TUS, TR)$ as described by (Pepe, 2003b) as:

$$Var(\log rSe(TUS, TR))$$

$$= \frac{n(Dp) \times Se \text{ of TUS} + n(Dp) \times Se \text{ of TR} - 2n((TUS + \cup TR +)|Dp)}{Se \text{ of TUS} \times Se \text{ of TR} \times n(Dp)^2}$$

$$Var(\log rSp(TUS, TR))$$

$$= \frac{n(Dn) \times (1 - Sp \text{ of TUS}) + n(Dn) \times (1 - Sp \text{ of TR}) - 2n((TUS + \cup TR +)|Dn)}{(1 - Sp \text{ of TUS}) \times (1 - Sp \text{ of TR}) \times n(Dn)^2}$$

Where $n(Dp)$ and $n(Dn)$ correspond to the number of calves with or without active BP, respectively; $n((TUS + \cup TR +)|Dp)$ corresponds to the number of calves with active BP positive on TUS and TR; and $n((TUS + \cup TR +)|Dn)$ corresponds to the number of calves without active BP classified as positive on TUS and TR. 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% confidence intervals for ratios of Se and Sp (Pepe, 2003b).

VI- Sequential used of TR and TUS

Considering the covariance between both tests, TR and TUS in series or in parallel were assessed as follows based on Dohoo et al. (2012):

- from parallel interpretation of both tests where \cup represent the union of 2 ensembles (OR Boolean operator):

$$Se(TUS \cup TR) = 1 - (1 - Se \text{ of TR}) * (1 - Se \text{ of TUS}) - Covp$$

$$Sp(TUS \cup TR) = Sp \text{ of TR} * Sp \text{ of TUS} + Covn$$

- from the series interpretation of both tests where \cap represents the intersection of 2 ensembles (AND Boolean operator):

$$Se(TUS \cap TR) = Se \text{ of TR} * Se \text{ of TUS} + Covp$$

$$Sp(TUS \cap TR) = 1 - (1 - Sp \text{ of TR}) * (1 - Sp \text{ of TUS}) - Covn$$

Confidence intervals were estimated using an exact binomial method (Pepe, 2003b).

SUPPLEMENTARY RESULTS

For each calf, results of component tests and classification by experts are shown in the following table.

CONCLUSION ARTICLE 4 : Dans cet ARTICLE 4, nous avons comparé les performances de l'échographie thoracique et de la radiographie thoracique pour DÉTECTER LES BPI ACTIVES nécessitant un traitement d'antibiotiques et/ou anti-inflammatoires. Tout comme avec les lésions pulmonaires dans l'article précédent, AUCUNE DIFFÉRENCE de Se et Sp n'a été détectée entre les deux tests d'imagerie médicale. Cependant, le faible nombre de veaux inclus dans cette étude a pu limiter la détection d'une différence entre les deux tests. Les résultats de l'ARTICLE 3 et l'ARTICLE 4 coïncident avec notre hypothèse affirmant que l'échographie thoracique était suffisante pour diagnostiquer la BPI chez les veaux de race laitière. Détaillons maintenant l'amélioration de la standardisation de ce test afin d'améliorer son utilisation sur le terrain.

Chapitre 4 – Amélioration de la technique de l'échographie thoracique pour détecter la bronchopneumonie infectieuse chez les veaux de race laitière

ARTICLE 5

Bayesian estimation of sensitivity and specificity of systematic thoracic ultrasound exam for diagnosis of bovine respiratory disease in pre-weaned calves

Publié dans le *Preventive Veterinary Medicine* 2019 Jan 1;162:38-45.

Berman J^a, Francoz D^a, Dufour S^b, Buczinski S^a

^aFrom the Département des sciences cliniques, Faculté de médecine vétérinaire, Université de Montréal, Saint-Hyacinthe, QC, Canada, J2S 2M2

^bDépartement de pathologie et microbiologie, Faculté de médecine vétérinaire, Université de Montréal, Saint-Hyacinthe, QC, Canada, J2S 2M2

Abbreviation: AUS = Thoracic Auscultation; BCI = Bayesian Credible Intervals; BRD = Bovine Respiratory Disease; Covp = covariance of TUS and CRSC in calves with active pneumonia; Covn = covariance of TUS and CRSC in calves without active pneumonia; CRSC = Calf Respiratory Scoring Chart; DEPTH = maximal lung consolidation depth on ultrasound; DIC = Deviance Information Criterion; Hap = Haptoglobin; LCM = Latent class model; LR₊ = Positive Likelihood Ratio; LR₋ = Negative Likelihood Ratio; NPV = Negative Predictive Value; PPV = Positive Predictive Value; Se = Sensitivity; Se_{TUS} = Sensitivity of Thoracic ultrasonography; Se_{CRSC} = Sensitivity of Clinical Respiratory Scoring Chart; Se_{HAP} = Sensitivity of haptoglobin; Sp = Specificity; Sp_{TUS} = Specificity of thoracic ultrasonography; Sp_{CRSC} = Specificity of clinical respiratory scoring chart; Sp_{HAP} = Specificity of haptoglobin; TUS = Thoracic ultrasonography.

ABSTRACT

Objectives: Among the different clinical presentations of bovine respiratory disease, active pneumonia, defined as an infection of the lower airway with signs of inflammation, is the most important to diagnose correctly so appropriate treatment can be initiated. Diagnostic tests that accurately identify cases of active pneumonia are lacking; however, thoracic ultrasonography (**TUS**) seems promising. The primary objective of this study was to estimate the accuracy of TUS compared to reference tests for the diagnosis of active pneumonia in pre-weaned calves, using a latent-class model method (**LCM**).

Methods: The tests used for comparison were the Wisconsin Clinical Respiratory Scoring Chart (**CRSC**, positive if ≥ 5) and serum haptoglobin concentration (**Hap**, positive if $\geq 15\text{mg/dL}$). Secondary objectives were to assess the incremental value on TUS accuracy of combining TUS of the right cranial part and caudal parts, and to determine the accuracy of various thresholds for depth of consolidation ($>0\text{ cm}$, $\geq 1\text{cm}$, or $\geq 3\text{cm}$) for diagnosis of active pneumonia. One population of veal calves ($n=209$) and one of dairy calves ($n=301$) were enrolled. TUS, CRSC and Hap were all performed on each calf on the same day. TUS was performed by screening the mid to ventral portion of the lung caudal of the heart (the caudal sites), as well as the right parenchyma cranial to the heart (the cranial site). The maximal depth of consolidation (**DEPTH**) on TUS was recorded and noted separately for caudal and cranial sites. Different TUS cases were defined according to site and DEPTH. The accuracy of TUS was estimated by LCM for three tests conducted in two populations.

Results: Prevalence of active pneumonia was low (0.05) in both populations. In general, higher minimal consolidation depth thresholds led to increased TUS specificity (**Sp**) estimates, with minimal effects on TUS sensitivity (**Se**). With a TUS DEPTH threshold of $\geq 3\text{cm}$, adding TUS of the cranial site had little effect on accuracy. Using the $\geq 3\text{cm}$ threshold with caudal sites only, posterior Se and Sp median estimates of 0.89 (95%BCI: 0.55, 1.0) and 0.95 (95%BCI: 0.92, 0.98), respectively, were obtained.

Conclusion: In conclusion, in populations with low active pneumonia prevalence, adding TUS of the cranial site did not enhance the performance of the test. We suggest using a DEPTH threshold of $\geq 3\text{cm}$ solely on caudal sites to detect active pneumonia.

Key words: Active pneumonia; ultrasonography; dairy calf; veal calf; Bayesian analysis.

I- INTRODUCTION

Bovine respiratory disease (**BRD**) is a multifactorial disease with both short-term and long-term economic impacts on bovine production (Pardon et al., 2012b, Teixeira et al., 2017b). In veal and dairy operations, the peak of clinical incidence of BRD occurs during the first two months of life (Svensson et al., 2007, Pardon et al., 2012b). Accurate diagnosis of BRD during this time is therefore necessary for two reasons: first, to establish an appropriate treatment plan and avoid the subsequent negative outcomes caused by not treating affected animals; and second, to avoid unjustified and inappropriate use of antimicrobials due to false-positive test results. However, a distinction between active pneumonia (i.e., infection (by virus or bacteria) of the lower airway with signs of inflammation) and non-active pneumonia (i.e., upper airway BRD, scarring, fibrosis of the lung) is essential to establishing an appropriate treatment plan.

Different strategies have been implemented to assess active pneumonia in pre-weaned calves. For producers, diagnosis of active pneumonia is typically based on visual signs of respiratory disease (e.g., anorexia, depression, nasal and ocular discharge, cough, and ear position), associated or not with an increased rectal temperature. In recent years, a systematic scoring method has been developed (McGuirk, 2008). The calf respiratory scoring chart (**CRSC**) is now widely used in North America and is a method that can be easily applied by producers to screen calves for active pneumonia (Buczinski et al., 2014, Love et al., 2014, Buczinski et al., 2015a). However, this diagnostic approach is only moderately accurate for identifying sick (sensitivity [**Se**] = 0.62; 95% Bayesian credible intervals [**BCI**]: 0.48, 0.76) and healthy animals (specificity [**Sp**] = 0.74; 95% BCI: 0.65, 0.83) (Buczinski et al., 2015a). For veterinarians, diagnosis of active pneumonia is typically based on the presence of visual signs of respiratory disease and abnormal lung sounds, such as increased bronchial sounds, crackles, wheezes, or the absence of respiratory sounds, on thoracic auscultation (**AUS**). But the accuracy of AUS, like that of the CRSC, is questionable, with a moderate Se (Se = 0.73; 95% BCI: 0.50, 0.96), and quite poor Sp (Sp = 0.53; 95% BCI: 0.43, 0.64) (Buczinski et al., 2016b). Thus, these two tests are imperfect, and supplementary tools are necessary to establish an accurate diagnosis of active pneumonia.

In the past few years, thoracic ultrasonography (TUS) has been developed and used to diagnose BRD in calves (Ollivett and Buczinski, 2016). Among the lesions that can be detected with TUS, depth of lung consolidation was one of the abnormalities most closely associated with a clinical outcome (e.g., euthanasia or death) in a case-control study involving feedlot calves, and was, above all, the easiest parameter to measure (Rademacher et al., 2014). Moreover, inter-observer agreement for identifying calves with this type of lesion was substantial (Buczinski et al., 2013). Ultrasonographical evidence of lung consolidation is apparent when inflammatory and infectious processes affect the lung parenchyma, resulting in a non-aerated lung (Babkine and Blond, 2009, Ollivett and Buczinski, 2016). When pulmonary consolidation is present, lung tissue appears hypoechoic, and its echo texture may resemble liver parenchyma. This differs from normal lung tissue, which is anechoic due to its gas content (Babkine and Blond, 2009, Ollivett and Buczinski, 2016). However, a consolidated lung is not systematically associated with active lung infection. It can also be found in cases of lung infarction, atelectasis, fibrosis or scarring (Sartori and Tombesi, 2010). There is no standardization among studies concerning interpretation of lung consolidation images on TUS. In fact, the thresholds for depth of lung consolidation that are associated with the presence of significant lesion vary between 0 cm (Teixeira et al., 2017b), to ≥ 1 cm (Buczinski et al., 2014, Buczinski et al., 2015a), or ≥ 3 cm (Ollivett, 2014). Different TUS techniques have been reported in the literature; scanning of the 3rd, 5th and 7th intercostal spaces on the right hemithorax (Abutarbush et al., 2012); scanning the bilateral hemithorax (caudal part of the lung) from the 11th to 3rd intercostal spaces (Babkine and Blond, 2009); or combining TUS techniques and adding examination of the right cranial part of the lung by scanning the 2nd and 1st intercostal spaces (Ollivett et al., 2015). Bilateral scanning of the hemithorax was proven useful (Buczinski et al., 2014), but the hypothetical added value of assessing the cranial part of the lung compared to the extra time needed for that measurement was not investigated.

Previously, only two studies described the accuracy of TUS by using necropsy as the gold standard (Rabeling et al., 1998, Ollivett et al., 2015). Chronic cases, with long-standing pneumonia (Rabeling et al., 1998), or subclinical cases with lung consolidation but no clinical signs (Ollivett et al., 2015), were enrolled. Whether pneumonia was active (i.e.,

there was infection of the lower airway with signs of inflammation) was not, however, considered. Thus, to date, no study has described the accuracy of TUS for diagnosing active pneumonia, compared to simply diagnosing lung lesions in general. This latter aspect is of importance, since, when using TUS to determine whether to treat with antimicrobials and or anti-inflammatories, one would certainly want to disregard the chronic, non-active lesions that can be found in calves that have recovered from a previous episode of respiratory disease.

The concentration of haptoglobin (**Hap**), an acute phase protein, was studied to determine its relationship to clinical cases of BRD (Abdallah et al., 2016). A cohort study in dairy calves reported that, with a threshold of ≥ 15 mg/dl, the Se and Sp of Hap to detect clinical BRD were 0.64 (95% CI: 0.50, 0.77) and 0.70 (95% CI: 0.60, 0.80), respectively (Svensson et al., 2007).

Therefore, although CRSC, TUS, and Hap can all be helpful for diagnosing active pneumonia in calves, none of these tests is a gold standard for that purpose. An alternative approach for assessing the accuracy of TUS for diagnosing active pneumonia is to use Bayesian latent class models (**LCM**). In these models, none of the methods are considered gold-standard, and the disease status (active pneumonia in our case) is designated as latent: existing, but not apparent nor detected (Branscum et al., 2005).

The main objective of this study was to estimate the accuracy of TUS (our test under investigation) compared to CRSC and Hap (our reference tests) for the diagnosis of active pneumonia in pre-weaned dairy and veal calves, using LCM. One of our secondary objectives was to assess the incremental effect on TUS accuracy of combining TUS of the right cranial part with the caudal part. Our hypotheses were that TUS would be an excellent test to detect active pneumonia, and that combining TUS of the right cranial part with the caudal part would increase TUS Se. Another secondary objective was to determine the accuracy of various depth of consolidation thresholds (≥ 0 cm, ≥ 1 cm, or ≥ 3 cm) for diagnosis of active pneumonia.

II- MATERIALS AND METHODS

The STARD-BLCM guidelines (Standards for Reporting of Diagnostic accuracy studies that use Bayesian Latent Class Models) were used to facilitate reporting on the design, conduct, and results of the current study (Kostoulas et al., 2017).

II-1- Study design

Data from two different prospective cross-sectional one-gate designs conducted in two different populations were used. Briefly, a “one-gate” design is one in which disease positive and negative cases are selected from the same population following the same diagnostic testing in both groups. It differs from a “two-gate” design, in which disease positive and negative cases are diagnosed using different inclusion criteria (Rutjes et al., 2005). One trial was conducted between July and October 2014 on veal calves. This specific study was a randomized clinical trial to assess the efficacy of tildipirosin (Zuprevo; Merck Animal Health, Kirkland, Québec, Canada) as a antimicrobial metaphylactic treatment for the prevention of BRD in veal calves (Berman et al., 2017). Another cross-sectional study was conducted between July and October 2015 on dairy calves. This study was performed to assess the prevalence of ultrasonographic lung lesions in herds visited by the ambulatory clinic of the Faculté de Médecine Vétérinaire, Université de Montréal (Buczinski et al., 2018a). For both studies, the experimental protocol was approved by the Comité d'éthique et d'utilisation des animaux de l'Université de Montréal (protocol: 14-Rech_1727, 15-Rech_1773).

II-2- Participants

II-2-1- Population of veal calves

This study (Berman et al., 2017) was performed in one commercial pre-weaned veal calf fattening unit in Québec (DéliMAX, Veaux Lourds Ltée, Saint-Hyacinthe, Québec, Canada). Calves (n=209) from one batch, coming from local auction markets, were enrolled upon arrival at the farm at the approximate age of seven days of life (exact age was not reported for each calf). On arrival at the farm (day 0), calves were immediately vaccinated, intranasally, against infectious bovine rhinotracheitis and para-influenza 3 virus (Nasalgen® IP, Merck Animal Health, Kirkland, Québec, Canada). We used the data sampled on day

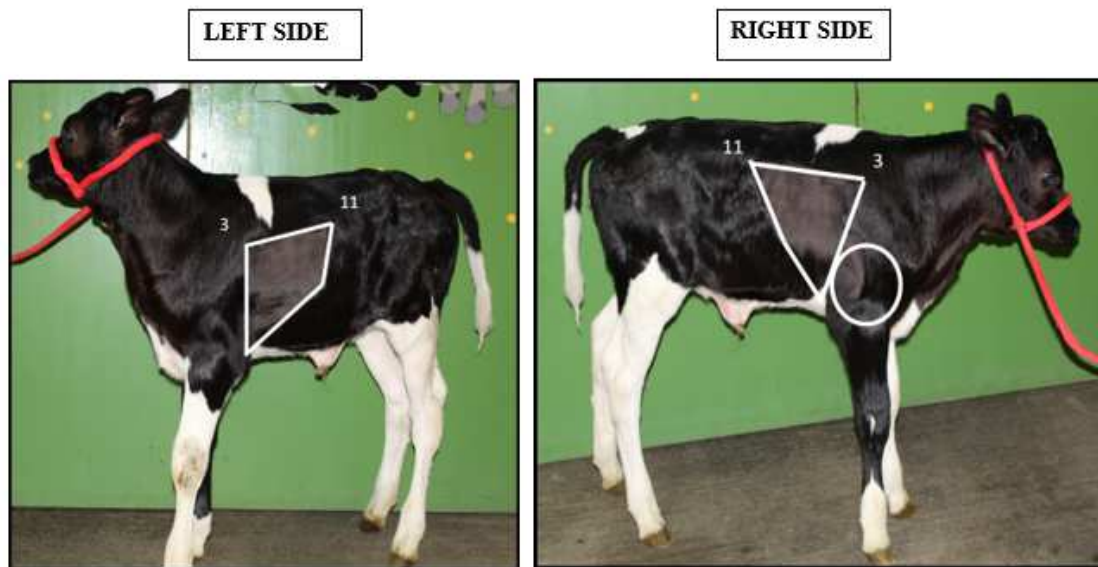
12 after arrival, before the administration of the antimicrobial treatment. At that time, veal calves were approximately 3 weeks old.

II-2-2- Population of dairy calves

The other study was performed with 39 herds, owned by clients of the ambulatory clinic of the Faculté de Médecine Vétérinaire, Université de Montréal (Buczinski et al., 2018a). The herds enrolled were selected by convenience. Among each herd, ≥ 6 and up to 12 pre-weaned heifers (median age of 38 days; interquartile range of 28–50 days) were enrolled. Data were collected on 301 dairy calves in the summer. These calves were part of another, larger study aimed at validating the Californian Clinical scoring system compared to TUS as a reference test (Buczinski et al., 2018c).

II-3- Test methods

Bilateral TUS, the test under investigation, was performed with a 7.5 MHz linear probe (Imago, ECM, Angoulême, France for veal calves and DRAMINSKI iScan, Olsztyn, Poland for dairy calves) by two experienced operators (JB and SB for veal calves and MEB and SB for dairy calves). The operators were blinded to CRSC results when conducting TUS. 70% isopropyl alcohol was sprayed on the area of interest to improve image quality, and the probe was directly applied on the thorax. The following areas were scanned: the mid to ventral portion of the right and left lungs caudal to the heart (the caudal sites) and the parenchyma cranial to the heart on the right side (the cranial site) (Ollivett and Buczinski, 2016). The maximal depth of consolidation on TUS (**DEPTH**; in cm) was recorded. DEPTH was calculated by manual count, using the 1 cm² grid of the ultrasound unit, and recorded separately for the caudal and cranial parts of the lung (caudal *vs.* cranial). Several TUS-based case definitions were investigated, considering a TUS test as positive as function of the maximal DEPTH observed (≥ 0 cm (Teixeira et al., 2017b); ≥ 1 cm (Buczinski et al., 2014, Buczinski et al., 2015a) and ≥ 3 cm (Ollivett, 2014, Buczinski et al., 2018a) and considering the examination of bilateral caudal sites only, *vs.* the right cranial site only *vs.* a combination of bilateral caudal sites and the right cranial site. Cranial and caudal TUS sites are described in **Figure 14**. The different TUS-based case definitions investigated are described in **Table 43**.



The caudal sites (white area) are represented from the 11th to 3rd intercostal spaces, on the left and right side.
 The Cranial site (white circle) is represented from the 2nd to 1st intercostal spaces, on the right side

Figure 14. – Illustration of thoracic ultrasonography (TUS) sites using the TUS-based case definitions (bilateral caudal sites only, vs. right cranial site only vs. a combination of bilateral caudal sites and right cranial site).

As previously described, a calf respiratory score (our first reference test) was performed on each calf in both populations, using CRSC (McGuirk, 2008). The test was done on the same day as TUS. A CRSC of five or more was used to define a BRD clinical case (McGuirk, 2008). Each calf was scored by one of two operators (CF or FJ for the veal calf population, and MEB or SB for the dairy calf population), who were blinded to TUS results. To decrease inter-observers' variability, all operators in both studies were trained by the same person (SB).

Blood samples were also collected for each calf and were centrifuged at 3500 rpm for 10 min at approximately 20°C within 24 h of collection. Serum was separated, and aliquots were stored at – 20°C and subsequently analyzed in duplicate for Hap concentration determination using a commercially available ELISA kit (Genway Biotech, San Diego, CA), already validated in cattle. These analyses were performed immediately after the TUS and CRSC data collection. Briefly, Hap present in samples reacts with anti-Hap antibodies

adsorbed to the surface of polystyrene micro-titre wells. After the removal of unbound proteins by washing, anti-Hap antibodies were conjugated with horseradish peroxidase. Following another wash, the enzyme binds to a chromogenic substrate, 3, 3', 5, 5'-tetramethylbenzidine. Peroxidase activity was measured by spectrophotometry. Results were calculated using the ELISA standard curve, reading the absorption at 450 nm and transferring the values into mg/ml, with a working range of 15.6 to 1000 ng/ml and a sensitivity of 6.258 ng/ml (Genway, 2017). Based on a cohort study of pre-weaned dairy calves (Svensson et al., 2007), an active inflammatory process was deemed present when the concentration of Hap was ≥ 15 mg/dl, and we used this cut-off to define active pneumonia using the Hap test (negative if < 15). This Hap test was used as a second reference test and could be hypothesized to be more specific for diagnosis of active lung lesions (i.e., it measures the presence of acute inflammation). All three tests (TUS, CRCS, and Hap) were performed on the same day, on each calf.

II-4- Statistical analysis

The calf was the unit of interest in this study. Descriptive results for each diagnostic test (TUS, CRSC, and Hap) were computed. Two-by-two tables were prepared to compare tests, and these tables were used to inform the LCM.

II-4-1- Main model (initial LCM)

In this study, our latent variable included two statuses: active pneumonia (defined as the presence of lung lesions with active inflammation); or non-active pneumonia (defined as the presence of lung lesions without inflammation, chronic cases, or a complete absence of lesions). An LCM for comparing three tests (TUS, CRSC, Hap) in two populations (veal and dairy calves) was used.

For the initial LCM, we made two important assumptions: 1) the three diagnostic tests were conditionally independent; and 2) The Se and Sp of the three diagnostic tests were constant across the two populations. Under these assumptions, 8 unknown parameters had to be estimated: Se and the Sp of TUS (\mathbf{Se}_{TUS} ; \mathbf{Sp}_{TUS}), Se and Sp of CRSC ($\mathbf{Se}_{\text{CRSC}}$; $\mathbf{Sp}_{\text{CRSC}}$), Se and Sp of Hap (\mathbf{Se}_{HAP} ; \mathbf{Sp}_{HAP}), prevalence of active pneumonia in veal calves, and prevalence of active pneumonia in dairy calves. There were 14 total degrees of freedom, and the model was therefore identifiable (Dendukuri et al., 2009).

II-4-2- Priors

Beta distributions were used as priors for all parameters of interest. A vague prior distribution ($\beta(1;1)$) was chosen for Se and Sp of TUS and CRSC. Informative priors were used for prevalence of active pneumonia in veal and dairy calves. The most probable value of active pneumonia prevalence and the range containing the true value with 95% confidence were obtained from the literature on veal and dairy calves (Svensson et al., 2007, Pardon et al., 2012b). Given the inaccuracy in definitions of this BRD status, low informative priors were used, with a mode at 0.05 and 95% certainty that prevalence was ≤ 0.25 , corresponding to a $\beta(1.7; 14.5)$ distribution. Considering the ability of the Hap test to identify active inflammatory processes, we used informative prior distributions for Hap test accuracy parameters based on available literature (Svensson et al., 2007). The most probable value for Hap Se was determined to be 0.65, while we were 95% sure that it was ≥ 0.50 , corresponding to a $\beta(21.0; 11.8)$ distribution. For Hap Sp, a mode of 0.70 was hypothesized, with 95% certainty that $Sp \geq 0.60$, corresponding to a $\beta(47.5; 20.9)$ distribution.

II-4-3- Sensitivity analysis - Assumption of conditional independence

The Hap test was assumed to be conditionally independent of the two other tests, since it is based on different biological principles than TUS and CRSC (dosage of an acute-phase protein). It is not clear, however, whether TUS and CRSC can be considered conditionally independent. To investigate the impact of that assumption, we also considered an LCM allowing for conditional TUS and CRSC dependency. Covariance parameters were added in this later model, and their distribution defined as follows, based on Branscum and al. (2005):

$$\text{Covp} \sim \text{dunif}((\text{Se}_{\text{CRSC}} - 1) * (1 - \text{Se}_{\text{TUS}}), \min(\text{Se}_{\text{CRSC}}, \text{Se}_{\text{TUS}}) - \text{Se}_{\text{CRSC}} * \text{Se}_{\text{TUS}})$$

$$\text{Covn} \sim \text{dunif}((\text{Sp}_{\text{CRSC}} - 1) * (1 - \text{Sp}_{\text{TUS}}), \min(\text{Sp}_{\text{CRSC}}, \text{Sp}_{\text{TUS}}) - \text{Sp}_{\text{CRSC}} * \text{Sp}_{\text{TUS}})$$

Where **Covp** (covariance of TUS and CRSC in calves with active pneumonia) and **Covn** (covariance of TUS and CRSC in calves without active pneumonia) are assumed to follow uniform (U) distributions, centered on 0 and with variation determined, for Covp, by the minimal Se between TUS and CRSC minus the product of TUS and CRSC sensitivities,

and, for Covn, by the minimal Sp between TUS and CRSC minus the product of TUS and CRSC specificities.

II-4-4- Sensitivity analysis - Assumption of constant Se and Sp across populations

In both populations, calves were of dairy breeds. However, calves in both populations were of different ages and exposed to different environments. Moreover, both TUS and CRSC measurements in the dairy and veal populations were conducted by different sets of observers. To investigate the impact of the assumption of constant test accuracy across populations, LCMs allowing for different TUS and different CRSC Se and Sp in veal and dairy calves were evaluated.

II-4-5- Evaluation of the test's practical application

Positive predictive value (**PPV**) is the probability that subjects with a positive screening test truly have the disease. Negative predictive value (**NPV**) is the probability that subjects with a negative screening test are truly healthy. Positive PPV and NPV of TUS were computed using the LCM models, as described by Dohoo et al. (2012), for a range of disease prevalence: 2.5%, 5%, 7.5%, 10%, 12.5%, 15%, 17.5% and 20%.

Positive (**LR+**) and negative (**LR-**) likelihood ratios for each TUS case-definition were computed as described by (Dohoo et al., 2012)), in order to illustrate how a TUS test result usefully changes the probability that active pneumonia exists.

II-4-6- Assessment of convergence and statistical software

Bayesian computations were implemented using a public domain software (WinBUGS, version 1.4.3, MRC biostatistics Unit, Cambridge, UK). Three chains were run, with different initial values. The first 5,000 iterations of each chain were discarded as burn-in, whereas the next 95,000 were used to obtain posterior distributions. Convergence of models was assessed by visual inspection of the history, density plots, and Brooks-Gelman-Rubin statistics. Autocorrelation plots were also checked to ensure proper behavior of the Markov chains. Posterior distributions of each parameter were reported as medians and corresponding 95% BCI. We used the DIC tool dialog box to estimate the deviance information criterion (**DIC**) for the different models. The model with the smallest DIC was

the model that would best predict a replicate dataset with the same structure as that currently observed (Spiegelhalter et al., 2002). LR+, LR-, PPV and NPV were computed directly with WinBUGS, using the LCM models, The results of PPV and NPV were illustrated with a graphic created using the Excel software program (Windows, Microsoft, Redmond WA).

III- RESULTS

III-1- Descriptive Statistics

Numbers of animals testing positive on the different tests in each population are presented in **Table 43**. The proportion of animals positive to all tests was low, but a large proportion of calves were negative on all three tests. The proportion of animals positive to TUS was greater than that of the two other tests (i.e., CRSC and Hap).

TUS threshold	Pop	TUS+, CRSC+, Hap+	TUS+, CRSC+, Hap-	TUS-, CRSC+, Hap+	TUS-, CRSC+, Hap-	TUS+, CRSC-, Hap+	TUS+, CRSC-, Hap-	TUS-, CRSC-, Hap+	TUS-, CRSC-, Hap-
Caudal >0cm	Veal	3	3	1	0	3	28	9	162
	Dairy	1	16	0	17	4	43	19	201
Cranial >0cm	Veal	3	2	1	1	4	21	8	169
	Dairy	1	20	0	13	6	71	17	173
Caudal or cranial >0cm	Veal	4	3	0	0	7	54	5	136
	Dairy	1	21	0	12	8	90	15	154
Caudal ≥1cm	Veal	3	3	1	0	3	21	9	169
	Dairy	1	15	0	18	3	36	20	208
Cranial ≥1cm	Veal	3	2	1	1	3	17	9	173
	Dairy	1	17	0	16	4	43	19	201
Caudal or cranial ≥1cm	Veal	4	3	0	0	4	39	8	151
	Dairy	1	18	0	15	5	60	18	184
Caudal ≥3cm	Veal	3	3	1	0	3	8	9	182
	Dairy	1	9	0	24	2	14	21	230
Cranial ≥3cm	Veal	2	1	2	2	2	10	10	180
	Dairy	1	12	0	21	3	15	20	229
Caudal or cranial ≥3cm	Veal	3	3	1	0	4	16	8	174
	Dairy	1	16	0	17	3	27	20	217

Cranial: right cranial part of the lung on ultrasound; Caudal: left and right caudal part of the lung on ultrasound

Tableau 43. – Number of calves with a given tests results profile using three tests (thoracic ultrasonography [TUS], concentration of haptoglobin [Hap], calf respiratory score [CRSC]) on the two populations (veal and dairy calves) as function of TUS case definition.

III-2- Comparison between TUS case definition

Results obtained from the initial Bayesian LCM using the different TUS sites and depth are shown in **Table 44**. All models converged to their posterior distributions. The TUS Se was relatively similar across the different TUS case-definitions used (there was a variation of less than 10% between posterior medians, and largely overlapping BCI, as shown in **Table 44**). In general, requesting higher minimal consolidation depth thresholds led to increased TUS Sp estimates, with minimal effects on TUS Se. With TUS consolidation depth thresholds of $>0\text{cm}$ or $\geq 1\text{cm}$, adding TUS of the right cranial site increased TUS Se, but with an important decrease in Sp. With a TUS consolidation depth threshold $\geq 3\text{cm}$, adding TUS of the cranial site had little effect on TUS Se and Sp estimates. The scenario with the highest median Sp to detect active pneumonia was when TUS was considered positive if lung consolidation $\geq 3\text{ cm}$ was observed caudal to the heart (without assessing the cranial site). Using this threshold, posterior Se and Sp median estimates of 0.89 (95% BCI: 0.55, 1.0) and 0.95 (95% BCI: 0.92, 0.98), respectively, were observed.

Likelihood ratios (LR₊ and LR₋) of TUS are presented in **Table 44**. LR₋ was low and relatively similar across the different TUS case-definitions. Adding TUS of the cranial site decreased the LR₊, regardless of depth of consolidation. The scenario with the highest LR₊ was when TUS was considered positive if lung consolidation $\geq 3\text{ cm}$ was observed caudal to the heart (without assessing the cranial site). Using this threshold, LR₋ and LR₊ of TUS were 0.12 (95% BIC: 0.01, 0.47) and 17.7 (95% BIC: 9.37, 40.5), respectively. In a population with a 5% prevalence of active pneumonia (i.e., in a population with a pre-test odd of 0.053), a positive TUS test would be associated with a 48% probability of truly being diseased (i.e., a post-test odd of 0.93). Similarly, a negative TUS test would be associated with a 99% probability of truly being non-diseased.

TUS	Sensitivity		Specificity		LR ₊		LR ₋		
	Median	BCI	Median	BCI	Median	BCI	Median	BCI	
DEPTH ≥ 0 cm									
Cranial	0.86	0.60, 0.99	0.81	0.76, 0.89	4.44	2.86, 7.68	0.17	0.01, 0.50	
Caudal	0.85	0.53, 0.99	0.84	0.80, 0.88	5.17	3.01, 7.22	0.17	0.01, 0.56	
Cranial/Caudal	0.90	0.65, 1.00	0.66	0.62, 0.71	2.67	1.89, 3.27	0.14	0.01, 0.52	
DEPTH ≥ 1 cm									
Cranial	0.85	0.54, 0.99	0.90	0.87, 0.93	6.32	3.80, 9.73	0.17	0.01, 0.53	
Caudal	0.85	0.52, 0.99	0.87	0.83, 0.91	6.46	3.76, 9.46	0.17	0.01, 0.56	
Cranial/Caudal	0.90	0.61, 1.00	0.77	0.73, 0.81	3.89	2.59, 5.01	0.13	0.01, 0.51	
DEPTH ≥ 3cm									
Cranial	0.81	0.45, 0.99	0.94	0.92, 0.97	14.4	6.98, 29.5	0.19	0.01, 0.59	
Caudal	0.89	0.55, 1.00	0.95	0.92, 0.98	17.7	9.37, 40.5	0.12	0.01, 0.47	
Cranial/Caudal	0.86	0.53, 0.99	0.90	0.87, 0.94	8.66	4.98, 14.4	0.15	0.01, 0.52	

DEPTH: maximal depth of consolidation; Cranial: right cranial part of the lung on ultrasound; Caudal: left and right caudal part of the lung on ultrasound; BCI: Bayesian Credible Interval.

Tableau 44. – Estimates of sensitivity (Se), specificity (Sp), positive (LR₊) and negative (LR₋) likelihood ratio of systematic thoracic ultrasonography (TUS) using Bayesian latent-class models of the main model for the different TUS case definitions in dairy and veal calves during the pre-weaning period.

III-3- Sensitivity analysis

Sensitivity analyses were conducted solely for a DEPTH ≥ 3 cm observed caudal to the heart. The prior distributions used, the posterior distributions observed in the LCM allowing different TUS or CRSC Se and Sp across populations, and the LCM allowing conditional dependency between TUS and CRSC, are summarized in **Table 45**. Adding conditional dependency terms had little effect on TUS Se and Sp estimates. The models assuming different TUS or CRSC Se and Sp estimates between populations (i.e., between veal and dairy calves) yielded results qualitatively similar to that of the LCM assuming common accuracy measures between populations (**Table 45**). The model with the lowest DIC was the LCM allowing for different CRSC accuracy measures between populations (DIC: 94.1).

Parameters	Initial LCM		LCM allowing tests dependency		LCM allowing different TUS Se and Sp across populations		LCM allowing different CRSC Se and Sp across populations	
	Priors	Posteriors	Priors	Posteriors	Priors	Posteriors	Priors	Posteriors
TUS								
Se	β (1;1)	0.89 (0.55, 1.00)	β (1;1)	0.82 (0.35, 0.99)	Dairy β (1;1) Veal β (1;1)	Dairy 0.66 (0.24, 0.98) Veal 0.87 (0.53, 0.99)	β (1;1)	0.88 (0.62; 0.99)
Sp	β (1;1)	0.95 (0.92, 0.98)	β (1;1)	0.93 (0.90, 0.96)	Dairy β (1;1) Veal β (1;1)	Dairy 0.94 (0.91, 0.98) Veal 0.96 (0.91, 0.99)	β (1;1)	0.95 (0.92; 0.97)
CRSC								
Se	β (1;1)	0.69 (0.40, 0.97)	β (1;1)	0.48 (0.12, 0.91)	β (1;1)	0.70 (0.40, 0.97)	Dairy β (1;1) Veal β (1;1)	0.70 (0.30; 0.98) 0.66 (0.31; 0.97)
Sp	β (1;1)	0.95 (0.92, 0.97)	β (1;1)	0.93 (0.90, 0.95)	β (1;1)	0.95 (0.92, 1.00)	Dairy β (1;1) Veal β (1;1)	0.91 (0.87; 0.94) 0.99 (0.98; 1.0)
Hap								
Se	β (21.0;11.8)	0.51 (0.36, 0.67)	β (21.0;11.8)	0.62 (0.43, 0.78)	β (21.0;11.8)	0.47 (0.30, 0.65)	β (21.0;11.8)	0.54 (0.39; 0.69)
Sp	β (47.5;20.9)	0.90 (0.88, 0.93)	β (47.5;20.9)	0.90 (0.88, 0.93)	β (47.5;20.9)	0.90 (0.88, 0.93)	β (47.5;20.9)	0.90 (0.88; 0.93)
Dairy prev	β (1.7;14.5)	0.05 (0.02, 0.10)	β (1.7;14.5)	0.03 (0.01, 0.08)	β (1.7;14.5)	0.05 (0.02, 0.09)	β (1.7;14.5)	0.04 (0.01; 0.08)
Veal prev	β (1.7;14.5)	0.05 (0.02, 0.09)	β (1.7;14.5)	0.02 (0.002, 0.06)	β (1.7;14.5)	0.06 (0.02, 0.17)	β (1.7;14.5)	0.05 (0.02; 0.09)
Cov Se				0.004 (-0.11, 0.14)				
Cov Sp				0.018 (0.004, 0.03)				
DIC		98.4		94.1		97.5		78.2

Dairy Prev: prevalence in the population of dairy calves; Veal Prev: prevalence in the population of veal calves

Cov Se: Covariance between the sensitivity of CRSC and TUS among active pneumonia positive calves

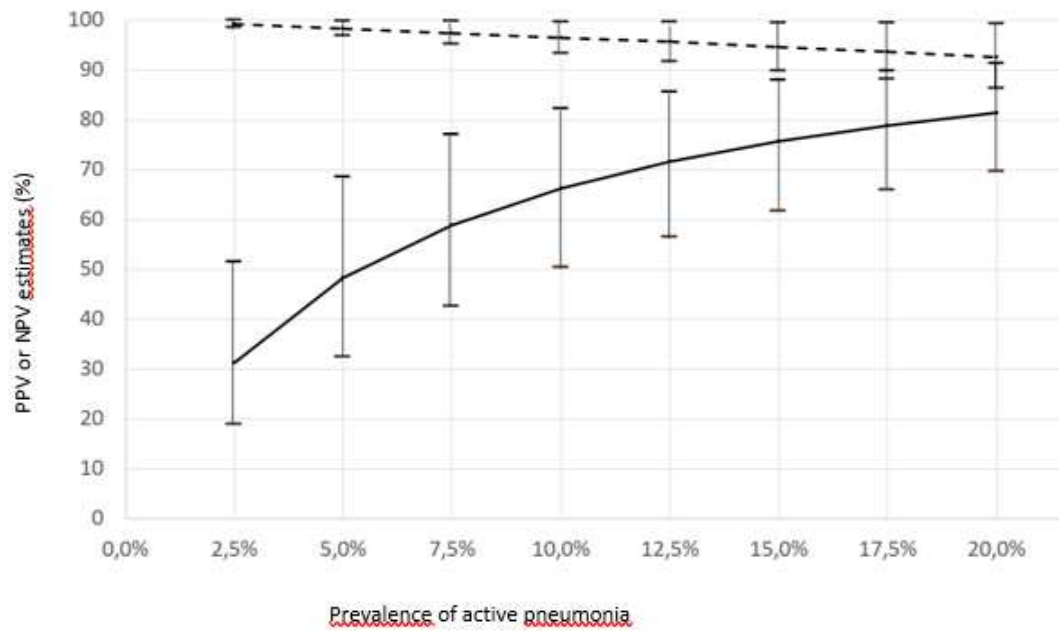
Cov Sp: Covariance between the specificity of CRSC and TUS among active pneumonia negative calves

DIC: Deviance Information Criterion

Tableau 45. – Posterior median and 95% credible intervals based on four plausible Bayesian latent class models (LCM) for the prevalence of active pneumonia in two different populations, and sensitivity (Se) and specificity (Sp) of the clinical respiratory score chart (CRSC), thoracic ultrasound (TUS) (with the TUS case definition of caudal ≥ 3 cm) and dosage of haptoglobin (Hap) used to diagnose active pneumonia in dairy and veal calves during the pre-weaning period.

III-4- Illustration of PPV and NPV of TUS

The PPV and NPV for the TUS case-definition requiring that lung consolidation ≥ 3 cm be observed caudal to the heart across disease prevalence are presented in **Figure 15**. For a variation of BRD prevalence from 2.5% to 20%, the NPV only decreased from 0.99 (95% BCI: 0.98, 1.00) to 0.92 (95% BCI: 0.86, 0.99). Conversely, the PPV was multiplied by a factor of 2.6 over that same range of disease prevalence (from 0.31; 95% BCI: 0.19; 0.51 to 0.82; 95% BCI: 0.69; 0.91).



The positive predictive values (PPV, line) and negative predictive values (NPV, dotted line) with 95% credible intervals are presented.

Figure 15. – Positive predictive values (PPV) and negative predictive values (NPV) of thoracic ultrasonography (using a consolidation threshold of ≥ 3 cm caudal to the heart), according to different plausible prevalence values of active pneumonia in dairy and veal calves during the pre-weaning period.

IV- DISCUSSION

In these calf populations, Se_{TUS} was little affected by the DEPTH threshold chosen, but Sp_{TUS} was increased when considering a ≥ 3 cm DEPTH threshold. In the current study, however, adding TUS of the cranial site did not increase the Se_{TUS} , as expected, but had a negative impact on the Sp_{TUS} . Thus, evaluating caudal sites solely and using a ≥ 3 cm DEPTH threshold seems like the most practical approach.

Previously, several thresholds have been used to define a positive TUS test. Some studies used a consolidation depth ≥ 1 cm to avoid misclassification with pleural thickening, pleural irregularity, or both (Buczinski et al., 2014, Buczinski et al., 2015a). Another study used a depth ≥ 3 cm by selecting the value between the 90th and 95th percentiles (2.25 cm and 3.5 cm, respectively) of TUS lung consolidation in calves with any amount of consolidation (Ollivett, 2014). Another study used a consolidation depth threshold of > 0 cm (Teixeira et al., 2017b). To the authors' knowledge, this is the first study that compares these different thresholds and investigates threshold accuracy to detect active pneumonia. In this study, consolidation depth ≥ 3 cm seemed to be the most specific threshold to detect active pneumonia. However, given the low prevalence of sick animals in both populations, our study lacks power to estimate Se with accuracy. Thus, it is difficult to know that the increase in Sp with the depth of consolidation (≥ 3 cm) is not achieved at the expense of Se. With this limitation in mind, a threshold of ≥ 3 cm can be suggested from our results. Further studies, in populations with higher prevalence of active pneumonia, are needed to confirm our findings.

In recent years, many studies have described the thoracic ultrasonography technique (Babkine and Blond, 2009, Ollivett and Buczinski, 2016). Globally, different internal structures can be observed during TUS: the lung (its right and left caudal parts as well as its right cranial part), the heart, and the thymus (on the left cranial aspect of the thorax) (Ollivett and Buczinski, 2016). The literature presents various techniques that have been used for performing TUS, from scanning only the 3rd, 5th and 7th intercostal spaces on the right hemithorax (Abutarbush et al., 2012) to scanning the bilateral hemithorax (the caudal part of the lung) from the 11th to 3rd intercostal spaces (Babkine and Blond, 2009). Recently, also scanning the 1st and 2nd intercostal spaces was described as a means of better assessing

the cranial aspect of the right cranial lung lobe, which can be a site of pneumonia (Ollivett et al., 2015, Teixeira et al., 2017b). In a recent study conducted by Ollivett (Ollivett et al., 2015), 4 of 5 of the most severely affected subclinical calves would have been misclassified as normal without the scanning of the right cranial part, because the lesions were visible only from the right 1st and 2nd ICS. However, in contrast to our initial hypothesis, in this study, systematically adding TUS of the right cranial part of the lung did not lead to an increase in Se. Rather, it led to a decreased Sp to detect active pneumonia. However, as for the thresholds, this finding must be interpreted carefully, given that the proportion of sick animals was relatively low in both populations making estimations of Se unprecise, contrary to Sp. Thus, it is possible that a difference existed between Se estimates in the different scenarios but was not detected because of the low prevalence of active pneumonia in both populations. Moreover, the description of this site on TUS and possible increase in Se was performed on subclinical calves (Ollivett et al., 2015). Thus, another hypothesis would be that adding TUS of the right cranial site led to observation of lung consolidation that was not necessarily active (e.g., fibrosis, scarring).

In this study, using the threshold of ≥ 3 cm caudal to the heart, Se and Sp of TUS to detect active pneumonia were 0.89 (95% BCI: 0.55, 1.0) and 0.95 (95% BCI: 0.92, 0.98), respectively. Other studies reported Se and Sp of TUS in active pneumonia. One study reported a Se of 0.79 (95% BCI: 0.66, 0.91) and Sp of 0.93 (95% BCI: 0.87, 0.97) (Buczinski et al., 2015b); another reported Se of 0.77 (95% BCI: 0.60, 0.89) and Sp of 0.93 (95% BCI: 0.87, 0.97) (Buczinski et al., 2016b). However, these studies were conducted using a Bayesian framework using informative priors elicited from experts, and with a limited number of calves (Buczinski et al., 2015a, Buczinski et al., 2016b): $n = 209$ and $n = 191$, respectively. As a result (of having a limited “n”), choice of priors may have had an important impact on the observed posterior distributions, and comparisons with the present findings are difficult (Lambert et al., 2005).

In this study, the initial LCM was based on two main assumptions: (1) the diagnostic tests were considered conditionally independent of each other; and (2) Se and Sp of the diagnostic tests were constant between both populations. To evaluate the impact of these assumptions, three alternative models were developed and compared to the initial LCM

(**Table 45**). Adding test dependency had little effect on TUS estimates. However, the inferior limit for BCI Covn was positive (0.004), suggesting that TUS and CRSC were conditionally dependent when calves did not have active pneumonia (Branscum et al., 2005). Nonetheless, Covp, Covn median estimates and Covn positivity were relatively small. Adding test dependency did affect CRSC Se median estimates.

When considering that TUS or CRSC Se and Sp may be different for veal vs. dairy, in general we obtained similar results in terms of point estimates and 95% confidence intervals for models, allowing for population-specific accuracy measures compared to the model, and assuming common accuracy measures across populations. The only exception was TUS Se point estimate in dairy calves (0.66), which seemed relatively different than that of the common accuracy measures model (0.89; **Table 45**). However, the imprecision of Se estimation precludes drawing conclusions on that matter. Thus, the LCM allowing for different TUS Se and Sp across populations could also be biologically relevant.

In a population with a low disease prevalence, it is desirable to have a test with an excellent Sp to avoid false positives and, thus, avoid inappropriate and costly treatment. In a population with high disease prevalence, it is the opposite; it is more helpful to have a test with good Se to detect sick animals and reduce the proportion of false-negative calf that could have long-term economic impacts. In this study, PPV tripled (0.3 to 0.8) when active pneumonia prevalence varied between 2.5 and 20.0%, and NPV stayed constant when active pneumonia prevalence varied between 2.5 and 20.0%. Hence, TUS could be used as an add-on test to confirm a suspicion of active pneumonia.

Finally, it is important to be aware of the limits of this study. The first such limitation is the relatively small difference in active pneumonia prevalence between populations (**Table 45**). Usually, jointly modeling results from different populations increases the precision of Se and Sp estimates. However, since the difference in active pneumonia prevalence in the two populations was small, using a two-population LCM did little to improve model identifiability (Branscum et al., 2005). Therefore, no substantial increase in the precision of estimates occurred. The second limitation is the absence of information about aetiological agents. Indeed, it is possible that our findings would have been different if

different bacteria/viruses were involved. It would have been interesting to know, for the purposes of generalizing the results of our study, which aetiological agents were involved. This should be investigated in future studies.

V- CONCLUSION

In conclusion, in this study, the presence of lung consolidation caudal to the heart with ≥ 3 cm of depth yielded very good Se and Sp estimates for diagnosing active pneumonia in calves. In populations with low prevalence of active pneumonia, adding TUS of the cranial site did not increase the performance of the test.

CONCLUSION ARTICLE 5 : Dans cet ARTICLE 5, nous avons standardisé l'échographique thoracique en définissant les sites à échographier et le seuil de profondeur de consolidation pulmonaire avec la meilleure Se et Sp pour DÉTECTER LES BPI ACTIVES nécessitant un traitement d'antibiotiques et/ou d'anti-inflammatoires. Ainsi, la présence de lésions de consolidation pulmonaire dans les SITES CAUDAUX AU CŒUR avec un SEUIL DE PROFONDEUR ≥ 3 cm permet de détecter avec une Se de 0.89 (95%ICB : 0.55, 1.0) et Sp 0.95 (95%ICB : 0.92, 0.98) les BPI actives chez les veaux de race laitière. Ces performances coïncident avec notre hypothèse suggérant que l'échographie thoracique était facile et adéquate pour diagnostiquer les veaux de race laitière à traiter avec des antibiotiques et/ou anti-inflammatoire sur le terrain

Chapitre 5 – Discussion générale

I- Résumé des résultats et réponse à nos objectifs

L'objectif général de cette thèse était d'améliorer les stratégies de détection de la BPI des veaux de race laitière nécessitant un traitement afin de mieux diagnostiquer les veaux malades et, ultimement, diminuer les impacts économiques de la maladie, améliorer le bien-être des veaux et favoriser l'emploi judicieux des antibiotiques en traitant uniquement les veaux ou les groupes de veaux qui en ont besoin.

Pour répondre à cet objectif, nous avons (1) amélioré l'échographie thoracique et le SCRD pour détecter des veaux et des groupes de veaux à traiter et apporter une définition diagnostique du statut actif de la BPI; (2) déterminer les performances du SCRD, l'échographie thoracique et la radiographie thoracique pour détecter ce statut chez les veaux de race laitière.

I-1- Amélioration des tests pour détecter les veaux de race laitière à traiter et apporter une définition diagnostique du statut actif de bronchopneumonie infectieuse

I-1-1- De la définition conceptuelle à la définition opérationnelle de la bronchopneumonie infectieuse

Une définition diagnostique exacte et facile d'utilisation était absente pour détecter les **veaux à traiter** de BPI en médecine bovine (Buczinski et Pardon, 2020). Un des résultats majeurs de cette thèse a été d'améliorer le SCRD (à l'échelle populationnelle) et l'échographie thoracique (à l'échelle individuelle) pour détecter les veaux de race laitière à traiter et ainsi apporter une définition diagnostique du statut actif de la BPI basé sur les résultats de ces tests. Cette définition opérationnelle pourra être facilement appliquée en pratique.

Une des limites de cette définition est qu'elle manque de précision sur la nature du traitement à donner ne permettant pas de distinguer les veaux à traiter avec des antibiotiques seulement, des anti-inflammatoires seulement ou les deux. Actuellement, la

pratique de traiter les veaux atteints de BPI seulement avec des anti-inflammatoires est assez peu réalisée (Buczinski et Pardon, 2020). En effet, en absence d'évidence, beaucoup de médecins vétérinaires partent du principe que la surinfection bactérienne survient systématiquement après une infection virale ou une pneumonie inflammatoire, justifiant l'emploi systématique d'antibiotiques lors d'un diagnostic de BPI contrairement à d'autres espèces animales ou en médecine humaine. D'autre part, pour des raisons économiques, éthiques et la pression des propriétaires, des antibiotiques sont très souvent donnés afin de privilégier la sécurité plutôt que le risque de ne pas traiter une infection bactérienne (Buczinski et Pardon, 2020). Des progrès scientifiques sont donc à faire pour changer cette pratique bien ancrée. Cependant, à défaut d'être précise sur la nature du traitement, cette définition constitue la première étape pour guider les médecins vétérinaires et les producteurs vers la diminution de l'utilisation des antibiotiques en définissant avec exactitude les veaux qui en auraient besoin de ceux qui n'en auraient pas besoin.

Dans cette discussion général je parlerai donc de **veaux à traiter** sans préciser la nature du traitement et du **statut actif de la BPI**, signifiant le statut de BPI à traiter avec des antibiotiques et/ou des anti-inflammatoires.

I-1-2- Les tests sélectionnés

Comme mentionné dans le **CHAPITRE 1**, les performances actuelles du SCRD et de l'auscultation thoracique sont similaires (**Tableaux 15 et 16**). Par conséquent, aucune donnée ne montre que l'ajout de l'auscultation thoracique après le SCRD améliore le diagnostic. L'implantation seulement de l'un des deux tests est donc suffisante pour améliorer les stratégies diagnostiques actuelles de la BPI chez les veaux de race laitière. Par sa plus grande simplicité d'interprétation et son accessibilité facilitée aux producteurs, nous avons décidé de nous limiter à l'amélioration du SCRD plutôt que l'auscultation thoracique pour détecter le statut actif de la BPI dans cette thèse (**CHAPITRE 2**).

En ce qui concerne les tests d'imagerie médicale, aucune donnée dans la littérature actuelle ne nous permettait de savoir quel test parmi l'échographie ou la radiographie thoracique était le plus performant pour détecter la BPI. Une des étapes préalables a donc été de comparer ces deux tests d'imagerie médicale afin de déterminer la meilleure des stratégies d'utilisation de ces tests pour optimiser la détection de veaux à traiter de la BPI :

échographie thoracique seule, radiographie thoracique seule, ou utilisation d'une combinaison des deux (**CHAPITRE 3**). Nos travaux ont montré l'absence de différence entre les performances de l'échographie thoracique et celles de la radiographie thoracique pour détecter les lésions pulmonaires et le statut actif de la BPI. L'utilisation de l'un ou de l'autre test seul est donc suffisante pour détecter les veaux à traiter. L'échographie thoracique étant plus populaire que la radiographie thoracique ainsi que plus facile à utiliser sur le terrain, moins coûteuse et répétable sans nécessité d'une expertise (Ollivett et Buczinski, 2016, Buczinski et al., 2013, Buczinski et al., 2018b), nous avons décidé de nous limiter uniquement à l'amélioration de l'échographie thoracique pour détecter le statut actif de la BPI dans cette thèse (**CHAPITRE 4**). Revenons maintenant sur l'amélioration de ces tests.

I-1-3- Amélioration du score clinique respiratoire diagnostique

Comme détaillé dans le **CHAPITRE 1**, les problématiques des SCRD précédemment rapportés chez les veaux de race laitière étaient : (1) leurs défaillances de développement et de validation externe pouvant impacter leurs performances dans des populations variées (**Tableau 11**) ; (2) l'absence de ligne directrice approuvée sur leur utilisation pour traiter un groupe de veaux ; (3) leur répétabilité faible à modéré entre différents opérateurs (**Tableau 20**).

i- Amélioration du développement du score clinique respiratoire diagnostique

Afin de permettre de détecter les veaux de race laitière à traiter de BPI de manière universelle, nous avons décidé de développer un nouveau SCRD pour détecter le statut actif de BPI, en suivant les recommandations de médecine humaine du TRIPOD (Moons et al., 2015). Le **Tableau 46** compare les caractéristiques du développement de ce nouveau score par rapport aux SCRD précédents.

Score clinique	Étape 1		Étape 2		Étape 3 : Validation interne	Étape 4 : Validation externe	Limites
	1a : Sélection des prédicteurs	1b : Évaluation des prédicteurs	2a : Mesure de la bronchopneumonie infectieuse	2b : Calcul des poids			
SCORE CIS (Amrine et al., 2013)	Sélection subjective	Évaluation subjective par une échelle de gravité.	Aucune mesure statistique de la relation des prédicteurs choisis avec la maladie		Aucune	Aucune	Absence de développement
SCORE DU WISCONSIN (McGuirk, 2008)	Sélection subjective	Évaluation subjective par une échelle de gravité.	Aucune mesure statistique de la relation des prédicteurs choisis avec la maladie.		Aucune	Validé dans différentes populations (Buczinski et al., 2015, Love et al., 2016) et races (Lowie et al., 2022)	<ul style="list-style-type: none"> ◆ Absence de développement et de validation interne ◆ Validation externe utilisant un test imparfait pour définir la maladie (biais de mauvaise classification) Love et al., (2016) et Lowie et al., (2022)
SCORE CALIFORNIEN (Love et al., 2014)	Sélection objective : choix des prédicteurs selon leur relation linéaire avec la présence de la maladie.	Sélection objective : choix de la technique d'évaluation selon la simplicité et les résultats de validation des modèles statistiques.	Deux modalités pour un veau d'être défini malade : 1. PCR positive pour BRSV, IBR, ou BVD Et/ou 2. Score de Wisconsin ≥ 5 + culture positive de bactéries aérobique ou <i>Mycoplasma spp.</i>	Analyse de régression logistique multivariée.	Aucune	Validé dans différentes populations (Love et al., 2016) et races (Lowie et al., 2022)	<ul style="list-style-type: none"> ◆ Biais de classification lors du développement et de la validation externe (définition de la maladie avec des tests imparfaits) ◆ Non limité au statut à traiter de la bronchopneumonie infectieuse ◆ Biais d'incorporation (utilisation du score du Wisconsin pour définir la maladie ce qui peut surestimer les poids due à la colinéarité entre le test utilisé pour définir la maladie et les mêmes signes cliniques évalués) ◆ Absence de validation interne
SCORE CALIFORNIEN Québec (Buczinski et al., 2018c)	Sélection objective : choix des mêmes prédicteurs que le score Californien.	Sélection subjective avec le choix de la même technique d'évaluation que le score Californien.	Analyse bayésienne de classe latente à partir des résultats de l'échographie thoracique.	Analyse de régression logistique multivariée avec des <i>priors</i> sur les coefficients selon un sondage d'expert	Par analyse Bayésienne	Aucune	<ul style="list-style-type: none"> ◆ Non limité au statut à traiter de la bronchopneumonie infectieuse ◆ Absence de validation externe
SCORE DE MAIER (Maier et al., 2019)	Sélection objective : choix des prédicteurs selon leur relation linéaire avec la présence de la maladie.	Sélection subjective : Présence-Absence.	Utilisation en parallèle de l'auscultation thoracique + l'échographie thoracique.	Analyse de régression logistique multivariée.	Par validation croisé	Aucune	<ul style="list-style-type: none"> ◆ Biais de classification (définition de la maladie avec des tests imparfaits) ◆ Non limité au statut à traiter de la bronchopneumonie infectieuse ◆ Absence de validation externe
SCORE DE BERMAN-VEAUX LOURDS	Sélection objective : choix des signes cliniques les plus répétables entre des professionnels de veaux lourds.	Sélection de l'échelle de gravité la plus répétable entre des professionnels de veaux lourds.	Analyse bayésienne de classe latente à partir des résultats de l'échographie thoracique (consolidation ≥ 3 cm) et du dosage de l'haptoglobine (≥ 0.25 g/l) Statut actif de la maladie.	Analyse de régression logistique multivariée avec des <i>priors</i> sur les coefficients.	Par analyse Bayésienne	Validation externe temporelle	<ul style="list-style-type: none"> ◆ Faible prévalence de bronchopneumonie infectieuse et ainsi imprécision des estimés de la Se individuelle

Tableau 46. – Comparaison des caractéristiques de développement des scores cliniques respiratoires diagnostiques précédemment décrits chez les veaux de race laitière avec celui développé et validé dans cette thèse.

En améliorant son développement, ce nouveau SCRD a l'avantage :

- **D'inclure des signes cliniques répétables** par les différents professionnels de l'industrie de veaux lourds (producteurs, médecins vétérinaires, techniciens), favorisant sa répétabilité et sa généralisation ;
- **D'avoir été validé dans d'autres populations** favorisant son extrapolation et son application à d'autres populations de veaux lourds ;
- **De détecter le statut actif de la BPI** permettant de mieux détecter les veaux à traiter à l'échelle d'un groupe.

ii- Amélioration de la simplification et de l'utilisation du score clinique respiratoire diagnostique

Pour favoriser l'implantation d'une stratégie diagnostique en contexte d'élevage actuel, les tests doivent être simples à utiliser. C'est pourquoi dans cette thèse, nous nous sommes efforcés de développer un SCRD simple d'utilisation. Le **Tableau 47** résume les caractéristiques de notre SCRD par rapport aux SCRD précédents. On constate que plusieurs critères définissent sa simplicité d'action par rapport aux scores précédents :

- **Le nombre de prédicteurs est limité** (seulement 3 *versus* 4 à 6 précédemment) ;
- **L'évaluation se réalise dans un premier temps à distance** en évaluant la toux spontanée et la position des oreilles, diminuant ainsi la perte de temps lié à la manipulation des veaux et les risques de transmission de maladie. Si une manipulation est nécessaire dans un second temps pour prendre la température rectale ou induire de la toux, l'utilisation du score reste simple, sécuritaire et rapide étant donné que cette évaluation est restreinte à un maximum de 10 veaux ;
- **Utilisation à un instant t précis**, deux semaines après l'arrivée en parc d'engraissement, ce qui simplifie son utilisation et favorise son utilisation en pratique ;
- **La rapidité d'interprétation** par l'absence de calcul puisque la présence de seulement 2 signes cliniques sur 3 est nécessaire pour définir un score positif. De plus, à l'échelle d'un groupe, le seuil d'interprétation est relativement constant : nécessité de 3 veaux positifs sur 10 pour détecter une prévalence de BPI active ≥ 0.10 . Ce seuil varie uniquement pour détecter des prévalences plus faibles (< 0.10) avec la nécessité de 4 veaux positifs sur 10 (**Figure 9**, ARTICLE 2).

Score clinique respiratoire diagnostique	Population de développement du score	Prédicteurs	Évaluation des prédicteurs	Poids ¹	Interprétation Du score total
SCORE CIS (Amrine et al., 2013)	Veaux laitiers sevrés expérimentalement inoculés avec <i>Mycoplasma bovis</i> .	4 prédicteurs Comportement Dépression Toux Respiration anormale	1= comportement habituel 2= légèrement malade (léger signe de dépression ou toux) 3= modérément malade (sévèrement déprimé, respiration anormale ou toux) 4= moribond	-	Seuil arbitraire ≥ 1 (Amrine et al., 2013)
SCORE DU WISCONSIN (McGuirk, 2008)	Veaux laitiers présevrés logés en groupe.	5 prédicteurs Température rectale Jetage nasal Jetage oculaire Toux Position des oreilles	Gravité (0-3)	♦ x 1 pour tous les prédicteurs	Seuil arbitraire Deux seuils décrits : ≥ 6 (Lago et al., 2006) ≥ 5 (McGuirk, 2008)
SCORE CALIFORNIEN (Love et al., 2014)	Veaux laitiers présevrés logés individuellement.	6 prédicteurs Les 5 même prédicteurs que pour le score de Wisconsin + l'ajout de la respiration anormale	Présence ou absence	♦ x 2 Jetage oculaire, toux, respiration anormale, température rectale (seuil 39.2°C) ♦ x 4 Jetage nasal ♦ x 5 Position des oreilles	Seuil statistiquement calculé pour optimiser les performances Seuil ≥5
SCORE CALIFORNIEN QUÉBEC (Buczinski et al., 2018c)	Veaux laitier présevrés logés individuellement ou en groupe.	6 prédicteurs Les 6 mêmes prédicteurs que le score Californien	Présence ou absence	♦ x 20 : respiration anormale ♦ x 16 : position des oreilles, toux ♦ x 10 : Jetage nasal ♦ x 7 : Température rectale ♦ x-1 : Écoulement oculaire	Seuil statistiquement calculé pour optimiser les performances et variable selon la prévalence dans un groupe Seuil variable de 9 à 16 Plus la prévalence est élevée, plus le seuil est revu à la baisse
SCORE DE MAIER (Maier et al., 2019)	Veaux sevrés logés en groupe.	6 Prédicteurs Toux Respiration anormale Note indice corporel Yeux enfoncés Température rectale Température ambiante (24h)	Présence ou absence	♦ x 4 Yeux enfoncés ♦ x 5 Note indice corporel faible ♦ x 2 Toux ♦ x 1 Respiration anormale, fluctuation de température diurne ♦ x 1 Température rectale	Seuil statistiquement calculé pour optimiser les performances ≥2 si inclusion de la température dans le score Seuil après l'ajout d'un prédicteur ≥1 sans la température
SCORE DE BERMAN-VEAUX LOURDS	Veaux présevrés logés individuellement.	3 Prédicteurs Toux Position des oreilles Température rectale	Toux : Présence ou absence Position des oreilles : Présence d'oreilles tombantes uni/bilatéral ou tête penchée Température rectale : seuil ≥39.7°C	♦ x10 Toux ♦ x 9 Position des oreilles ♦ x 6 Température rectale	Seuil statistiquement calculé pour réduire les pertes économiques d'un mauvais diagnostique (MCT) . Individuel : ≥15 soit la présence de 2 signes sur les 3 . Groupe : 3 veaux positifs sur 10 veaux

¹dans cette colonne, chaque poids est indiqué par le signe multiplicateur (x) signifiant qu'il faut multiplier le nombre indiqué par 1 si le signe est présent ou par 0 si absent

Tableau 47. – Comparaison des caractéristiques des scores cliniques respiratoires diagnostique précédemment décrits chez les veaux de race laitière avec le score clinique respiratoire diagnostique développé et validé dans cette thèse.

I-1-4- Amélioration de la technique d'échographie thoracique

Plusieurs études ont permis de simplifier l'utilisation de l'échographie thoracique dans le passé (Ollivett et al., 2011, Buczinski et al., 2014, Pravettoni et al., 2021). Cependant, aucune d'entre elles ne se limitait au statut actif de la BPI et ne permettait de définir objectivement les sites à échographier et le seuil de profondeur à utiliser (**Tableau 9, CHAPITRE 1**). Les résultats du **CHAPITRE 4** ont permis de standardiser la technique échographique dans une population de faible prévalence en montrant que la détection de lésions de consolidation pulmonaire de profondeur ≥ 3 cm lors de l'échographie des sites caudaux au cœur étaient les paramètres les plus performants pour détecter le statut actif de la BPI.

I-2 Estimation de l'exactitude des tests diagnostiques pour détecter le statut actif de la bronchopneumonie infectieuse

Comme détaillé dans le **CHAPITRE 1**, l'estimation de l'exactitude des tests diagnostiques couramment utilisés chez les veaux de race laitière était : (1) parfois réalisée de manière inexacte avec la présence de biais ; (2) non restreinte au statut actif de la BPI ; (3) absente pour certains tests comme la radiographie thoracique. Les travaux de cette thèse apportent une estimation plus exacte des performances du SCRD, de l'échographie thoracique et de la radiographie thoracique pour détecter le statut actif de la BPI.

I-2-1- Les performances du score clinique respiratoire diagnostique pour détecter la prévalence de bronchopneumonie infectieuse active dans un lot

Dans cette thèse, un SCRD a été développé et validé pour détecter des lots de veaux lourds avec une certaine prévalence du statut actif de BPI, permettant de décider de l'administration ou non d'antibiotiques en prévention et notamment en métaphylaxie. Dans le passé, une seule étude avait tenté de détecter les lots prévalents de BPI mais sans résultat concluant (Leruste et al., 2012). Cet outil diagnostique est donc le premier développé avec succès pour cet effet en médecine bovine.

Étant donné que les performances des SCRD précédemment décrits étaient rapportées à l'échelle individuelle, il est difficile de comparer les performances de notre score avec les

études précédentes. Cependant, avec des VPP ≥ 0.94 et VPN ≥ 0.95 pour détecter des prévalences ≥ 0.10 du statut actif de BPI, on peut affirmer que les performances de notre score sont satisfaisantes pour détecter les lots potentiellement problématiques, nécessitant une surveillance ou l'administration d'antibiotiques au groupe.

À titre d'exemple, si on considère la prévalence problématique de BPI de 0.20 rapportée dans l'étude de Leruste et al. (2012), en appliquant notre score sur 10 veaux d'un lot, deux semaines après l'arrivée dans le parc d'engraissement, la présence de ≥ 3 veaux positifs nous permettrait d'affirmer que le lot à 99 % de chance d'avoir une prévalence de BPI active ≥ 0.20 . La présence de < 3 veaux positifs permettrait d'affirmer que le lot à 95 % de chance de ne pas avoir cette prévalence. De telles performances sont prometteuses pour favoriser une administration raisonnée d'antibiotiques en l'élevage de veaux lourds.

I-2-2- Détermination des performances de l'échographie thoracique

Les performances de l'échographie thoracique ont déjà été précédemment décrites chez les veaux de race laitière (**Tableau 17, CHAPITRE 1**). Cependant, leurs estimations étaient : soit inexactes à cause de biais présents dans les études, soit estimées chez des veaux avancés dans la maladie ou sous cliniques dont le traitement n'est pas forcément rentable ou nécessaire. Les travaux de notre thèse apportent, pour la première fois, des valeurs de performances de l'échographie thoracique pour détecter les veaux de race laitière atteints du statut actif de BPI et nécessitant un traitement d'antibiotiques et/ou anti-inflammatoires. Le **Tableau 48** illustre l'ajout de nos travaux aux connaissances actuelles de performances de l'échographie thoracique chez les veaux de race laitière.

Étude/ Design	Population de veaux	Définition d'un test index positif	Confirmation de la bronchopneumonie infectieuse (Test comparateur)	Performances	Limites
Rabeling et al. (1998) / <i>two-gate</i>	Génisses de remplacement sevrées atteintes de polyarthrite ou de BPI grave.	Queue de comète, abcès, consolidation (> 0 cm)..	Test <i>gold standard</i> = Nécropsie	Se = 0.85 Sp = 0.98	◆ Biais de spectre, les cas étant des cas sévères et chroniques de la BPI
(Jung and Bostedt, 2004) / <i>two-gate</i>	Génisses de remplacement nouveau-nés de 0 à 14 jours.	Queue de comète, abcès, effusion pleurale, pneumothorax, consolidation (> 0 cm)	Test isolé imparfait = Radiographie thoracique	Se = 0.77 Sp = 1.0	◆ Biais de classification : test comparateur imparfait ◆ Biais de spectre – <i>two-gate design</i>
(Ollivett et al., 2015) / <i>one-gate</i>	Veaux mâles présevrés sans signes cliniques (score de Wisconsin < 5).	Queue de comète, abcès, effusion pleurale, pneumothorax, consolidation (> 0 cm).	Test <i>gold standard</i> = Nécropsie	Se = 0.94 (IC95% : 0.69; 1) Sp = 1.0 (IC95% : (0.64; 1)	◆ Biais de spectre incluant uniquement des veaux sous cliniques ◆ Imprécision sur les estimations de Se et Sp
(Buczinski et al., 2015b) / <i>one-gate</i>	Génisses de remplacement présevrées logées provenant de deux localisations (New York, prévalence élevée (Ollivett et al., 2011); Québec, prévalence faible (Buczinski et al., 2014).	Consolidation (\geq 1 cm).	Analyse Bayésienne de classe latente réalisée pour estimer les performances du score de Wisconsin et en utilisant des <i>priors</i> d'expert sur les performances de l'échographie thoracique.	Se = 0.79 (ICB95%: 0.66; 0.91) Sp = 0.94 (ICB95%: 0.88; 0.98)	◆ Les <i>priors</i> sur les performances de l'échographie thoracique et le faible nombre de veaux ont peu influencé les résultats (n= 191) ◆ Imprécision sur l'estimation de la Se
(Buczinski et al., 2016b) / <i>one-gate</i>	209 veaux lourds âgés d'environ 3 semaines d'âge.	Consolidation (\geq 1 cm).	Analyse Bayésienne de classe latente réalisée pour estimer les performances de l'auscultation thoracique et en utilisant des <i>priors</i> d'expert sur les performances de l'échographie thoracique.	Se = 0.77 (ICB95%: 0.60; 0.89) Sp = 0.93 (ICB95%: 0.87; 0.97)	◆ Les <i>priors</i> sur les performances de l'échographie thoracique et le faible nombre de veaux ont peu influencé les résultats (n=209) ◆ Imprécision sur l'estimation de la Se
(Berman et al., 2019) / <i>one-gate</i>	209 veaux lourds pré-sevrés d'environ 3 semaines d'âge et 301 génisses de remplacement pré sevrées.	Différents seuils de consolidation (0, 1 et 3 cm) et différents sites (caudal <i>versus</i> crânial au cœur).	Analyse Bayésienne de classe latente réalisée pour estimer les performances de l'échographie thoracique et en utilisation des <i>priors</i> sur le score de Wisconsin et le dosage de l'haptoglobine.	Meilleures performances avec un seuil de 3 cm, caudal Se = 0.89 (ICB95%: 0.55; 1.0) Sp = 0.95 (ICB95%: 0.92; 0.98)	◆ Imprécision sur l'estimation de la Se car réalisé dans une population de faible de prévalence
(Berman et al., 2021b) / <i>one-gate</i>	50 veaux hospitalisés pesant moins de 100 kg.	Toutes les lésions Consolidation (\geq 1 cm).	Panel diagnostic incluant trois experts en médecine bovine avec un accord interopérateurs de 0.58.	BPI active Se = 0.84 (IC95%: 0.60; 0.97) Sp = 0.74 (IC95%: 0.57; 0.86)	◆ Imprécision sur les estimations de Se et Sp (nombre faible de veaux et manque de puissance)
(Berman et al., 2020) / <i>one-gate</i>	50 veaux hospitalisés pesant moins de 100 kg.	Toutes les lésions Consolidation (\geq 1 cm).	Analyse Bayésienne en deux étapes en utilisation le CT scan comme test <i>gold standard</i> sur les veaux positifs aux tests d'imagerie médicale pour confirmation.	Lésions pulmonaires Se = 0.81 (ICB95%: 0.65; 0.92) Sp = 0.90 (ICB95%: 0.81; 0.96)	◆ Imprécision sur les estimations de Se et Sp (nombre faible de veaux et manque de puissance)

BPI : Bronchopneumonie infectieuse; IC95% : intervalle de confiance à 95%; ICB95% : intervalle de crédibilité à 95%; Se : sensibilité; Sp : Spécificité.

Tableau 48. – Données actualisées des sensibilités et spécificités de l'échographie thoracique pour détecter la bronchopneumonie infectieuse des veaux de race laitière à la suite de nos travaux de thèse.

I-2-3- Détermination des performances de la radiographie thoracique

Les travaux de cette thèse ont permis, pour la première fois, de déterminer les performances de la radiographie thoracique pour détecter la BPI chez les veaux de race laitière (**Tableau 49**). Ces données permettent de quantifier le potentiel diagnostique de la radiographie thoracique pour détecter des lésions pulmonaires ainsi que le statut actif de la BPI chez les veaux de race laitière.

Étude/ Caractéristique	Population de veaux	Définition d'un test index positif	Confirmation de la bronchopneumonie infectieuse (Test comparateur)	Performances
(Berman et al., 2021b) <i>/ one-gate</i>	50 veaux hospitalisés pesant moins de 100 kg.	Toutes les lésions	Panel diagnostic incluant trois experts en médecine bovine avec un accord inter-opérateur de 0.58.	BPI active Se = 0.89 (IC95%: 0.67; 0.99) Sp = 0.58 (IC95%: 0.39; 0.75)
(Berman et al., 2020) / <i>one-gate</i>	50 veaux hospitalisés pesant moins de 100 kg.	Toutes les lésions	Analyse Bayésienne en deux étapes en utilisation le CT scan comme gold standard les veaux positifs aux tests d'imagerie médicale.	Lésions pulmonaires Se = 0.86 (ICB95%: 0.62; 0.99) Sp = 0.89 (ICB95%: 0.67; 0.99)

BPI : Bronchopneumonie infectieuse; IC95% : intervalle de confiance à 95%; ICB95% : intervalle de crédibilité à 95%; Se : sensibilité; Sp : Spécificité

Tableau 49. – Sensibilités et spécificités de la radiographie thoracique pour détecter la bronchopneumonie infectieuse des veaux de race laitière.

Dans le passé, une seule étude avait rapporté les performances de la radiographie thoracique pour détecter des lésions chez des vaches adultes hospitalisées (> 13 mois) présentant soit des signes respiratoires (64 %), soit des signes cardiaques (19 %) (Masseau et al., 2008). Dans cette étude, la Se et la Sp étaient de 0.94 (IC95% : 0.87 ; 1.0) et 0.50 (IC95% : 0.10 ; 0.90), respectivement. Ces résultats chez les adultes sont similaires à nos résultats obtenus pour détecter le statut actif de BPI chez les veaux de race laitière. La détermination de ces performances chez le veau dans cette thèse permettra à l'avenir d'améliorer le test et les stratégies de détection en milieu hospitalier.

II- Validité des résultats

II-1- L'exactitude des performances des tests sélectionnés

L'inexactitude dans l'estimation des performances d'un test diagnostique est causée par différents biais secondaires à la conception de l'étude (**Tableau 14, CHAPITRE 1**). L'un des biais dominants qui affectait l'exactitude des performances des tests rapportés dans le

passé était le biais de mauvaise classification en l'absence de *gold standard* pour définir la BPI. Afin de contrôler ce biais dans nos travaux, nous avons construit un test standard de référence pour définir le statut actif de la BPI. Pour cela, nous avons procédé en deux étapes :

Étape 1 : Sélection de la méthode de construction du test standard de référence pour définir le statut actif de BPI. Aucun résultat de tests diagnostiques ne permettait, avant nos travaux, de distinguer les veaux atteints de BPI active *versus* non active. Le test standard composite n'était donc pas la méthode recommandée (Reitsma et al., 2009) (**Annexe 9**). Une analyse bayésienne de classe latente a donc été utilisée, avec le statut actif de BPI comme classe latente (ARTICLE 2, ARTICLE 3 et ARTICLE 5). Afin de favoriser une vision plus clinique, un panel diagnostique, définissant le statut actif de la BPI selon l'avis d'expert, a été utilisé pour l'ARTICLE 4.

Étape 2 : Sélection des tests imparfaits à inclure pour construire le test standard de référence. Dans cette thèse, le statut actif de la BPI se définit comme une infection accompagnée d'une inflammation d'une partie ou de la totalité du poumon. Une combinaison de tests permettant : (1) de visualiser des lésions pulmonaires et (2) de mettre en évidence l'inflammation pulmonaire, permettait de définir ce statut. Le **Tableau 50** fournit une liste non exhaustive des tests disponibles chez les veaux de race laitière pour répondre à ces critères. Les tests utilisés dans nos travaux sont accompagnés d'un astérisque (*).

Tests mettant en évidence l'inflammation pulmonaire	Tests mettant en évidence la présence de lésions dans le poumon
Dosage de biomarqueurs de l'inflammation (haptoglobine *, SAA, procalcitonine etc.) Examen et score clinique * Hématologie * Biochimie * Gaz veineux * Cytologie pulmonaire	Échographie thoracique * Radiographie thoracique * Tomodensitométrie (CT scan) *

Tableau 50. – Principaux tests disponibles pour définir le statut actif de la bronchopneumonie infectieuse.

Ainsi, les tests étudiés dans cette thèse ont été développés et leurs exactitudes déterminées en les comparant avec ce test standard de référence afin de limiter le biais de mauvaise classification (Reitsma et al., 2009).

Les autres biais ont également été pris en compte permettant de favoriser l'exactitude de nos résultats. Le **Tableau 51** résume les moyens utilisés pour limiter ces biais et favoriser l'exactitude de nos résultats (Schmidt et Factor, 2013, Buczinski et O'Connor, 2016).

Biais	Définition	Impact potentiel	Moyens utilisés dans nos travaux
Biais de spectre	Survient si la population ne représente pas le spectre complet de la maladie (cas référés, sélection de cas extrêmes).	Surestimation de la Se et Sp, les cas où le diagnostic est plus difficile étant exclus.	Utilisation d'un échantillonnage <i>one-gate</i> permettant d'avoir un spectre représentatif de la maladie.
Biais de classification	Survient quand le test comparateur n'est pas un <i>gold standard</i> 100% exact.	Sous-estimation potentielle; dépend de la corrélation entre le test index et le test comparateur.	Utilisation d'un test standard de référence permettant de limiter ce biais : analyse de classe latente et panel diagnostique. L'utilisation du panel diagnostique a permis d'avoir une vision plus clinique et proche de la réalité lors de la prise en charge d'un veau hospitalisé.
Biais de revue diagnostique	Survient quand il y a connaissance des résultats du test index lors de la réalisation et l'interprétation du test comparateur.	Généralement une surestimation des performances du test index; la connaissance des résultats du test index peut influencer l'interprétation du test comparateur.	La prise de données a été réalisée à l'aveugle.
Biais d'incorporation	Survient quand il y a incorporation du test à investiguer dans la construction du test comparateur.	Généralement une surestimation des performances du test index.	Les tests étudiés n'ont pas été incorporés dans la construction du test comparateur.
Biais de vérification partielle	Survient quand la totalité des individus ne sont pas testés. Les résultats du test index a un impact sur la probabilité de réaliser le test comparateur.	L'absence de test comparateur chez les individus non testés peut sur ou sous-estimer les performances du test index.	Lorsque la totalité des individus ne pouvaient être testés (coût lié au CT scan), une méthode statistique bayésienne de classe latente en deux étapes a permis de prendre en compte les erreurs possibles de cette vérification partielle.
Biais de vérification différentielle	Survient quand différents tests comparateurs sont appliqués aux individus.	L'exactitude des différents tests comparateurs peuvent différer et doivent être pris en compte dans les analyses d'exactitude du test index	Les mêmes tests comparateurs ont été appliqués aux individus.
Biais de revue clinique	Survient quand les informations cliniques de l'individu recevant le test sont manquantes	Dans ce cas-là, les résultats du test index ont besoin d'un contexte spécifique.	Les informations sur les différentes populations d'étude ont été détaillées.
Biais de résultats indéterminés ou d'exclusion d'individu	Survient quand certains individus sont exclus de l'étude à cause de l'impossibilité de réaliser ou d'interpréter le test index ou le test comparateur.	Tendance à une surestimation des performances.	Aucun individu n'a été exclu à cause d'une impossibilité de réaliser ou d'interpréter le test index.

Tableau 51. – Moyens utilisés pour limiter les biais pouvant affecter l'exactitude des résultats des travaux de notre thèse.

II-2- La répétabilité des tests sélectionnés

Afin d'améliorer la détection clinique du statut actif de BPI, nous nous sommes efforcés de favoriser la répétabilité des tests sélectionnés. La technique et l'interprétation des lésions de consolidation pulmonaire de l'échographie thoracique avaient déjà été prouvées

répétables (Buczinski et al., 2013, Buczinski et al., 2018b). En revanche, la détermination de la répétabilité des SCRD n'a pas été réalisée pour tous les SCRD disponibles et, lorsqu'elle était déterminée, elle présentait des limites : comme la présence de biais (Paradoxe du Kappa) ou l'absence d'évaluation chez les producteurs (Amrine et al., 2013, Buczinski et al., 2016a, Aly et al., 2014). Le **Tableau 52** revient sur ces limites et permet de comparer les résultats de répétabilité des SCRD précédents à nos travaux de thèse.

Précédemment, les SCRD disponibles avaient d'abord été développés avant d'en étudier leurs répétabilités (Amrine et al., 2013, Buczinski et al., 2016a, Aly et al., 2014). L'inconvénient de cette démarche est qu'en cas de mauvaise répétabilité du score, l'unique façon de l'améliorer est de former ses utilisateurs, ce qui nécessite des moyens temporels et financiers supplémentaires. À l'inverse, dans cette thèse, nous avons d'abord déterminé la répétabilité individuelle des prédicteurs potentiels de notre score avant de décider de les inclure ou non. De cette manière, les prédicteurs de notre score sont répétables garantissant un SCRD potentiellement répétable avant son implantation. De plus, pour la première fois, l'étude de la répétabilité a inclus des professionnels autres que des médecins vétérinaires à savoir des producteurs et des techniciens susceptibles également d'utiliser ce test.

Autre point de divergence avec les études précédentes, nous nous sommes efforcés dans cette thèse de déterminer la répétabilité de chaque signe clinique le plus précisément et exactement possible en : (1) calculant au préalable le nombre de veaux et d'opérateurs nécessaires pour optimiser la précision d'estimation des indicateurs de répétabilité et utiliser l'échelle de Gwet (2008) pour les interpréter (**Tableau 19, CHAPITRE 1**) ; (2) utilisant plusieurs indicateurs comme recommandé pour juger au mieux de la répétabilité (Morris et al., 2008, O'Leary et al., 2014, Walsh et al., 2014) ; (3) en limitant le paradoxe du Kappa en fixant la prévalence à 0.5 (Dohoo et al, 2012) et en utilisant un indicateur paradoxe-résistant (Gwet AC1) (Gwet, 2014).

Études	Individus	Opérateurs	Évaluation	Indicateurs	Interprétation	Limites
Score CIS (Amrine et al., 2013)	178 génisses de remplacement sevrées expérimentalement inoculées par <i>Mycoplasma bovis</i> .	♦ 9 opérateurs (médecins vétérinaires) d'expériences variables ♦ Évaluation le même jour	Binaire – CIS positif si malade (≥ 1), négatif si non malade ($=0$)	Kappa de Cohen médian (κ)	$\kappa = 0.16$ (Intervalle de confiance (IC) 95%: 0.10; 0.24) Échelle de Landis et Koch (1977) : Accord pauvre à léger	♦ Utilisation d'un seul indicateur ♦ Pas d'ajustement au paradoxe du Kappa ♦ Restreint aux médecins vétérinaires
Score du Wisconsin (Buczinski et al., 2016a)	Veaux lourds présevrés en trois visites (V) : ♦ V1 : 10 veaux aléatoires à l'arrivée en parc engraissement ; ♦ V2 : 10 veaux aléatoires à 12 jours d'engraissement ; ♦ V3 : 20 veaux à 21 jours d'engraissement ; 10 veaux avec des lésions échographiques de consolidation pulmonaire et 10 veaux aléatoires.	♦ 2 (V1 et V2) ou 3 opérateurs (V3) (un médecin vétérinaire et deux étudiants vétérinaires), tous peu expérimentés ♦ Évaluation le même jour à 4 heures d'intervalle.	Score de Wisconsin, ordinal pour chaque signe clinique de 0 à 3 et considéré malade selon différents seuils (≥ 4 , ≥ 5 , et ≥ 6)	♦ Pourcentage d'accord médian (Pa) ♦ Kappa de Cohen médian pondéré ($w\kappa$) ♦ Fleiss (F) pour opérateurs multiples	Échelle de Landis et Koch (1977) Signes cliniques : accord pauvre à modéré Toux Pa=0.78 $w\kappa = 0.1$ Température Pa = 0.63 $w\kappa = 0.6$ Jetage nasal Pa = 0.62 $w\kappa = 0.24$ Jetage oculaire Pa=0.63 $w\kappa = 0.11$ Oreilles Pa=0.85 $w\kappa = -0.04$ Les différents seuils : accord pauvre à juste ≥ 4 : léger F = 0.35 ≥ 5 : pauvre F = 0.06 ≥ 6 : pauvre F = 0.13	♦ Pas de moyen de juger de la précision autour de la valeur de l'indicateur ♦ Biais du paradoxe du Kappa exprimé statistiquement par l'index de prévalence et jugé important ♦ Restreint aux médecins vétérinaires
Score Californien (Aly et al., 2014)	100 génisses de remplacement présevrées suspectées malades par auscultation thoracique et/ou échographie thoracique.	Un seul opérateur (médecin vétérinaire) expérimenté qui réalise le score Californien et le score de Wisconsin.	Accord comparé avec le score de Wisconsin	Kappa de Cohen (κ)	$\kappa = 0.85$ (Déviation standard : 0.099) Échelle de Fleiss et al. (1981) : Accord presque parfait entre le score Californien et Wisconsin	♦ Score Californien développé à partir du score de Wisconsin donc accord anticipé ♦ Pas de calcul de la répétabilité inter opérateur du score Californien ♦ Un seul opérateur médecin vétérinaire
(Berman et al., 2021a)	30 vidéos de chaque signes cliniques de BPI (jetage nasal, jetage oculaire, toux, respiration anormale, positions des oreilles), dont 15 avec des signes cliniques et 15 sans (prévalence de 0,5).	6 producteurs, 6 médecins vétérinaires et 6 techniciens, tous expérimentés dans l'évaluation de la BPI en élevage de veaux lourds.	Étude de différentes combinaisons de variations pour chaque signe clinique : ordinal de 0 à 2 ou 3 selon les signes, et binaire (0,1/2,3; 0/1,2,3; 0/1,2)	♦ Pourcentage d'accord médian (Pa) ♦ Kappa de Cohen médian (κ) ♦ Gwet AC1 (AC1) pour opérateurs multiples	Échelle de Gwet (2008) Signes cliniques : accord pauvre à modéré Toux (0/1,2) Pa = 0.90 (IC95%: 0.73;0.98) $\kappa = 0.80$ (IC95%: 0.58;1.00) AC1 = 0.85 (IC95%: 0.75;0.95) Jetage nasal (0,1/2,3) Pa = 0.83 (IC95%: 0.65;0.94) $\kappa = 0.45$ (IC95%: 0.07;0.84) AC1 = 0.71 (IC95%: 0.57;0.85) Jetage oculaire (0,1/2,3) Pa = 0.80 (IC95%: 0.61;0.92) $\kappa = 0.44$ (IC95%: 0.07;0.86) AC1 = 0.77 (IC95%: 0.65;0.89) Oreilles (0,1/2,3) Pa = 0.83 (IC95%: 0.65;0.94) $\kappa = 0.67$ (IC95%: 0.42;0.92) AC1 = 0.60 (IC95%: 0.44;0.76) Respiration (0/1,2) Pa = 0.70 (IC95%: 0.51;0.92) $\kappa = 0.42$ (IC95%: 0.13;0.71) AC1 = 0.59 (IC95%: 0.47;0.71)	♦ Utilisation de vidéos et de signes cliniques séparés s'éloignant de l'évaluation dans la vraie vie

Tableau 52. – Données actuelles de répétabilité des signes cliniques de la bronchopneumonie infectieuse chez les veaux de race laitière à la suite de nos travaux de thèse.

II-3- La précision des tests sélectionnés

Une des limites de nos travaux est le manque de précision dans l'estimation de certains paramètres, comme en témoignent les larges IC et ICB rapportés autour de la Se du SCRD (0.31 [ICB95% : 0.14 ; 0.70] ; ARTICLE 2) ou encore, de la Se d'échographie thoracique (0.89 [ICB95% : 0.55 ; 1.0] ; ARTICLE 5). Ce manque de précision est lié à une prévalence faible du statut actif de la BPI dans les populations d'étude (prévalence de BPI de 0.05). Cette faible prévalence signifie une faible proportion de veaux malades. L'estimation des tests pour détecter le statut actif de la BPI chez ces veaux malades, à savoir la Se, est alors réalisée sur une faible proportion de veaux, rendant l'estimation moins précise. Une des conséquences est que ce manque de précision autour de la Se de l'échographie thoracique a probablement limité la détection de différence entre les sites et les seuils dans l'ARTICLE 5 par manque de puissance statistique (Dohoo et al., 2012). Cette prévalence est faible dans ces études car les échantillonnages ont été réalisés avant le pic d'incidence de BPI afin d'une part, de favoriser l'amélioration de nos tests lors de détection précoce, et d'autre part, échantillonner une population ayant reçu peu d'antibiotiques afin de ne pas influencer nos résultats. Des études supplémentaires dans des populations à prévalence plus élevée sont nécessaires pour valider les sites et le seuil trouvés dans nos travaux. À noter qu'en revanche, cette limite ne s'applique pas pour la valeur de Sp car la faible prévalence signifie en contrepartie une forte proportion de veaux non malades permettant d'estimer avec précision les Sp et de favoriser la puissance statistique pour détecter des différences entre elles. Cependant, l'application de nos tests sur ces populations de faibles prévalences a pu rendre l'échographie thoracique plus spécifique que ce qu'elle serait dans une population de plus grandes prévalences.

On retrouve également un manque de puissance statistique pour comparer et détecter une différence entre l'échographie et la radiographie thoracique. Ce manque de puissance n'est cette fois-ci pas dû à la prévalence faible de BPI, mais à un nombre limité de veaux inclus dans les ARTICLES 3 et 4. Ainsi, il est possible qu'une différence existe entre la Se et la Sp des deux tests, mais qu'elle n'est pas pu être mise en évidence par ce manque de puissance statistique (Dohoo et al., 2012). Pour illustration, la Sp de l'échographie pour détecter le statut actif de BPI (ARTICLE 4) était presque significativement supérieur (McNemar's; $p = 0.06$) à celle de la radiographie ($Sp = 0.74$ [95%CI : 0.57 ; 0,86] *versus*

Sp = 0,58 [95%CI : 0,39 ; 0,75], respectivement). Des études supplémentaires, incorporant un nombre plus important de veaux, permettraient d'augmenter la puissance statistique nécessaire pour confirmer ou infirmer une telle différence entre les deux tests.

II-4- La validation externe de nos résultats

La validation externe de nos résultats n'a pas été étudiée dans certains de nos travaux. De nombreux facteurs pourraient faire varier nos résultats lors de leurs implications dans une autre population. La connaissance de ces facteurs est indispensable pour juger de la pertinence de l'application de ces résultats et aiguiller les études futures pour les améliorer et les valider. Voici la liste non exhaustive de ces facteurs :

La prévalence de la maladie : Comme mentionné précédemment, la faible prévalence dans nos populations d'étude dans les ARTICLES 2 et 5 a pu influencer nos résultats. Il est possible que les performances obtenues du SCRD et de l'échographie thoracique soient différentes dans une population de plus forte prévalence ou la Se sera plus précisément estimée, mais également privilégiée pour appliquer ces tests.

L'agent microbien : Dans nos études sur l'échographie thoracique (ARTICLES 3-5), la nature de l'agent microbien en cause de la BPI n'a pas été déterminée. Or, Ollivett et al. (2015) ont montré que la majorité des veaux sous cliniques (6 sur 7) atteints de BPI virale n'avaient soit aucune consolidation pulmonaire, soit des lésions de consolidation < 1 cm. À l'inverse, les veaux atteints de BPI bactérienne avaient des profondeurs plus importantes, majoritairement ≥ 3 cm. Nos résultats sont donc susceptibles de varier selon la nature de l'agent microbien en cause, virus *versus* bactéries. De la même manière, il est possible que des différences de lésions existent entre les différentes bactéries comme c'est le cas macroscopiquement (**Partie I; I-2-3**). Par exemple, *Mannheimia haemolytica*, *Mycoplasma bovis* et *Histophilus somni* peuvent être responsables d'inflammation fibrineuse de l'espace pleural. Les lésions d'effusion pleurale à l'échographie thoracique seraient donc un bon indicateur de ces agents et seraient plus représentées si ces bactéries sont prévalentes. Il en serait de même pour le BRSV et les lésions de pneumothorax ou encore *Trueperella pyogenes* et les abcès pulmonaires (Panciera et Confer, 2010). La nature de l'agent peut également influencer les signes cliniques inclus dans notre score. En effet, dans notre population de développement de notre SCRD, *Mycoplasma bovis* souvent

retrouvé lors d'otite, était l'agent dominant (**Tableau 28**, ARTICLE 2). Les professionnels inclus dans l'étude ont donc pu être plus au fait du signe clinique d'otite à savoir la position anormale des oreilles, et être plus en accord pour le détecter dans l'ARTICLE 1.

Le type de lésion pulmonaire : Lors de la comparaison entre l'échographie et la radiographie thoracique (ARTICLES 3 et 4), la faible prévalence de certaines lésions pulmonaires n'a pas permis de savoir si un des deux tests était meilleur pour détecter un certain type de lésion. Il est possible que pour certaines lésions, l'un des deux tests soit meilleur pour les détecter, comme c'est le cas chez les chevaux de la radiographie thoracique pour les pneumothorax (Partlow et al, 2017) ou l'échographie thoracique pour l'effusion pleurale (Reef et al, 1991). La stratégie d'utilisation d'un test par rapport à l'autre serait alors variable selon le type de lésion pulmonaire suspectée.

L'âge des veaux : Dans cette thèse, l'échographie thoracique a été améliorée chez des veaux âgés < 2 mois. À cet âge, l'accès au site crânial au cœur est facilité et la profondeur du poumon favorise une exposition des lésions pulmonaires en surface optimisant leurs détections par échographie (Babkine et Blond, 2009). Il est donc possible que la faculté de détection des lésions soit plus limitée chez des veaux plus âgés et que la radiographie thoracique soit plus performante pour détecter le statut actif de BPI ou les lésions. De la même manière, notre SCRD a été développé pour être utilisé chez les veaux présevrés, deux semaines après l'arrivée. Son efficacité sur des veaux plus âgés n'est pas garantie en raison de leurs présentations cliniques qui peuvent être différentes (Woolums, 2015). Son utilisation plus tardivement lors de la période d'engraissement reste donc à être validée.

Les conditions d'élevage : Notre SCRD a été développé dans les veaux lourds. Son applicabilité ne peut être garantie chez les génisses de remplacement où les conditions d'élevage sont différentes. De la même manière, les résultats de cette thèse n'ont pas été validés chez les veaux de race de boucherie dont la conformation et les conditions d'élevage peuvent faire varier nos résultats.

L'opérateur : la répétabilité des signes cliniques de la BPI a été déterminée chez des opérateurs expérimentés dans l'évaluation de la maladie chez les veaux lourds. Il est possible que nos résultats diffèrent avec des opérateurs moins expérimentés, ou encore travaillant avec des génisses de remplacement.

III- Des résultats de notre thèse vers l'élaboration d'algorithmes décisionnels

Un algorithme décisionnel se définit comme un outil d'aide à la décision. En médecine humaine, des algorithmes décisionnels sont déjà utilisés pour améliorer le diagnostic et traiter les pneumonies infectieuses (Christ-Crain et al., 2006, Lippi et al., 2011). On retrouve de tels algorithmes décisionnels chez les veaux de race laitière pour la gestion des diarrhées néonatales (Gomez al., 2017). Cependant, de tels algorithmes sont actuellement absents pour la BPI. En améliorant les tests pour détecter les **veaux à traiter**, les résultats de cette thèse pourraient être utilisés pour élaborer des algorithmes décisionnels pour mieux détecter et traiter la BPI à l'échelle individuelle et à l'échelle d'un groupe de veau. Développons un peu plus cette perspective dans les paragraphes suivants.

III-1- À l'échelle individuelle

Nos résultats montrent que l'échographie thoracique (site caudal au cœur, profondeur de consolidation pulmonaire ≥ 3 cm) est très spécifique pour détecter le statut actif de BPI (Sp = 0.95 [ICB95% : 0.92 ; 0.98]). Autre caractéristique, bien que la VPP soit plus variable, la VPN de l'échographie thoracique reste $> 0,90$ quelle que soit la prévalence (**Figure 15**, ARTICLE 5). Cliniquement, cela signifie que ce test a peu de chance de classer un veau non malade comme atteint du statut actif de BPI et donc de le traiter inutilement (FP). Par conséquent, une des stratégies diagnostiques prometteuses pour détecter les veaux atteints de BPI serait d'appliquer à chaque veau un premier test de dépistage sensible pour détecter le plus de veaux malades possible, puis d'utiliser l'échographie thoracique par la suite pour confirmer ou infirmer la maladie et instaurer ou non un traitement.

La **Figure 16** propose une ébauche d'un algorithme décisionnel de traitement de la BPI à l'échelle individuelle. Développons en détail les particularités de cet algorithme et les recherches futures à envisager pour le développer.

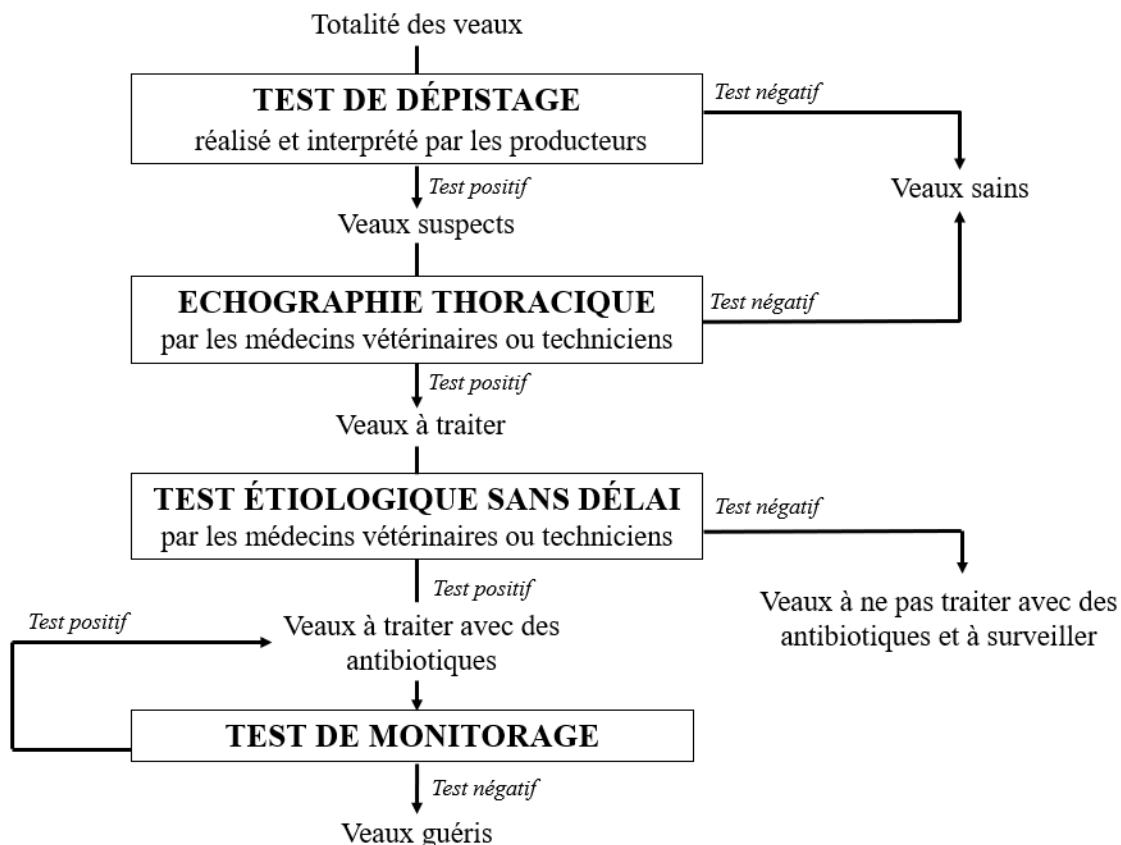


Figure 16. – Proposition d’un algorithme décisionnel pour mieux détecter et traiter un veau de race laitière atteint de bronchopneumonie infectieuse.

III-1-1- Les utilisateurs

Dans les conditions d’élevages actuels, une des approches les plus appropriées serait que le premier test de dépistage soit réalisé et interprété par les producteurs pour, par la suite, faire intervenir les médecins vétérinaires, ou des techniciens entraînés, pour confirmer la suspicion à l’échographie thoracique. Nos différences de répétabilité des signes cliniques entre les producteurs, les techniciens et les médecins vétérinaires ont rejoint ceux de Sischo et al. (2019), dénonçant un manque de standardisation d’évaluation entre ces professions. Dans le cas d’utilisation d’un algorithme décisionnel, il est essentiel que son élaboration et son utilisation impliquent l’enseignement et la collaboration des professionnels afin de favoriser une utilisation optimale de cet outil.

III-1-2- Les tests de dépistage de la bronchopneumonie infectieuse

Idéalement, un test de dépistage permettrait de détecter la BPI précocement. Pour ce faire, il doit être simple d'utilisation, peu coûteux, répétable et avoir une grande Se afin d'identifier une grande proportion de veaux malades tout en gardant une Sp modérée pour garder confiance au test (Wolfger et al., 2015). Le **Tableau 53** répertorie les caractéristiques des principaux tests étudiés pour dépister précocement la BPI par les producteurs chez les veaux de race laitière.

	Sensibilité	Spécificité	Répétabilité	Référence
SCRD	0.48 à 0.68*	0.72 à 0.86**	Pauvre à légère	(Amrine et al., 2013) (Buczinski et al., 2015b) (Love et al., 2016) (Buczinski et al., 2018c) (Maier et al., 2019)
Signe clinique isolé***				
Température rectale > 39.4°C	0.26	0.89	Pauvre à modérée	(Lowie et al., 2022) (Buczinski et al., 2016a)
Tachypnée (≥ 43 mvt/min)	0.36	0.83		
Toux spontanée	0.40	0.86		
Jetage nasal	0.13	0.94		
Mesure du changement individuel de comportement (automatisé ou non automatisé)	0.23 à 0.68	0.43 à 0.95	Non investiguée	Revue de littérature de (Cramer et Ollivett, 2020)
Détecteurs de toux pour un groupe de veau via un microphone installé dans l'enclos	0.50	0.99	Non investiguée	(Vandermeulen et al., 2016, Carpentier et al., 2018)

*Calcul de la moyenne des bornes inférieures et supérieures des sensibilités de dépistage disponible pour les scores de Wisconsin, Californien et de Maier.

**Calcul de la moyenne des bornes inférieures et supérieures des spécificités des SCRD rapportés chez les veaux de race laitière.

***Performances calculées par rapport à la présence de consolidation pulmonaire.

Tableau 53. – Performances des principaux tests de dépistage de la bronchopneumonie infectieuse chez les veaux de race laitière.

On constate qu'actuellement aucun test ne présente les caractéristiques adéquates citées précédemment pour dépister avec exactitude les veaux de race laitière atteints de BPI. Une amélioration de ces tests ou le développement de nouveaux outils sont à prioriser à l'avenir. L'extension des recherches vers l'extrapolation d'outils déjà utilisés dans les bovins de boucherie serait prometteuse comme : (1) le **stéthoscope électronique** (*Whisper®*) utilisé par les producteurs et dont les performances rapportées sont intéressantes avec une Se de 0.92 (ICB95% = 0.71 ; 0.99) et une Sp de 0.90 (ICB95 % : 0.64 ; 0.99) (Mang et al., 2015) ; ou encore (2) la mesure de **la température** précédant les signes cliniques (Timsit et al.,

2011a, b) par thermographie infrarouge (Schaefer et al, 2007 ; 2012) ou par capteur d'oreille (McCorkell et al., 2014). Autre moyen à l'amélioration, la poursuite de l'identification **de signes cliniques** pour détecter les veaux ayant besoin d'être traité qui permettrait de sensibiliser les producteurs à l'observation et la détection ces signes cliniques spécifiquement et le développement d'éventuels outils pour les mesurer (p. ex. ; détecteur de toux individuel).

III-1-3- Le traitement à donner

Un algorithme décisionnel aboutit à une action, en l'occurrence ici, l'administration d'un traitement. Bien que nos résultats permettent de définir un veau atteint du statut actif de la BPI et donc nécessitant un traitement, nous ne sommes pas en mesure de savoir si ce traitement est un traitement d'antibiotiques, d'anti-inflammatoire ou les deux. Une des pistes de recherche pour connaître lequel donner serait d'avoir des renseignements sur l'étiologie de la BPI au moment du diagnostic et non *a posteriori* comme le proposent les tests actuels (**Tableau 3, CHAPITRE 1**).

Pour répondre à cela, l'utilisation d'un autre test après le test de confirmation d'échographie thoracique pourrait être envisagée comme la **cytologie pulmonaire** en développement actuellement (Van Leenen et al., 2020) ou la découverte future d'un **biomarqueur spécifique** d'une infection bactérienne comme la procalcitonine en médecine humaine (Christ-Crain et al., 2006, Pinzone et al., 2014). Une autre piste d'amélioration serait de ne pas utiliser un autre test, mais de raffiner l'échographie thoracique afin d'associer **des lésions échographiques à des agents étiologiques spécifiques** (voir **II-4** de ce chapitre). La connaissance et le suivi des lésions typiques à un agent étiologique permettra également dans le futur de voir si tous les agents ou lésions ont besoin d'un traitement d'antibiotiques pour guérir ou être autolimitants comme c'est le cas des lésions de petites tailles à *Rhodococcus equi* chez les poulains (Arnold-Lehna et al., 2020). Les agents microbiens responsables de BPI (**Tableau 1, CHAPITRE 1**) pourront alors être classés selon la nécessité ou non d'un traitement pour favoriser la guérison, ce qui facilitera l'utilisation raisonnée des antibiotiques.

III-1-4- Suivi dans le temps

Au-delà de savoir quel veau nécessite un traitement d'antibiotiques ou non, la durée de celui-ci est également essentielle pour optimiser la guérison de la maladie et devrait constituer une indication à inclure dans l'algorithme décisionnel. En pratique, une durée de 5 à 10 jours de traitement d'antibiotiques est souvent utilisée. Cependant, les évidences scientifiques justifiant cette durée sont limitées et les critères permettant de juger du succès du traitement sont peu définis (Apley, 2015, DeDonder et Apley, 2015). De plus, historiquement, le critère pour juger la fin d'un traitement était basé sur la disparition des signes cliniques liés à la BPI (Buczinski et Pardon, 2020). Or, d'après les travaux d'Ollivett et al. (2013), on sait que les lésions pulmonaires détectables à l'échographie persistent au-delà des signes cliniques. Arrêter le traitement selon des critères cliniques ne garantit donc pas une guérison du poumon (Binversie et al., 2020). C'est pourquoi une étude récente par Jourquin et al. (2022) a utilisé la disparition des lésions de consolidation comme critère d'arrêt du traitement d'antibiotiques, comme cela est pratiqué en médecine humaine (Bouhemad et al., 2010, Bello and Blanco, 2019). Les résultats de cette étude se sont avérés prometteurs montrant que les veaux traités selon ce critère de durée n'avaient pas de recrudescence de la maladie à l'échographie 14 jours post-traitement avec des antibiotiques. Des études à plus long terme sont nécessaires pour juger de l'implantation d'une telle pratique dans le suivi des traitements et utiliser l'échographie thoracique comme critère de succès thérapeutique. L'échographie thoracique pourra alors être intégrée comme test de monitoring.

III-2- À l'échelle d'un groupe de veaux

Les travaux de notre thèse apportent, pour la première fois, un algorithme décisionnel pour décider de l'instauration d'un traitement d'antibiotiques de groupe résumé dans le **Tableau 54**. Même si des études de validation externe sont encore nécessaires (voir **II-4** de ce chapitre), cet outil simple à appliquer est prometteur pour mieux traiter la BPI en élevage de veaux lourds.

Test utilisé	Utilisation	Résultat	Traitement
SCRD	Sur 10 veaux, 2 semaines après l'arrivée en parc d'engraissement par des producteurs, techniciens ou médecins vétérinaires	Présence de ≥ 3 veaux sur 10 avec 2 des 3 prédictors suivants : toux spontanée ou induite, position anormales des oreilles, température augmentée ≥ 39.5	Administration d'un traitement d'antibiotiques de groupe
		Présence de < 3 veaux sur 10 avec 2 des 3 prédictors suivants : toux spontanée ou induite, position anormales des oreilles, température augmentée ≥ 39.5	Surveillance du lot

Tableau 54. – Proposition d'un algorithme décisionnel pour décider d'un traitement d'antibiotiques d'un lot de veaux lourds.

IV- Impacts anticipés de nos résultats

Bien que non investigués dans nos travaux, les résultats de cette thèse pourraient ultimement avoir deux impacts anticipés : (1) diminuer les pertes économiques de la maladie et améliorer le bien-être des veaux en détectant mieux les veaux malades ; (2) réduire l'emploi des antibiotiques en traitant uniquement les veaux ou groupe de veaux qui en ont besoin. Développons d'avantage ces deux hypothèses.

IV-1- Les impacts anticipés sur l'économie de l'industrie des veaux de race laitière

Dans cette thèse, nous avons amélioré les stratégies diagnostiques pour détecter les veaux de race laitière atteints du statut actif de BPI. L'implantation de nos résultats en pratique permettra donc de mieux détecter les veaux malades et ainsi diminuer les impacts économiques à long terme énumérés **Partie II** du **CHAPITRE 1**.

IV-1-1- le gain économique anticipé

Afin d'avoir une idée du gain économique anticipé chez les génisses de remplacement et les veaux lourds, prenons comme exemple les prévisions rapportées en France par Delabougliose et al. (2017) (**Tableau 55**). A noter que la prévalence de BPI en France avoisine les 0.20 chez les génisses de remplacements et les 0.50 chez les veaux lourds (Brscic et al., 2012). Ces chiffres sont similaires à ceux rapporté en Amérique du Nord (**CHAPITRE 1, Partie II, II**).

Pourcentage d'augmentation de productivité*	Pourcentage théorique de réduction d'incidence de BPI		
	20%	50%	100% (éradication)
Secteur de génisses de remplacement	0 (IC95% : 0; 0.1)	0.1 (IC95% : 0.1; 0.1)	0.2 (IC95% : 0.1; 0.3)
Secteur de veaux lourds	0.5 (IC95% : 0.4; 0.6)	1.3 (IC95% : 1.1; 1.4)	2.6 (IC95% : 2.2; 2.9)
Bénéfice de l'éradication de la BPI au niveau national (millions euro/année)			
Secteur de génisses de remplacement	14.5 (IC95% : 8.6-21.4)		
Secteur de veaux lourds	21.4 (IC95% : 18.3-24.5)		

* La productivité = (la valeur de gain de tous les produits perçus en un an) – (les coûts supplémentaires (achat d'animaux, les prix des traitements)) / (le total des énergies métabolisables requises (fourrage, ensilage, concentrés, le lait de remplacement))

Tableau 55. – Changements prédictifs de productivité en réponse à la réduction de l'incidence de bronchopneumonie infectieuse (BPI) et bénéfice de l'éradication dans le secteur de génisse de remplacement et de veaux lourds. Adapté de Delabouglise et al. (2017).

D'après ce tableau, on constate que, la réduction de l'incidence de la BPI (allant de 20 % à l'éradication) est avantageuse économiquement, particulièrement pour le secteur de veaux lourds. Loin de pouvoir éradiquer la maladie avec nos travaux, on peut tout de même penser que l'implantation de notre SCRD permettrait de traiter un lot de forte prévalence avant son pic d'incidence (à 2 semaines *versus* le pic qui apparaît entre 3 et 5 semaines) et ainsi en diminuer la magnitude. Cette diminution d'un certain % pourrait donc apporter un gain économique à l'industrie.

IV-1-2- Les études futures à envisager

Afin de tester l'hypothèse que l'implantation de notre SCRD pourrait apporter un gain économique à l'industrie de veaux lourds, une future étude pourrait être réalisée en regardant si, lorsqu'on a 3 veaux positifs sur 10 dans un lot, deux semaines après l'arrivée, et que l'on traite le lot avec des antibiotiques en métaphylaxie (*versus* le lot pas traité), il y a alors un impact positif sur la morbidité et mortalité. L'étude pourrait consister en un essai clinique randomisé comparant la morbidité et la mortalité des lots, avec 3 veaux positifs sur 10 deux semaines après l'arrivée, traités *versus* non-traités. Dans le futur, on pourrait même envisager de comparer le poids à l'abattage comme approximation des revenus des lots appliquant notre SCRD (groupe exposé) *versus* des lots ne l'appliquant pas (groupe

non exposé). Ces idées de futures études permettraient de quantifier précisément les bénéfices de cet outil et ainsi favoriser son implantation.

IV-2- Les impacts anticipés sur l'utilisation des antibiotiques chez les veaux de race laitière

IV-2-1- À l'échelle individuelle

Dans notre thèse, on rapporte une Sp de 0.95 (95%ICB : 0.92 ; 0.98) pour détecter le statut actif de BPI de l'échographie thoracique (site caudal, seuil ≥ 3 cm). Concrètement, cela signifie que dans une population de prévalence de BPI de 0.20, seulement 2 % à 8 % des veaux seront déclarés malades, et donc potentiellement traités avec des antibiotiques alors qu'ils n'en ont pas besoin. L'un des impacts anticipés d'inclure notre définition d'interprétation de l'échographie thoracique élaborée dans cette thèse dans un algorithme décisionnel ou après une suspicion clinique serait donc une réduction de la consommation d'antibiotiques chez les veaux de race laitière en traitant uniquement les veaux qui en ont besoin. De telles réductions après l'implantation d'un algorithme décisionnel ont été démontrées en médecine humaine pour les pneumonies (Bosso et Drew, 2011) ou encore, en médecine équine pour les pneumonies à *Rhodococcus equi* (Arnold-Lehna et al., 2020). Récemment, l'implantation d'un algorithme décisionnel pour traiter les veaux en diarrhée a diminué l'utilisation d'antibiotiques de 80 % (Gomez et al., 2017).

Ainsi, comme dans ces autres maladies ou espèces animales, la poursuite de la recherche vers l'élaboration d'un algorithme décisionnel comme suggéré **Figure 16**, permettra de favoriser une meilleure utilisation des antibiotiques pour la BPI chez les veaux de race laitière, à l'échelle individuelle.

IV-2-2- À l'échelle d'un groupe de veaux

Comme développé extensivement dans notre revue de littérature, **Partie II** du **CHAPITRE 1**, l'administration d'antibiotiques de groupe en métaphylaxie constitue la majeure consommation d'antibiotiques chez les veaux lourds (Pardon et al., 2012a, Fertner et al., 2016, Lava et al., 2016b, Schnyder et al., 2019). Chez les bovins de boucherie, l'administration d'antibiotiques à l'échelle d'un groupe en métaphylaxie a montré un intérêt uniquement dans les groupes où la prévalence de BPI est élevée (Baptiste et Kyvsgaard, 2017). Cette caractéristique se retrouve également chez les veaux lourds avec

des études montrant que le traitement de groupe n'était pas bénéfique dans les lots de faibles prévalences (Lava et al., 2016b, Berman et al., 2017). Pour la première fois en médecine bovine, nos travaux ont permis d'élaborer un outil simple et accessible aux producteurs afin de mettre en place un traitement sélectif des lots de veaux lourds avec des antibiotiques. Nous avons conscience que la décision de traiter un lot est multifactorielle dépendant, entre autres, de la qualité des veaux, de la sévérité des facteurs de risque exposé ou encore du prix à l'achat. Cependant, grâce à ce nouveau SCRD, la prévalence de statut actif de BPI dans le lot pourra être estimée et pourra être utilisée comme critère décisionnel pour traiter ou non le lot au complet avec des antibiotiques.

Des études sont nécessaires afin de valider notre SCRD du point de vue clinique et pas seulement du point de vue théorique. Une étude temporelle comparant les consommations d'antibiotiques avant et après l'implantation du score permettra de quantifier la réduction d'antibiotiques et promouvoir ou non l'implantation universelle de ce score dans l'industrie de veaux lourds. Une application mobile utilisable par les producteurs pourrait alors être développée afin de favoriser son utilisation.

CONCLUSION

Dans cette thèse, nous avons perfectionné les stratégies diagnostiques pour détecter la BPI chez les veaux de race laitière et contribué à améliorer les principaux points problématiques actuels de la BPI chez les veaux de race laitière à savoir :

1- **L'inexactitude de la définition de la BPI** pour détecter les veaux à traiter ;

En effet, dans cette thèse, nous avons défini pour la première fois, le statut actif de la BPI à partir d'un test standard de référence. Par la suite, nous avons développé un SCRD, comparé la radiographie et l'échographie thoracique et amélioré l'échographie thoracique pour détecter ce statut à partir de ce test standard de référence. Les résultats des tests obtenus permettent ainsi d'apporter une définition diagnostique de la BPI pour détecter les veaux à traiter avec des antibiotiques et / ou des anti-inflammatoires et ainsi faciliter la prise de décision thérapeutique de la maladie.

2- **Le manque de développement et de standardisation des tests** couramment utilisés pour détecter la maladie;

Dans cette thèse, nous avons amélioré le développement d'un SCRD et la standardisation de l'échographie thoracique pour détecter le statut actif de la BPI chez les veaux de race laitière.

3- **Le manque de répétabilité entre les opérateurs et les performances sous-optimales** des tests couramment utilisés pour détecter et traiter la maladie.

Dans cette thèse, nous avons optimisé la répétabilité d'un SCRD lors de son développement et estimé des performances diagnostiques acceptables de l'échographie thoracique et de la radiographie thoracique pour détecter le statut actif BPI chez les veaux de race laitière. Ainsi, en contribuant à l'amélioration de ces différents points, les travaux de cette thèse représentent une réelle avancée dans le diagnostic des BPI. Les impacts de ces améliorations sur l'économie et sur la consommation d'antibiotiques sont à anticiper et à investiguer dans le futur.

Références bibliographiques

Aarestrup, F. M. 2005. Veterinary drug usage and antimicrobial resistance in bacteria of animal origin. *Basic Clin Pharmacol Toxicol* 96(4):271-281.

Abdallah, A.A., A. M. Abdelaal, A. R. El-Sheikh, H. Selim, and S. Buczinski. 2019. Determination of the dynamics of respiratory diseases using thoracic ultrasonographic examination in preweaned dairy calves. *Can Vet J* 60(8):859.

Abdallah, A.A., J. Hewson, D. Francoz, H. Selim, and S. Buczinski. 2016. Systematic Review of the Diagnostic Accuracy of Haptoglobin, Serum Amyloid A, and Fibrinogen versus Clinical Reference Standards for the Diagnosis of Bovine Respiratory Disease. *J Vet Intern Med* 30(4):1356-1368.

Abutarbush, S. M., C. M. Pollock, B. K. Wildman, T. Perrett, O. C. Schunicht, R. K. Fenton, S. J. Hannon, A. R. Vogstad, G. K. Jim, and C. W. Booker. 2012. Evaluation of the diagnostic and prognostic utility of ultrasonography at first diagnosis of presumptive bovine respiratory disease. *Can J Vet Res* 76(1):23-32.

Ackermann, M. R., R. Derscheid, and J. A. Roth. 2010. Innate immunology of bovine respiratory disease. *Vet Clin North Am Food Anim Pract* 26(2):215.

Adrion, C. and U. Mansmann. 2012. Bayesian model selection techniques as decision support for shaping a statistical analysis plan of a clinical trial: an example from a vertigo phase III study with longitudinal count data as primary endpoint. *BMC Med Res Methodol* 12(1):137.

Altman, D. G. 1991. *Practical Statistics for Medical Research*. Ed. Chapman and Hall, London, England.

Aly, S. S., W. J. Love, D. R. Williams, T. W. Lehenbauer, A. Van Eenennaam, C. Drake, P. H. Kass, and T. B. Farver. 2014. Agreement between bovine respiratory disease scoring systems for pre-weaned dairy calves. *Anim Health Res Rev* 15(2):148-150.

Ames, T. R., R. Markham, J. Opuda-Asibo, J. Leininger, and S. Maheswaran. 1985. Pulmonary response to intratracheal challenge with *Pasteurella haemolytica* and *Pasteurella multocida*. *Can J Comp Med* 49(4):395.

Amrine, D. E., B. J. White, R. Larson, D. E. Anderson, D. A. Mosier, and N. Cernicchiaro. 2013. Precision and accuracy of clinical illness scores, compared with pulmonary consolidation

scores, in Holstein calves with experimentally induced *Mycoplasma bovis* pneumonia. *Am J Vet Res* 74(2):310-315.

APHIS, U. 2018. Dairy 2014: Health and Management Practices on US Dairy Operations, 2014. Published by USDA, Fort Collins, CO USA: Report 3:62-77.

Apley, M. 2014. The clinical syndrome of BRD: what it is and what it is not. *Anim Health Res Rev* 15(2):135-137.

Apley, M. D. 2015. Treatment of calves with bovine respiratory disease: duration of therapy and posttreatment intervals. *Vet Clin North Am Food Anim Pract* 31(3):441-453.

Arnold-Lehna, D., M. Venner, L. J. Berghaus, R. Berghaus, and S. Giguère. 2020. Changing policy to treat foals with *Rhodococcus equi* pneumonia in the later course of disease decreases antimicrobial usage without increasing mortality rate. *Equine Vet J* 52(4):531-537.

Balls, M., 1999. The funding of research on the three rs in the EU. *ALTEX* 16 (4), 282–284.

Babkine, M. and L. Blond. 2009. Ultrasonography of the bovine respiratory system and its practical application. *Vet Clin North Am Food Anim Pract* 25(3):633-649.

Bach, A., C. Tejero, and J. Ahedo. 2011. Effects of group composition on the incidence of respiratory afflictions in group-housed calves after weaning. *J Dairy Sci* 94(4):2001-2006.

Bahler, C., A. Steiner, A. Luginbuhl, A. Ewy, H. Posthaus, D. Strabel, T. Kaufmann, and G. Regula. 2012. Risk factors for death and unwanted early slaughter in Swiss veal calves kept at a specific animal welfare standard. *Res Vet Sci* 92(1):162-168.

Ballou, M., D. Hanson, C. Cobb, B. Obeidat, M. Sellers, A. Pepper-Yowell, J. Carroll, T. Earleywine, and S. Lawhon. 2015. Plane of nutrition influences the performance, innate leukocyte responses, and resistance to an oral *Salmonella enterica* serotype Typhimurium challenge in Jersey calves. *J Dairy Sci* 98(3):1972-1982.

Baptiste, K. E. and N. C. Kyvsgaard. 2017. Do antimicrobial mass medications work? A systematic review and meta-analysis of randomised clinical trials investigating antimicrobial prophylaxis or metaphylaxis against naturally occurring bovine respiratory disease. *Pathog Dis* 75(7).

Barkema, H. W., M. A. von Keyserlingk, J. P. Kastelic, T. Lam, C. Luby, J.-P. Roy, S. J. LeBlanc, G. P. Keefe, and D. F. Kelton. 2015. Invited review: Changes in the dairy industry affecting dairy cattle health and welfare. *J Dairy Sci* 98(11):7426-7445.

Barone, 2001. Appareil respiratoire. p. 743–5 In: Splanchnologie II: Anatomie comparée des mammifères domestiques (Tome 3). 3ème édition, ed: Vigog, Paris, France.

Baruch, J., N. Cernicchiaro, C. A. Cull, K. F. Lechtenberg, J. S. Nickell, and D. G. Renter. 2019. Performance of multiple diagnostic methods in assessing the progression of bovine respiratory disease in calves challenged with infectious bovine rhinotracheitis virus and *Mannheimia haemolytica*. *J Anim Sci* 97(6):2357-2367.

Bassel, L. L., S. Tabatabaei, and J. L. Caswell. 2020. Host tolerance to infection with the bacteria that cause bovine respiratory disease. *Vet Clin North Am Food Anim Pract* 36(2):349-359.

Bello, G. and P. Blanco. 2019. Lung ultrasonography for assessing lung aeration in acute respiratory distress syndrome: a narrative review. *J Ultrasound Med* 38(1):27-37.

Berge, A. C. B., D. A. Moore, and W. M. Sischo. 2006. Field trial evaluating the influence of prophylactic and therapeutic antimicrobial administration on antimicrobial resistance of fecal *Escherichia coli* in dairy calves. *Appl Environ Microbiol* 72(6):3872-3878.

Berkvens, D., Speybroeck, N., Praet, N., Adel, A., Lesaffre, E., 2006. Estimating disease prevalence in a Bayesian framework using probabilistic constraints. *Epidemiol* 17 (2), 145–153.

Berman, J., D. Francoz, A. Abdallah, S. Dufour, and S. Buczinski. 2021a. Evaluation of inter-rater agreement of the clinical signs used to diagnose bovine respiratory disease in individually housed veal calves. *J Dairy Sci* 104(11):12053-65.

Berman, J., D. Francoz, J. Dubuc, and S. Buczinski. 2017. A randomised clinical trial of a metaphylactic treatment with tildipirosin for bovine respiratory disease in veal calves. *BMC Vet Res* 13(1):176.

Berman, J., D. Francoz, S. Dufour, and S. Buczinski. 2019. Bayesian estimation of sensitivity and specificity of systematic thoracic ultrasound exam for diagnosis of bovine respiratory disease in pre-weaned calves. *Prev Vet Med* 162:38-45.

Berman J, Masseau I, Fecteau G, Buczinski, S. and Francoz, D. 2020. 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* 184:105-153.

Berman, J., Masseau, I., Fecteau, G., Buczinski, S. and Francoz, D., 2021b. Comparison of thoracic ultrasonography and thoracic radiography to detect active infectious bronchopneumonia in hospitalized dairy calves. *J Vet Int Med* 35(4):2058-2068.

Binversie, E., P. Ruegg, D. Combs, and T. Ollivett. 2020. Randomized clinical trial to assess the effect of antibiotic therapy on health and growth of preweaned dairy calves diagnosed with respiratory disease using respiratory scoring and lung ultrasound. *J Dairy Sci* 103(12):11723-11735.

Bleul, U. 2009. Respiratory distress syndrome in calves. *Vet Clin North Am Food Anim Pract* 25(1):179-193, vii.

Bokkers, E. and P. Koene. 2001. Activity, oral behaviour and slaughter data as welfare indicators in veal calves: a comparison of three housing systems. *Appl Anim Behav Sci* 75(1):1-15.

Booker, C. W. and B. V. Lubbers. 2020. Bovine Respiratory Disease Treatment Failure: Impact and Potential Causes. *Vet Clin North Am Food Anim Pract* 36(2):487-496.

Bos, M. E., F. J. Taverne, I. M. van Geijlswijk, J. W. Mouton, D. J. Mevius, D. J. Heederik, and N. V. M. Authority. 2013. Consumption of antimicrobials in pigs, veal calves, and broilers in the Netherlands: quantitative results of nationwide collection of data in 2011. *PLoS One* 8(10):e77525.

Bosso, J. and R. H. Drew. 2011. Application of antimicrobial stewardship to optimise management of community acquired pneumonia. *Int J Clin Pract* 65(7):775-783.

Bouhemad, B., Z.-H. Liu, C. Arbelot, M. Zhang, F. Ferarri, M. Le-Guen, M. Girard, Q. Lu, and J.-J. Rouby. 2010. Ultrasound assessment of antibiotic-induced pulmonary reaeration in ventilator-associated pneumonia. *Crit Care Med* 38(1):84.

Bourcier, J. E., J. Paquet, M. Seinger, E. Gallard, J. P. Redonnet, F. Cheddadi, D. Garnier, J. M. Bourgeois, and T. Geeraerts. 2014. Performance comparison of lung ultrasound and chest x-ray for the diagnosis of pneumonia in the ED. *Am J Emerg Med* 32(2):115-118.

Branscum, A. J., I. A. Gardner, and W. O. Johnson. 2005. Estimation of diagnostic-test sensitivity and specificity through Bayesian modeling. *Prev Vet Med* 68(2-4):145-163.

Braun, U., M. Flückiger, D. Sicher, and D. Theil. 1995. Suppurative pleuropneumonia and a pulmonary abscess in a ram: ultrasonographic and radiographic findings. *Schweiz Archiv Tierheilkd* 137(6):272-278.

Braun, U., C. Gerspach, and C. Brammertz. 2018. The frequency of abnormal ultrasonographic findings in the lungs of 129 calves with bronchopneumonia. *Schweiz Archiv Tierheilkd* 160(12):737-741.

Braun, U., N. Pusterla, and M. Fluckiger. 1997. Ultrasonographic findings in cattle with pleuropneumonia. *Vet Rec* 141(1):12-17.

Bringhenti, L., M. Pallu, J. Silva, T. Tomazi, A. C. Tomazi, M. X. Rodrigues, L. M. Duarte, T. R. Bilby, and R. C. Bicalho. 2021. Effect of metaphylactic administration of tildipirosin on the incidence of pneumonia and otitis and on the upper respiratory tract and fecal microbiome of preweaning Holstein calves. *J Dairy Sci* 104(5):6020-6038.

Brooks, S.P., Gelman, A., 1998. General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat* 7 (4), 434–455.

Brown, L. D., T. T. Cai, and A. DasGupta. 2001. Interval estimation for a binomial proportion. *Stat Sci* 101-117.

Brcsic, M., H. Leruste, L. Heutinck, E. Bokkers, M. Wolthuis-Fillerup, N. Stockhofe, F. Gottardo, B. Lensink, G. Cozzi, and C. Van Reenen. 2012. Prevalence of respiratory disorders in veal calves and potential risk factors. *J Dairy Sci* 95(5):2753-2764.

Buczinski, S., D. Achard, and E. Timsit. 2021a. Effects of calfhood respiratory disease on health and performance of dairy cattle: A systematic review and meta-analysis. *J Dairy Sci* 104(7):8214-8227.

Buczinski, S., M. E. Borris, and J. Dubuc. 2018a. Herd-level prevalence of the ultrasonographic lung lesions associated with bovine respiratory disease and related environmental risk factors. *J Dairy Sci* 101(3):2423-2432.

Buczinski, S., C. Buathier, A. M. Belanger, H. Michaux, N. Tison, and E. Timsit. 2018b. Inter-rater agreement and reliability of thoracic ultrasonographic findings in feedlot calves, with or without naturally occurring bronchopneumonia. *J Vet Intern Med* 32(5):1787-1792.

Buczinski, S. and N. Dendukuri. 2015a. 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* 119(3-4):227-231.

Buczinski, S., C. Faure, S. Jolivet, and A. Abdallah. 2016a. Evaluation of inter-observer agreement when using a clinical respiratory scoring system in pre-weaned dairy calves. *N Z Vet J* 64(4):243-247.

Buczinski, S., G. Fecteau, L. Blouin, and M. Villettaz-Robichaud. 2021b. Factors affecting dairy calf price in auction markets in Québec, Canada: 2008–2019. *J Dairy Sci* 104(4):4635-4649.

Buczinski, S., G. Fecteau, J. Dubuc, and D. Francoz. 2018c. Validation of a clinical scoring system for bovine respiratory disease complex diagnosis in preweaned dairy calves using a Bayesian framework. *Prev Vet Med* 156:102-112.

Buczinski, S., G. Forte, and A. M. Belanger. 2013. 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* 96(7):4523-4528.

Buczinski, S., G. Forte, D. Francoz, and A. M. Belanger. 2014. Comparison of thoracic auscultation, clinical score, and ultrasonography as indicators of bovine respiratory disease in preweaned dairy calves. *J Vet Intern Med* 28(1):234-242.

Buczinski, S., J. Menard, and E. Timsit. 2016b. Incremental Value (Bayesian Framework) of Thoracic Ultrasonography over Thoracic Auscultation for Diagnosis of Bronchopneumonia in Preweaned Dairy Calves. *J Vet Intern Med* 30(4):1396-1401.

Buczinski, S. and A. M. O'Connor. 2016c. Specific Challenges in Conducting and Reporting Studies on the Diagnostic Accuracy of Ultrasonography in Bovine Medicine. *Vet Clin North Am Food Anim Pract* 32(1):1-18.

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(2):399-423.

Buczinski, S., R. Rademacher, H. Tripp, M. Edmonds, E. Johnson, and S. Dufour. 2015b. Assessment of l-lactatemia as a predictor of respiratory disease recognition and severity in feedlot steers. *Prev Vet Med* 118(4):306-318.

Burn, C. C. and A. A. Weir. 2011. Using prevalence indices to aid interpretation and comparison of agreement ratings between two or more observers. *Vet J* 188(2):166-170.

Burow, E., C. Simoneit, B.-A. Tenhagen, and A. Käsbohrer. 2014. Oral antimicrobials increase antimicrobial resistance in porcine *E. coli*—A systematic review. *Prev Vet Med* 113(4):364-375.

Byrt, T., J. Bishop, and J. B. Carlin. 1993. Bias, prevalence and kappa. *J Clin Epidemiol* 46(5):423-429.

Callan, R. J. and F. B. Garry. 2002. Biosecurity and bovine respiratory disease. *Vet Clin North Am Food Anim Pract* 18(1):57-77.

Carpentier, L., D. Berckmans, A. Youssef, D. Berckmans, T. van Waterschoot, D. Johnston, N. Ferguson, B. Earley, I. Fontana, and E. Tullo. 2018. Automatic cough detection for bovine respiratory disease in a calf house. *biosyst Eng* 173:45-56.

Caswell, J. 2014. Failure of respiratory defenses in the pathogenesis of bacterial pneumonia of cattle. *Vet Pathol* 51(2):393-409.

Catry, B., J. Dewulf, T. Goffin, A. Decostere, F. Haesebrouck, and A. De Kruif. 2007. Antimicrobial resistance patterns of *Escherichia coli* through the digestive tract of veal calves. *Microb Drug Resist* 13(2):147-150.

Catry, B., J. Dewulf, D. Maes, B. Pardon, B. Callens, M. Vanrobaeys, G. Opsomer, A. de Kruif, and F. Haesebrouck. 2016. Effect of antimicrobial consumption and production type on antibacterial resistance in the bovine respiratory and digestive tract. *PLoS One* 11(1):e0146488.

Catry, B., F. Haesebrouck, S. D. Vlieghe, B. Feyen, M. Vanrobaeys, G. Opsomer, S. Schwarz, and A. D. Kruif. 2005. Variability in acquired resistance of *Pasteurella* and *Mannheimia* isolates from the nasopharynx of calves, with particular reference to different herd types. *Microb Drug Resist* 11(4):387-394.

Celestino, M. L., L. Fernandes, P. R. Menta, D. Paiva, T. L. Ribeiro, T. Silva, T. R. Bilby, R. C. Neves, M. A. Ballou, and V. S. Machado. 2020. The effect of metaphylactic use of tildipirosin for the control of respiratory disease in long-distance transported dairy calves. *Front Vet Sci* 7:632.

Chavez, M. A., N. Shams, L. E. Ellington, N. Naithani, R. H. Gilman, M. C. Steinhoff, M. Santosham, R. E. Black, C. Price, M. Gross, and W. Checkley. 2014. Lung ultrasound for the diagnosis of pneumonia in adults: a systematic review and meta-analysis. *Respir Res* 15:50.

Cheung, A., S. Dufour, G. Jones, P. Kostoulas, M. Stevenson, N. Singanallur, and S. Firestone. 2021. Bayesian latent class analysis when the reference test is imperfect. *Rev Sci Tech (Off Int Epizoot)* 40(1):271-286.

Christ-Crain, M., D. Stolz, R. Bingisser, C. Muller, D. Miedinger, P. R. Huber, W. Zimmerli, S. Harbarth, M. Tamm, and B. Muller. 2006. Procalcitonin guidance of antibiotic therapy in community-acquired pneumonia: a randomized trial. *Am J Respir Crit Care Med* 174(1):84-93.

Christensen, J. and I. A. Gardner. 2000. Herd-level interpretation of test results for epidemiologic studies of animal diseases. *Prev Vet Med* 45(1-2):83-106.

Coetzee, J. F., D. R. Magstadt, P. K. Sidhu, L. Follett, A. M. Schuler, A. C. Krull, V. L. Cooper, T. J. Engelken, M. D. Kleinhenz, and A. M. O'Connor. 2019. Association between antimicrobial drug class for treatment and retreatment of bovine respiratory disease (BRD) and frequency of resistant BRD pathogen isolation from veterinary diagnostic laboratory samples. *PloS one* 14(12):e0219104.

Cohen, J. 1960. A Coefficient of Agreement for Nominal Scales *Educ Psychol Meas.* SAGE Publications Inc 20:37-46.

Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, Irwig L, Levine D, Reitsma JB, De Vet HC, Bossuyt PM. 2016. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ* 6(11):e012799.

Cook, A., R. J. Reid-Smith, R. J. Irwin, S. A. McEwen, V. Young, and C. Ribble. 2011. Antimicrobial resistance in *Campylobacter*, *Salmonella*, and *Escherichia coli* isolated from retail grain-fed veal meat from southern Ontario, Canada. *J Food Prot* 74(8):1245-1251.

Correa, M. T., C. R. Curtis, H. N. Erb, and M. E. White. 1988. Effect of calfhood morbidity on age at first calving in New York Holstein herds. *Prev Vet Med* 6(4):253-262.

Cramer, M. and T. Ollivett. 2019. Growth of preweaned, group-housed dairy calves diagnosed with respiratory disease using clinical respiratory scoring and thoracic ultrasound—A cohort study. *J Dairy Sci* 102(5):4322-4331.

Cramer, C. and T. L. Ollivett. 2020. Behavior assessment and applications for BRD diagnosis: preweaned dairy calves. *Anim Health Res Rev* 21(2):188-191.

Crawley, J., C. Prospero, H. C. Baggett, W. A. Brooks, M. Deloria Knoll, L. L. Hammitt, S. R. C. Howie, K. L. Kotloff, O. S. Levine, S. A. Madhi, D. R. Murdoch, K. L. O'Brien, D. M. Thea, J. O. Awori, C. Bunthi, A. N. DeLuca, A. J. Driscoll, B. E. Ebruke, D. Goswami, M. M. Hidgon, R. A. Karron, S. Kazungu, N. Kourouma, G. Mackenzie, D. P. Moore, A. Mudau, M. Mwale, K. Nahar, D. E. Park, B. Piralam, P. Seidenberg, M. Sylla, D. R. Feikin, J. A. G. Scott, and P. S. Group. 2017. Standardization of Clinical Assessment and Sample Collection Across All PERCH Study Sites. *Clin Infect Dis* 64(suppl_3):S228-S237.

Cuevas-Gómez, I., M. McGee, J. M. Sánchez, E. O'Riordan, N. Byrne, T. McDanel, 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(1):1-9.

Curtis, R., L. Viel, S. McGuirk, O. Radostits, and F. Harris. 1986. Lung sounds in cattle, horses, sheep and goats. *Can Vet J* 27(4):170.

Dagleish, M., J. Finlayson, C. Bayne, S. MacDonald, J. Sales, and J. Hodgson. 2010. Characterization and time course of pulmonary lesions in calves after intratracheal infection with *Pasteurella multocida* A: 3. *J Comp Pathol* 142(2-3):157-169.

DeDonder, K. D. 2008. Lung auscultation as a predictor of lung lesions and bovine respiratory disease outcome in feed yard cattle. PhD dissertation. Kansas State University.

DeDonder, K. D. and M. D. Apley. 2015. A review of the expected effects of antimicrobials in bovine respiratory disease treatment and control using outcomes from published randomized clinical trials with negative controls. *Vet Clin North Am Food Anim Pract* 31(1):97-111.

De Oliveira, B. A. F. D., N. C. Gaeta, B. L. M. Ribeiro, M. A. R. Alemán, L. M. Marques, J. Timenetsky, P. A. Melville, J. A. Marques, V. Marvulle, and L. Gregory. 2016. Determination of bacterial aetiologic factor on tracheobronchial lavage in relation to clinical signs of bovine respiratory disease. *J Med Microbiol* 65(10):1137-1142.

Delabougliuse, A., A. James, J.-F. Valarcher, S. Hagglund, D. Raboisson, and J. Rushton. 2017. Linking disease epidemiology and livestock productivity: The case of bovine respiratory disease in France. *PloS one* 12(12):e0189090.

Dendukuri, N., A. Hadgu, and L. Wang. 2009. Modeling conditional dependence between diagnostic tests: a multiple latent variable model. *Stat Med* 28(3):441-461.

Dendukuri N, Joseph L. 2001. Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. *Biometrics*. 57(1):158-67.

Diaz F. 2014. OIE Standard on principles and methods of validation of diagnostic assays for infectious diseases. In: *Proceedings of the OIE Regional Workshop for OIE National Focal Points for Veterinary Products, Tokyo, Japan 2014 Dec 3* (pp. 3-5).

Di Labio, E., G. Regula, A. Steiner, R. Miserez, A. Thomann, and U. Ledergerber. 2007. Antimicrobial resistance in bacteria from Swiss veal calves at slaughter. *Zoonoses and public health* 54(9-10):344-352.

Diesel, D., J. Lebel, and A. Tucker. 1991. Pulmonary particle deposition and airway mucociliary clearance in cold-exposed calves. *Am J Vet Res* 52(10):1665-1671.

Dohoo, I. R., S. W. Martin, H. Stryhn, and I. R. Dohoo. 2012. Screening and diagnostic tests. p91- 134 In: *Veterinary Epidemiologic Research*. Second edition. ed. I. R. Dohoo, ed. VER Inc., Charlottetown, Prince Edward Island, Canada.

Donner, A. and M. Eliasziw. 1987. Sample size requirements for reliability studies. *Stat Med* 6(4):441-448.

Donovan, G. A., I. R. Dohoo, D. M. Montgomery, and F. L. Bennett. 1998. Calf and disease factors affecting growth in female Holstein calves in Florida, USA. *Prev Vet Med* 33(1-4):1-10.

Doyle, D., B. Credille, T. Lehenbauer, R. Berghaus, S. Aly, J. Champagne, P. Blanchard, B. Crossley, L. Berghaus, and S. Cochran. 2017. Agreement among 4 sampling methods to identify respiratory pathogens in dairy calves with acute bovine respiratory disease. *J Vet Int Med* 31(3):954-959.

Dubrovsky, S., A. Van Eenennaam, B. Karle, P. Rossitto, T. W. Lehenbauer, and S. S. Aly. 2019. Epidemiology of bovine respiratory disease (BRD) in preweaned calves on California dairies: The BRD 10K study. *J Dairy Sci* 102(8):7306-7319.

Dunn, T. R., T. L. Ollivett, D. L. Renaud, K. E. Leslie, S. J. LeBlanc, T. F. Duffield, and D. F. Kelton. 2018. The effect of lung consolidation, as determined by ultrasonography, on first-lactation milk production in Holstein dairy calves. *J Dairy Sci*. 101,1-7.

Dunson, D. B. 2001. Commentary: practical advantages of Bayesian analysis of epidemiologic data. *Am J Epidemiol* 153(12):1222-1226.

Earley, B., K. B. Sporer, and S. Gupta. 2017. Invited review: Relationship between cattle transport, immunity and respiratory disease. *Animal* 11(3):486-492.

Earley, B., A. Arguello, E. O’Riordan, P. Crosson, A. Cappelleri, and M. McGee. 2019. Antimicrobial drug usage from birth to 180 days of age in Irish dairy calves and in suckler beef calves. *J Appl Anim Res* 47(1):474-485.

Farrow, C. S. 1999. Bovine pneumonia: Its radiographic appearance. *Vet Clin North Am Food Anim Pract* 15(2):301-358.

Fertner, M., N. Toft, H. L. Martin, and A. Boklund. 2016. A register-based study of the antimicrobial usage in Danish veal calves and young bulls. *Prev Vet Med* 131:41-47.

Fleiss, J. L., J. Cohen, and B. S. Everitt. 1969. Large sample standard errors of kappa and weighted kappa. *Psychol Bull* 72(5):323.

Fleiss, J. L., B. Levin, and M. C. Paik. 1981. The measurement of interrater agreement. *Stat methods for rates and proportions* 2(212-236):22-23.

Flock, M. 2004. Diagnostic ultrasonography in cattle with thoracic disease. *Vet J* 167(3):272-280.

Fowler, J., S. M. Stieger-Vanegas, J. A. Vanegas, G. Bobe, and K. P. Poulsen. 2017. Comparison of Thoracic Radiography and Computed Tomography in Calves with Naturally Occurring Respiratory Disease. *Front Vet Sci* 4:101.

Francoz, D., G. Fecteau, A. Desrochers, and M. Fortin. 2004. Otitis media in dairy calves: a retrospective study of 15 cases (1987 to 2002). *Can Vet J* 45(8):661.

Fulton, R. W., C. Purdy, A. W. Confer, J. Saliki, R. W. Loan, R. E. Briggs, and L. J. Burge. 2000. Bovine viral diarrhoea viral infections in feeder calves with respiratory disease: interactions with *Pasteurella* spp., parainfluenza-3 virus, and bovine respiratory syncytial virus. *Can J Vet Res* 64(3):151.

Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian anal* 1(3):515-534.

Genway Biotech INC. 2017. Bovine Haptoglobin ELISA Kit. <https://www.genwaybio.com/bovine-haptoglobin>

Godinho, K. S., P. Sarasola, E. Renoult, N. Tilt, S. Keane, G. D. Windsor, T. G. Rowan, and S. J. Sunderland. 2007. Use of deep nasopharyngeal swabs as a predictive diagnostic method for natural respiratory infections in calves. *Vet Rec* 160(1):22-25.

Gold, R., Reichman, M., Greenberg, E., Ivanidze, J., Elias, E., Tsiouris, A.J., Comunale, J. P., Johnson, C.E., Sanelli, P.C., 2010. Developing a new reference standard: is validation necessary? *Acad. Radiol.* 17 (9), 1079–1082.

Gomez, D. E., L. G. Arroyo, Z. Poljak, L. Viel, and J. S. Weese. 2017. Implementation of an algorithm for selection of antimicrobial therapy for diarrhoeic calves: Impact on antimicrobial treatment rates, health and faecal microbiota. *Vet J* 226:15-25.

Gosselin, V. B., D. Francoz, M. Babkine, A. Desrochers, S. Nichols, E. Doré, C. Bédard, J. Parent, J.-H. Fairbrother, and G. Fecteau. 2012. A retrospective study of 29 cases of otitis media/interna in dairy calves. *Can Vet J* 53(9):957.

Gorden, P. J. and P. Plummer. 2010. Control, management, and prevention of bovine respiratory disease in dairy calves and cows. *Vet Clin North Am Food Anim Pract* 26(2):243-259.

Graveland, H., J. A. Wagenaar, H. Heesterbeek, D. Mevius, E. Van Duinkerken, and D. Heederik. 2010. Methicillin resistant *Staphylococcus aureus* ST398 in veal calf farming: human MRSA carriage related with animal antimicrobial usage and farm hygiene. *PloS one* 5(6):e10990.

Greiner, M. 1996. Two-graph receiver operating characteristic (TG-ROC): update version supports optimisation of cut-off values that minimise overall misclassification costs. *J Immunol Methods* 191(1):93-94.

Griffin, D., M. Chengappa, J. Kuszak, and D. S. McVey. 2010. Bacterial pathogens of the bovine respiratory disease complex. *Vet Clin North Am Food Anim Pract* 26(2):381-394.

Gustafson, P. 2005. On model expansion, model contraction, identifiability and prior information: two illustrative scenarios involving mismeasured variables. *Stat Sci* 20(2):111-140.

Guyton, A. a. H., JE. 2011. Pathophysiology of specific pulmonary abnormalities. p 518-519 In: *Textbook of medical physiology*. 12th Ed. Saunders, ed. Elsevier, Philadelphia, USA.

Gwet, K. L. 2008. Computing inter-rater reliability and its variance in the presence of high agreement. *Br J Math Stat Psychol* 61(1):29-48.

Gwet, K. L. 2014. Agreement coefficients for nominal ratings: a review. p27-73 In: *Handbook of inter-rater reliability: the definitive guide to measuring the extent of agreement among raters*. 4th ed. K. L. Gwet, ed. Advances analytics, LLC, Gaithersburg, USA.

Hayes, G., K. Mathews, S. Kruth, G. Doig, and C. Dewey. 2010. Illness severity scores in veterinary medicine: what can we learn? *J Vet Intern Med* 24(3):457-466.

Hilden, J. 2014. Commentary: On NRI, IDI, and “good-looking” statistics with nothing underneath. *Epidemiol* 25(2):265-267.

Hu, Q.J., Shen, Y.C., Jia, L.Q., Guo, S.J., Long, H.Y., Pang, C.S., Wen, F.Q., 2014. Diagnostic performance of lung ultrasound in the diagnosis of pneumonia: a bivariate meta-analysis. *Int. J. Clin. Exp. Med.* 7 (1), 115.

Hubbuch, A., R. Peter, B. Willi, S. Hartnack, C. Müntener, H. Naegeli, and C. Gerspach. 2021. Comparison of antimicrobial prescription patterns in calves in Switzerland before and after the launch of online guidelines for prudent antimicrobial use. *BMC Vet Res* 17(1):1-14.

Hussein, H. A., C. Binici, and R. Staufenbiel. 2018. Comparative evaluation of ultrasonography with clinical respiratory score in diagnosis and prognosis of respiratory diseases in weaned dairy buffalo and cattle calves. *J Anim Sci Technol* 60:29.

Iorio, G., M. Capasso, G. De Luca, S. Prisco, C. Mancusi, B. Lagana, and V. Comune. 2015. Lung ultrasound in the diagnosis of pneumonia in children: proposal for a new diagnostic algorithm. *PeerJ* 3:e1374.

Jarrige, N., G. Cazeau, E. Morignat, M. Chantepedrix, and E. Gay. 2017. Quantitative and qualitative analysis of antimicrobial usage in white veal calves in France. *Prev Vet Med* 144:158-166.

Jericho, K., A. Lejeune, and G. Tiffin. 1986. Bovine herpesvirus-1 and *Pasteurella haemolytica* aerobiology in experimentally infected calves. *Am J Vet Res* 47(2):205-209.

Jones, C. and A. Webster. 1984. Relationships between counts of nasopharyngeal bacteria, temperature, humidity and lung lesions in veal calves. *Res Vet Sci* 37(2):132-137.

Jones, C. 1987. Proliferation of *Pasteurella haemolytica* in the calf respiratory tract after an abrupt change in climate. *Res Vet Sci* 42(2):179-186.

Jones, G. F., D. A. Feeney, and C. Mews. 1998. Comparison of radiographic and necropsy findings of lung lesions in calves after challenge exposure with *Pasteurella multocida*. *Am J Vet Res* 59(9):1108-1112.

Jourquin, S., J. Bokma, L. De Cremer, K. van Leenen, N. Vereecke, and B. Pardon. 2022. Randomized field trial comparing the efficacy of florfenicol and oxytetracycline in a natural outbreak of calf pneumonia using lung reaeration as a cure criterion. *J Vet Int Med* 36(2):820-828.

Jung, C. and H. Bostedt. 2004. Thoracic Ultrasonography Technique in Newborn Calves and Description of Normal and Pathological Findings. *Vet Radiol Ultrasound* 45(4):331-335.

Kaneene, J. B. and H. S. Hurd. 1990. The national animal health monitoring system in Michigan. III. Cost estimates of selected dairy cattle diseases. *Prev Vet Med* 8(2-3):127-140.

Khan, M., D. Weary, and M. Von Keyserlingk. 2011. Invited review: Effects of milk ration on solid feed intake, weaning, and performance in dairy heifers. *J Dairy Sci* 94(3):1071-1081.

Kim, S., Lee, W., 2017. Does McNemar's test compare the sensitivities and specificities of two diagnostic tests? *Stat Methods Med Res* 26, 142-154.

Kim, N. N. and D. Marikar. 2019. Antibiotic prescribing for upper respiratory tract infections: NICE guidelines. *Archives of Disease in Childhood-Education and Practice*.

Kostoulas, P., S. S. Nielsen, A. J. Branscum, W. O. Johnson, N. Dendukuri, N. K. Dhand, N. Toft, and I. A. Gardner. 2017. STARD-BLCM: Standards for the Reporting of Diagnostic accuracy studies that use Bayesian Latent Class Models. *Prev Vet Med* 138:37-47.

Kottner, J., L. Audige, S. Brorson, A. Donner, B. J. Gajewski, A. Hrobjartsson, C. Roberts, M. Shoukri, and D. L. Streiner. 2011. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol* 64(1):96-106.

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(10):4014-4025.

Lambert, P. C., A. J. Sutton, P. R. Burton, K. R. Abrams, and D. R. Jones. 2005. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. *Stat Med* 24(15):2401-2428.

Landis, J. R. and G. G. Koch. 1977. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics* 33(2):363-374.

Lava, M., B. Pardon, G. Schupbach-Regula, K. Keckeis, P. Deprez, A. Steiner, and M. Meylan. 2016a. Effect of calf purchase and other herd-level risk factors on mortality, unwanted early slaughter, and use of antimicrobial group treatments in Swiss veal calf operations. *Prev Vet Med* 126:81-88.

Lava, M., G. Schupbach-Regula, A. Steiner, and M. Meylan. 2016b. Antimicrobial drug use and risk factors associated with treatment incidence and mortality in Swiss veal calves reared under improved welfare conditions. *Prev Vet Med* 126:121-130.

Leruste, H., M. Brscic, L. F. Heutinck, E. K. Visser, M. Wolthuis-Fillerup, E. A. Bokkers, N. Stockhofe-Zurwieden, G. Cozzi, F. Gottardo, B. J. Lensink, and C. G. van Reenen. 2012. The relationship between clinical signs of respiratory system disorders and lung lesions at slaughter in veal calves. *Prev Vet Med* 105(1-2):93-100.

Lillie, L. E. 1974. The bovine respiratory disease complex. *Can Vet J* 15(9):233-242.

Lippi, G., T. Meschi, and G. Cervellin. 2011. Inflammatory biomarkers for the diagnosis, monitoring and follow-up of community-acquired pneumonia: clinical evidence and perspectives. *Eur J Int Med* 22(5):460-465.

Liu, J., F. Chen, H. Yu, P. Zeng, and L. Liu. 2014. A two-stage Bayesian method for estimating accuracy and disease prevalence for two dependent dichotomous screening tests when the status of individuals who are negative on both tests is unverified. *BMC Med Res Methodol* 14:110.

Losinger, W. C. and A. J. Heinrichs. 1996. Management variables associated with high mortality rates attributable to respiratory tract problems in female calves prior to weaning. *JAVMA* 209(10):1756-1759.

Lombard, J., F. Garry, S. Tomlinson, and L. Garber. 2007. Impacts of dystocia on health and survival of dairy calves. *J Dairy Sci* 90(4):1751-1760.

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.

Love, W. J., T. W. Lehenbauer, A. L. Van Eenennaam, C. M. Drake, P. H. Kass, T. B. Farver, and S. S. Aly. 2016. 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* 28(2):119-128.

Lubbers, B. V., M. D. Apley, J. F. Coetzee, D. A. Mosier, D. S. Biller, D. E. Mason, and P. N. Henao-Guerrero. 2007. Use of computed tomography to evaluate pathologic changes in the lungs of calves with experimentally induced respiratory tract disease. *Am J Vet Res* 68(11):1259-1264.

Lundborg, G., E. Svensson, and P. Oltenacu. 2005. Herd-level risk factors for infectious diseases in Swedish dairy calves aged 0–90 days. *Prev Vet Med* 68(2-4):123-143.

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(8):7329-7344.

Mang, A., S. Buczinski, C. Booker, and E. Timsit. 2015. Evaluation of a computer-aided lung auscultation system for diagnosis of bovine respiratory disease in feedlot cattle. *J Vet Int Med* 29(4):1112-1116.

Marcato, F., H. van den Brand, B. Kemp, B. Engel, M. Wolthuis-Fillerup, and K. van Reenen. 2020. Effects of pretransport diet, transport duration, and type of vehicle on physiological status of young veal calves. *J Dairy Sci* 103(4):3505-3520.

Marquou, S., L. Blouin, H. Djakite, R. Laplante, and S. Buczinski. 2019. Health parameters and their association with price in young calves sold at auction for veal operations in Québec, Canada. *J Dairy Sci* 102(7):6454-6465.

Massé, J., H. Lardé, J. M. Fairbrother, J.-P. Roy, D. Francoz, S. Dufour, and M. Archambault. 2021. Prevalence of antimicrobial resistance and characteristics of *Escherichia coli* isolates from fecal and manure pit samples on dairy farms in the province of Québec, Canada. *Front Vet Sci* 8:438.

Masseau, I., Fecteau, G., Breton, L., Helie, P., Beaugard, G., Blond, L., 2008. Radiographic detection of thoracic lesions in adult cows: a retrospective study of 42 cases (1995-2002). *Can. Vet. J.* 49, 261–267.

Masset, N., S. Assié, N. Herman, T. Jozan, and V. Herry. 2022. Ultrasonography of the cranial part of the thorax is a quick and sensitive technique to detect lung consolidation in veal calves. *Vet Med and Sci* 8(3):1229-1239.

McCorkell R, Wynne-Edwards K, Windeyer C, Schaefer A. 2014. Limited efficacy of Fever Tag® temperature sensing ear tags in calves with naturally occurring bovine respiratory disease or induced bovine viral diarrhoea virus infection. *Can Vet J* 55(7):688.

McGill, J. L. and R. E. Sacco. 2020. The immunology of bovine respiratory disease: recent advancements. *Veterinary Clinics of North America: Food Animal Practice* 36(2):333-348.

McGuirk, S. M. 2008. Disease management of dairy calves and heifers. *Vet Clin North Am Food Anim Pract* 24(1):139-153.

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(2):145-147.

McInturff, P., W. O. Johnson, D. Cowling, and I. A. Gardner. 2004. Modelling risk when binary outcomes are subject to error. *Stat Med* 23(7):1095-1109.

Mellor, D. and K. Stafford. 2004. Animal welfare implications of neonatal mortality and morbidity in farm animals. *Vet J* 168(2):118-133.

Miller, W. M., J. W. Harkness, M. S. Richards, and D. G. Pritchard. 1980. Epidemiological studies of calf respiratory disease in a large commercial veal unit. *Res Vet Sci* 28(3):267-274.

Mischel, L. J. 2019. Watch and learn? Using EDpuzzle to enhance the use of online videos. *Manag Teach Rev* 4(3):283-289.

Moons, K. G., D. G. Altman, J. B. Reitsma, J. P. Ioannidis, P. Macaskill, E. W. Steyerberg, A. J. Vickers, D. F. Ransohoff, and G. S. Collins. 2015. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 162(1):W1-73.

Moons, K., F. E. Harrell, and E. W. Steyerberg. 2002. Should scoring rules be based on odds ratios or regression coefficients? *J Clin Epidemiol* 55(10):1054-1055.

Morin, M., Dubuc J, Francoz D, Buczinski S. 2020. Randomized clinical trial of intranasal vaccination with Once® PMH IN. BOVILIS® Coronavirus Vaccine placebo on prevalence of lung consolidation diarrhea and health parameters in veal calves.

Morris, R., P. MacNeela, A. Scott, P. Treacy, A. Hyde, J. O'Brien, D. Lehwaldt, A. Byrne, and J. Drennan. 2008. Ambiguities and conflicting results: the limitations of the kappa statistic in establishing the interrater reliability of the Irish nursing minimum data set for mental health: a discussion paper. *Int J Nurs Stud* 45(4):645-647.

Murray, G. M., S. J. More, T. A. Clegg, B. Earley, R. G. O'Neill, D. Johnston, J. Gilmore, M. Nosov, M. C. McElroy, and T. J. Inzana. 2018. Risk factors associated with exposure to bovine respiratory disease pathogens during the peri-weaning period in dairy bull calves. *BMC Vet Res* 14(1):1-11.

Naylor, J. M., R. M. Streeter, and P. Torgerson. 2012. Factors affecting rectal temperature measurement using commonly available digital thermometers. *Res Vet Sci* 92(1):121-123.

Noffsinger, T., K. Brattain, G. Quakenbush, and G. Taylor. 2014. Field results from Whisper® stethoscope studies. *Anim Health Res Rev* 15(2):142-144.

O'Leary, S., M. Lund, T. J. Ytre-Hauge, S. R. Holm, K. Naess, L. N. Dalland, and S. M. McPhail. 2014. Pitfalls in the use of kappa when interpreting agreement between multiple raters in reliability studies. *Physiotherapy* 100(1):27-35.

Ohlerth, S., H. Augsburger, M. Abe, S. Ringer, L. Hatz, and U. Braun. 2014. Computed tomography of the thorax in calves from birth to 105 days of age. *Schweiz Arch Tierheilkd* 156(10):489-497.

Ollivett, T. L. 2014. Understanding the Diagnosis and Risk Factors for Respiratory Disease in Dairy Calves. PhD Dissertation. The university of Guelph.

Ollivett, T. L. and S. Buczinski. 2016. On-Farm Use of Ultrasonography for Bovine Respiratory Disease. *Vet Clin North Am Food Anim Pract* 32(1):19-35.

Ollivett, T., A. Burton, R. Bicalho, and D. Nydam. 2011. Use of rapid thoracic ultrasonography for detection of subclinical and clinical pneumonia in dairy calves. Pages 148-148 in *Proc. American Association of Bovine Practitioners Proceedings of the Annual Conference*.

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(6):1728-1734.

Ollivett, T., J. Hewson, R. Shubotz, and J. Caswell. 2013. Ultrasonographic progression of lung consolidation after experimental infection with *Mannheimia haemolytica* in Holstein bull calves. Pages 147-147 in *Proc. American Association of Bovine Practitioners Proceedings of the Annual Conference*.

Olson, A., W. M. Sisco, A. C. B. Berge, A. Adams-Progar, and D. A. Moore. 2019. A retrospective cohort study comparing dairy calf treatment decisions by farm personnel with veterinary observations of clinical signs. *J Dairy Sci* 102(7):6391-6403.

OMSA, 2019. Accès en ligne au code terrestre. https://www.woah.org/fr/ce-que-nous-faisons/normes/codes-et-manuels/acces-en-ligne-au-code-terrestre/?id=169&L=1&htmfile=chapitre_aw_introduction.htm

Osman, R., N. Malmuthuge, P. Gonzalez-Cano, and P. Griebel. 2018. Development and function of the mucosal immune system in the upper respiratory tract of neonatal calves. *Annu Rev Anim Biosci* 6:141-155.

Pancier, R. J. and A. W. Confer. 2010. Pathogenesis and pathology of bovine pneumonia. *Vet Clin North Am Food Anim Pract* 26(2):191-214.

Pardon, B. 2019a. A quick scan lung ultrasound method and flow chart as a decision aid for bronchopneumonia diagnosis. In *XXIV congreso internacional ANEMBE de medicina bovina 2019* (pp. 258-260). Asociación Nacional de Especialistas en Medicina Bovina de España (ANEMBE).

Pardon, B., J. Alliet, R. Boone, S. Roelandt, B. Valgaeren, and P. Deprez. 2015. Prediction of respiratory disease and diarrhea in veal calves based on immunoglobulin levels and the serostatus for respiratory pathogens measured at arrival. *Prev Vet Med* 120(2):169-176.

Pardon, B., B. Catry, R. Boone, H. Theys, K. De Bleecker, J. Dewulf, and P. Deprez. 2014. Characteristics and challenges of the modern Belgian veal industry. *Vlaams Diergeneeskd Tijdschr* 83(4):155-163.

Pardon, B., B. Catry, J. Dewulf, D. Persoons, M. Hostens, K. De Bleecker, and P. Deprez. 2012a. Prospective study on quantitative and qualitative antimicrobial and anti-inflammatory drug use in white veal calves. *J Antimicrob Chemother* 67(4):1027-1038.

Pardon, B., K. De Bleecker, M. Hostens, J. Callens, J. Dewulf, and P. Deprez. 2012b. Longitudinal study on morbidity and mortality in white veal calves in Belgium. *BMC Vet Res* 8:26.

Pardon, B., M. Hostens, L. Duchateau, J. Dewulf, K. De Bleecker, and P. Deprez. 2013. Impact of respiratory disease, diarrhea, otitis and arthritis on mortality and carcass traits in white veal calves. *BMC Vet Res* 9:79.

Pardon, B., S. Buczinski, and P. R. Deprez. 2019b. Accuracy and inter-rater reliability of lung auscultation by bovine practitioners when compared with ultrasonographic findings. *Vet Rec* 185(4):109.

Partlow, J., F. David, L. M. Hunt, F. Relave, L. Blond, M. Pinilla, and J. P. Lavoie. 2017. Comparison of Thoracic Ultrasonography and Radiography for the Detection of Induced Small Volume Pneumothorax in the Horse. *Vet Radiol Ultrasound* 58(3):354-360.

Pepe, M. S. 2003a. Incomplete data and imperfect reference tests. p168-213 In: *The statistical evaluation of medical tests for classification and prediction*. ed. M. S. Pepe, ed. Oxford University Press Inc, New York, New York, USA.

Pepe, M. S. 2003b. Measures of accuracy for binary tests. p14-34 In: *The statistical evaluation of medical tests for classification and prediction*. ed. M. S. Pepe, ed. Oxford University Press Inc, New York, New York, USA.

Pereyra, V. G., M. Pol, F. Pastorino, and A. Herrero. 2015. Quantification of antimicrobial usage in dairy cows and preweaned calves in Argentina. *Prev Vet Med* 122(3):273-279.

Perez, E., J. Noordhuizen, L. Van Wuijkhuis, and E. Stassen. 1990. Management factors related to calf morbidity and mortality rates. *Livest Prod Sci* 25(1-2):79-93.

Perino, L. J. and M. D. Apley. 1998. Clinical trial design in feedlots. *Vet Clin North Am Food Anim Pract* 14(2):343-365.

Pinzone, M. R., B. Cacopardo, L. Abbo, and G. Nunnari. 2014. Duration of antimicrobial therapy in community acquired pneumonia: less is more. *Sci World J* 2014.

Portillo, T. 2014. Pen riding and evaluation of cattle in pens to identify compromised individuals. Pages 5-8 in *Proc. Proceedings of the Forty-Seventh Annual Conference of the American Association of Bovine Practitioners*, Albuquerque, New Mexico, USA, 18-20 September 2014. American Association of Bovine Practitioners.

Pravettoni, D., S. Buczinski, G. Sala, V. Ferrulli, F. Bianchi, and A. Boccardo. 2021. Diagnostic accuracy of focused lung ultrasonography as a rapid method for the diagnosis of respiratory disease in dairy calves. *J Dairy Sci* 104(4):4929-4935.

Producteurs bovins du Québec, 2020. Coup d'oeil. <http://bovin.qc.ca/>

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

Rabeling, B., J. Rehage, D. Dopfer, and H. Scholz. 1998. Ultrasonographic findings in calves with respiratory disease. *Vet Rec* 143(17):468-471.

Radaelli, E., M. Luini, G. Loria, R. Nicholas, and E. Scanziani. 2008. Bacteriological, serological, pathological and immunohistochemical studies of *Mycoplasma bovis* respiratory infection in veal calves and adult cattle at slaughter. *Res Vet Sci* 85(2):282-290.

Rademacher, R. D., S. Buczinski, H. M. Tripp, M. D. Edmonds, and E. G. Johnson. 2014. Systematic thoracic ultrasonography in acute bovine respiratory disease of feedlot steers: impact of lung consolidation on diagnosis and prognosis in a case-control study. *Bov Pract* 48(1).

Reef, V. B., M. G. Boy, C. F. Reid, and A. Elser. 1991. 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* 198(12):2112-2118.

Reinhold, P., B. Rabeling, H. Günther, and D. Schimmel. 2002. Comparative evaluation of ultrasonography and lung function testing with the clinical signs and pathology of calves inoculated experimentally with *Pasteurella muftocida*. *Vet Rec* 150(4):109-114.

Reitsma, J. B., A. W. Rutjes, K. S. Khan, A. Coomarasamy, and P. M. Bossuyt. 2009. A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. *J Clin Epidemiol* 62(8):797-806.

Renaud, D. L., T. F. Duffield, S. J. LeBlanc, S. Ferguson, D. B. Haley, and D. F. Kelton. 2018. Risk factors associated with mortality at a milk-fed veal calf facility: A prospective cohort study. *J Dairy Sci* 101(3):2659-2668.

Renaud, D., T. Duffield, S. LeBlanc, D. Haley, and D. Kelton. 2017. Management practices for male calves on Canadian dairy farms. *J Dairy Sci* 100(8):6862-6871.

Rérat, M., S. Albin, V. Jaquier, and D. Hüsey. 2012. Bovine respiratory disease: efficacy of different prophylactic treatments in veal calves and antimicrobial resistance of isolated Pasteurellaceae. *Prev Vet Med* 103(4):265-273.

Rhodes, V., E. G. Ryan, C. J. Hayes, C. McAloon, L. O'Grady, S. Hoey, J. F. Mee, B. Pardon, B. Earley, and C. G. McAloon. 2021. Diagnosis of respiratory disease in preweaned dairy calves using sequential thoracic ultrasonography and clinical respiratory scoring: Temporal transitions and association with growth rates. *J Dairy Sci* 104(10):11165-11175.

Roberts, J. A., P. Kruger, D. L. Paterson, and J. Lipman. 2008. Antibiotic resistance—what's dosing got to do with it? *Crit Care Med* 36(8):2433-2440.

Rossini, K. L. 2004. Effects of calfhood respiratory and digestive disease on calfhood morbidity and first lactation production and survival rates. PhD Dissertation. Virginia Tech.

Rutjes, A. W., J. B. Reitsma, A. Coomarasamy, K. S. Khan, and P. M. Bossuyt. 2007. Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technol Assess* 11(50):iii, ix-51.

Rutjes, A. W., J. B. Reitsma, J. P. Vandenbroucke, A. S. Glas, and P. M. Bossuyt. 2005. Case-control and two-gate designs in diagnostic accuracy studies. *Clin Chem* 51(8):1335-1341.

Sans, P. and G. d. Fontguyon. 2009. Veal calf industry economics. *Rev Med Vet* 160(8-9):420-424.

Santman-Berends, I., A. J. G. de Bont-Smolenaars, L. Roos, A. G. J. Velthuis, and G. van Schaik. 2018. Using routinely collected data to evaluate risk factors for mortality of veal calves. *Prev Vet Med* 157:86-93.

Sargeant, J. M., T. E. Blackwell, S. W. Martin, and R. R. Tremblay. 1994a. Production indices, calf health and mortality on seven red veal farms in Ontario. *Can J Vet Res* 58(3):196-201.

Sargeant, J. M., T. E. Blackwell, S. W. Martin, and R. R. Tremblay. 1994b. Production practices, calf health and mortality on six white veal farms in Ontario. *Can J Vet Res* 58(3):189-195.

Sartori, S. and P. Tombesi. 2010. Emerging roles for transthoracic ultrasonography in pulmonary diseases. *World J Radiol* 2(6):203-214.

Schaefer, A., N. Cook, C. Bench, J. Chabot, J. Colyn, T. Liu, E. Okine, M. Stewart, and J. Webster. 2012. The non-invasive and automated detection of bovine respiratory disease onset in receiver calves using infrared thermography. *Res Vet Sci* 93(2):928-935.

Schaefer AL, Cook NJ, Church JS, Basarab J, Perry B, Miller C, Tong AK. 2007. The use of infrared thermography as an early indicator of bovine respiratory disease complex in calves. *Res Vet Sci* 83(3):376-84.

Schlageter-Tello, A., E. A. Bokkers, P. W. Groot Koerkamp, T. Van Hertem, S. Viazzi, C. E. Romanini, I. Halachmi, C. Bahr, D. Berckmans, and K. Lokhorst. 2014. Effect of merging levels of locomotion scores for dairy cows on intra- and interrater reliability and agreement. *J Dairy Sci* 97(9):5533-5542.

Schnyder, P., L. Schönecker, G. Schüpbach-Regula, and M. Meylan. 2019. Effects of management practices, animal transport and barn climate on animal health and antimicrobial use in Swiss veal calf operations. *Prev Vet Med* 167:146-157.

Schmidt, R. L. and R. E. Factor. 2013. Understanding sources of bias in diagnostic accuracy studies. *Arch Pathol Lab Med* 137(4):558-565.

Schönecker, L., P. Schnyder, G. Overesch, G. Schüpbach-Regula, and M. Meylan. 2019. Associations between antimicrobial treatment modalities and antimicrobial susceptibility in *Pasteurellaceae* and *E. coli* isolated from veal calves under field conditions. *Vet microbiol* 236:108363.

Schwarz, L. A. and A. S. Tidwell. 1999. Alternative imaging of the lung. *Clin Tech Small Anim Pract* 14(4):187-206.

Scott, P. R. 2013. Clinical presentation, auscultation recordings, ultrasonographic findings and treatment response of 12 adult cattle with chronic suppurative pneumonia: case study. *Ir Vet J* 66(1):1-10.

Sharon, K., Y. Liang, N. B. Sanchez, J. Carroll, P. Broadway, and M. Ballou. 2015. Preweaning plane of nutrition and *Mannheimia haemolytica* dose influence inflammatory responses to a combined bovine herpesvirus-1 and *Mannheimia haemolytica* challenge in postweaned Holstein calves. *Proc. ADSA JAM*:27.

Shimbo, G., M. Tagawa, K. Matsumoto, M. Tomihari, and K. Miyahara. 2018a. Effects of scatter correction processing on image quality of portable thoracic radiography in calves. *Japanese J Vet Res* 66(2):105-112.

Shimbo, G., M. Tagawa, K. Matsumoto, M. Tomihari, and K. Miyahara. 2018b. Evaluation of grid performance for thoracic radiography in calves using a portable x-ray unit. *Japanese J Vet Res* 66(1):13-19.

Shimbo, G., M. Tagawa, K. Matsumoto, M. Tomihari, M. Yanagawa, Y. Ueda, H. Inokuma, and K. Miyahara. 2018c. Three-legged radiographic view for evaluating cranioventral lung region in standing calves with bovine respiratory disease. *J Vet Med Sci*:18-0333.

Shivley, C., J. Lombard, N. Urie, C. Koprak, M. Santin, T. Earleywine, J. Olson, and F. Garry. 2018. Preweaned heifer management on US dairy operations: Part VI. Factors associated with average daily gain in preweaned dairy heifer calves. *J Dairy Sci* 101(10):9245-9258.

Šimundić, A.-M. 2009. Measures of diagnostic accuracy: basic definitions. *Ejifcc* 19(4):203.

Sischo, W., D. Moore, R. Pereira, L. Warnick, D. Moore, J. Vanegas, S. Kurtz, K. Heaton, D. Kinder, and J. Siler. 2019. Calf care personnel on dairy farms and their educational opportunities. *J Dairy Sci* 102(4):3501-3511.

Sivula, N., T. Ames, W. Marsh, and R. Werdin. 1996. Descriptive epidemiology of morbidity and mortality in Minnesota dairy heifer calves. *Prev Vet Med* 27(3-4):155-171.

Skinner, J. G., R. A. Brown, and L. Roberts. 1991. Bovine haptoglobin response in clinically defined field conditions. *Vet Rec* 128(7):147-149.

Slocombe, R. F., N. E. Robinson, F. J. Derksen, and C. B. Carrig. 1988. Thoracic radiographic appearance during lung inflation in neonatal calves. *Vet Res Commun* 12(2-3):259-272.

Smith, R. A., D. L. Step, and A. R. Woolums. 2020. Bovine Respiratory Disease: Looking Back and Looking Forward, What Do We See? *Vet Clin North Am Food Anim Pract* 36(2):239-251.

Spiegelhalter, D. J., N. G. Best, B. R. Carlin, and A. van der Linde. 2002. Bayesian measures of model complexity and fit. *J Roy Stat Soc B* 64:583-616.

Stanton, A., D. Kelton, S. LeBlanc, S. Millman, J. Wormuth, R. Dingwell, and K. Leslie. 2010. The effect of treatment with long-acting antibiotic at postweaning movement on respiratory disease and on growth in commercial dairy calves. *J Dairy Sci* 93(2):574-581.

Stanton, A. L., D. F. Kelton, S. J. LeBlanc, J. Wormuth, L. K. Fox, and K. E. Leslie. 2013. Effects of tulathromycin on incidence of various diseases and growth of young heifers. *J Am Vet Med Assoc* 243(2):267-276.

Stanton, A. L., D. F. Kelton, S. J. LeBlanc, J. Wormuth, and K. E. Leslie. 2012. The effect of respiratory disease and a preventative antibiotic treatment on growth, survival, age at first calving, and milk production of dairy heifers. *J Dairy Sci* 95(9):4950-4960.

Stull, C. L. and S. P. McDonough. 1994. Multidisciplinary approach to evaluating welfare of veal calves in commercial facilities. *J Anim Sci* 72(9):2518-2524.

Surén, P., K. Try, J. Eriksson, B. Khoshnewiszadeh, and K. O. Wathne. 2008. Radiographic follow-up of community-acquired pneumonia in children. *Acta Paediatr* 97(1):46-50.

Svensson, C. and P. Liberg. 2006. The effect of group size on health and growth rate of Swedish dairy calves housed in pens with automatic milk-feeders. *Prev Vet Med* 73(1):43-53.

Svensson, C., P. Liberg, and J. Hultgren. 2007. Evaluating the efficacy of serum haptoglobin concentration as an indicator of respiratory-tract disease in dairy calves. *Vet J* 174(2):288-294.

Svensson, C., K. Lundborg, U. Emanuelson, and S.-O. Olsson. 2003. Morbidity in Swedish dairy calves from birth to 90 days of age and individual calf-level risk factors for infectious diseases. *Prev Vet Med* 58(3-4):179-197.

Tegtmeier, C. and J. Arnbjerg. 2000. 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* 47(3):229-234.

Teixeira, A., J. McArt, and R. Bicalho. 2017a. Efficacy of tildipirosin metaphylaxis for the prevention of respiratory disease, otitis and mortality in pre-weaned Holstein calves. *Vet J* 219:44-48.

Teixeira, A. G., J. A. McArt, and R. C. Bicalho. 2017b. Thoracic ultrasound assessment of lung consolidation at weaning in Holstein dairy heifers: Reproductive performance and survival. *J Dairy Sci* 100(4):2985-2991.

Tejero, C. and A. Bach. 2016. 0104 The hidden cost of a hidden disease: growth performance of calves as affected by bovine respiratory disease diagnosed using ultrasonography. *J Anim Sci* 94(suppl_5):48-48.

Terra, R. L. and Reynolds, J. P. 2015. Ruminant history, physical examination, welfare assessment, and records. p7 In: *Large animal internal medicine. Fifth edition.* ed. B.P. Smith, ed. Elsevier Mosby, St. Louis, Missouri, USA

Theurer, M., B. White, R. Larson, and T. Schroeder. 2015. A stochastic model to determine the economic value of changing diagnostic test characteristics for identification of cattle for treatment of bovine respiratory disease. *J Anim Sci* 93(3):1398-1410.

Timsit, E., S. Assié, and N. Bareille. 2009. Failure to detect bovine respiratory disease in newly received young beef cattle by visual appraisal. in Proc. 12. Symposium of the International Society for Veterinary Epidemiology and Economics.

Timsit, E., S. Assié, R. Quiniou, H. Seegers, and N. Bareille. 2011a. Early detection of bovine respiratory disease in young bulls using reticulo-rumen temperature boluses. *Vet J* 190(1):136-142.

Timsit, E., N. Bareille, H. Seegers, A. Lehebel, and S. Assié. 2011b. Visually undetected fever episodes in newly received beef bulls at a fattening operation: occurrence, duration, and impact on performance. *J Anim Sci* 89(12):4272-4280.

Timsit, E., H. Christensen, N. Bareille, H. Seegers, M. Bisgaard, and S. Assié. 2013. Transmission dynamics of *Mannheimia haemolytica* in newly-received beef bulls at fattening operations. *Vet Microbiol* 161(3-4):295-304.

Timsit, E., N. Dendukuri, I. Schiller, and S. Buczinski. 2016. Diagnostic accuracy of clinical illness for bovine respiratory disease (BRD) diagnosis in beef cattle placed in feedlots: A systematic literature review and hierarchical Bayesian latent-class meta-analysis. *Prev Vet Med* 135:67-73.

Timsit, E., R. Leguillette, B. J. White, R. L. Larson, and S. Buczinski. 2018. Likelihood ratios: an intuitive tool for incorporating diagnostic test results into decision-making. *J Am Vet Med Assoc* 252(11):1362-1366.

Timsit, E., N. Tison, C. W. Booker, and S. Buczinski. 2019. Association of lung lesions measured by thoracic ultrasonography at first diagnosis of bronchopneumonia with relapse rate and growth performance in feedlot cattle. *J Vet Int Med* 33(3):1540-1546.

Toll, D., K. Janssen, Y. Vergouwe, and K. Moons. 2008. Validation, updating and impact of clinical prediction rules: a review. *J Clin Epidemiol* 61(11):1085-1094.

Urie, N., J. Lombard, C. Shivley, C. Kopral, A. Adams, T. Earleywine, J. Olson, and F. Garry. 2018a. Preweaned heifer management on US dairy operations: Part I. Descriptive characteristics of preweaned heifer raising practices. *J Dairy Sci* 101(10):9168-9184.

Urie, N., J. Lombard, C. Shivley, C. Koprak, A. Adams, T. Earleywine, J. Olson, and F. Garry. 2018b. Preweaned heifer management on US dairy operations: Part V. Factors associated with morbidity and mortality in preweaned dairy heifer calves. *J Dairy Sci* 101(10):9229-9244.

Van der Mei, J. and T. Van den Ingh. 1987. Lung and pleural lesions of veal calves at slaughter and their relationship with carcass weight. *Vet Q* 9(3):203-207.

Van Driessche, L., B. Valgaeren, L. Gille, F. Boyen, R. Ducatelle, F. Haesebrouck, P. Deprez, and B. Pardon. 2017. A deep nasopharyngeal swab versus nonendoscopic bronchoalveolar lavage for isolation of bacterial pathogens from preweaned calves with respiratory disease. *J Vet Int Med* 31(3):946-953.

Van Donkersgoed, J., C. S. Ribble, L. G. Boyer, and H. Townsend. 1993. Epidemiological study of enzootic pneumonia in dairy calves in Saskatchewan. *Can J Vet Res* 57(4):247.

Van Houten, C. B., C. A. Naaktgeboren, L. Ashkenazi-Hoffnung, S. Ashkenazi, W. Avis, I. Chistyakov, T. Corigliano, A. Galetto, I. Gangoiti, A. Gervais, D. Glikman, I. Ivaskeviciene, A. A. Kuperman, L. Lacroix, Y. Loeffen, F. Luterbacher, C. B. Meijssen, S. Mintegi, B. Nasrallah, C. Papan, A. M. C. van Rossum, H. Rudolph, M. Stein, R. Tal, T. Tenenbaum, V. Usonis, W. de Waal, S. Weichert, J. G. Wildenbeest, K. M. de Winter-de Groot, T. F. W. Wolfs, N. Mastboim, T. M. Gottlieb, A. Cohen, K. Oved, E. Eden, P. D. Feigin, L. Shani, L. J. Bont, and I. consortium. 2019. Expert panel diagnosis demonstrated high reproducibility as reference standard in infectious diseases. *J Clin Epidemiol* 112:20-27.

Van Leenen, K. 2020a. Subclinical pneumonia and airway inflammation in calves and their association with particulate matter, endotoxin and other barn climate parameters. PhD Dissertation. Ghent University.

Van Leenen, K., L. Van Driessche, L. De Cremer, C. Masmeijer, F. Boyen, P. Deprez, and B. Pardon. 2020b. Comparison of bronchoalveolar lavage fluid bacteriology and cytology in calves classified based on combined clinical scoring and lung ultrasonography. *Prev Vet Med* 176:104901.

Vandendriessche, S., W. Vanderhaeghen, F. V. Soares, M. Hallin, B. Catry, K. Hermans, P. Butaye, F. Haesebrouck, M. J. Struelens, and O. Denis. 2013. Prevalence, risk factors and genetic diversity of methicillin-resistant *Staphylococcus aureus* carried by humans and animals across livestock production sectors. *J Antimicrob Chemother* 68(7):1510-1516.

Vandermeulen, J., C. Bahr, D. Johnston, B. Earley, E. Tullo, I. Fontana, M. Guarino, V. Exadaktylos, and D. Berckmans. 2016. Early recognition of bovine respiratory disease in calves using automated continuous monitoring of cough sounds. *Comput Electron Agric* 129:15-26.

Verhoeff, J., A. Wierda, and A. P. van Nieuwstadt. 1985. Correlation of a disease scoring system with arterial PO₂ values in respiratory syncytial virus infection in calves. *Vet Q* 7(2):106-111.

Virtala, A.-M., G. Mechor, Y. Gröhn, and H. Erb. 1996. The effect of calfhoo diseases on growth of female dairy calves during the first 3 months of life in New York State. *J Dairy Sci* 79(6):1040-1049.

Virtala, A. M., Y. T. Grohn, G. D. Mechor, and H. N. Erb. 1999. The effect of maternally derived immunoglobulin G on the risk of respiratory disease in heifers during the first 3 months of life. *Prev Vet Med* 39(1):25-37.

Walsh, P., J. Thornton, J. Asato, N. Walker, G. McCoy, J. Baal, J. Baal, N. Mendoza, and F. Banimahd. 2014. Approaches to describing inter-rater reliability of the overall clinical appearance of febrile infants and toddlers in the emergency department. *PeerJ* 2:e651.

Wang, Z., Dendukuri, N., Zar, H.J., Joseph, L., 2017. Modeling conditional dependence among multiple diagnostic tests. *Stat. Med.* 36 (30), 4843–4859.

Ward, J. L., G. R. Lisciandro, and T. C. DeFrancesco. 2018. Distribution of alveolar-interstitial syndrome in dogs and cats with respiratory distress as assessed by lung ultrasound versus thoracic radiographs. *J Vet Emerg Crit Care (San Antonio)* 28(5):415-428.

Warnick, L., H. Erb, and M. White. 1994. The association of calfhoo morbidity with first-lactation calving age and dystocia in New York Holstein herds. *Kenya Vet* 18:177-179.

Weary, D., J. Huzzey, and M. Von Keyserlingk. 2009. Board-invited review: Using behavior to predict and identify ill health in animals. *J Anim Sci* 87(2):770-777.

Wilson, D. W., J. Lofstedt and J. Lakritz. 2015. Alterations in respiratory function. p40-75 In: *Large animal internal medicine*. Fifth edition. ed. B.P. Smith, ed. Elsevier Mosby, S. Louis, Missouri, USA.

Wilson, L. L., J. L. Smith, D. L. Smith, D. L. Swanson, T. R. Drake, D. R. Wolfgang, and E. F. Wheeler. 2000. Characteristics of veal calves upon arrival, at 28 and 84 days, and at end of the production cycle. *J Dairy Sci* 83(4):843-854.

Winder, C. B., D. F. Kelton, and T. F. Duffield. 2016. Mortality risk factors for calves entering a multi-location white veal farm in Ontario, Canada. *J Dairy Sci* 99(12):10174-10181.

Windeyer, M. C., K. E. Leslie, S. M. Godden, D. C. Hodgins, K. D. Lissemore, 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(2):231-240.

Windeyer, M. C., E. Timsit, and H. Barkema. 2017. Bovine respiratory disease in pre-weaned dairy calves: Are current preventative strategies good enough? *Vet J* 224:16-17.

WHO, 2021. Antimicrobial resistance. <https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance>

Wolfger B, Timsit E, White BJ, Orsel K. 2015. A systematic review of bovine respiratory disease diagnosis focused on diagnostic confirmation, early detection, and prediction of unfavorable outcomes in feedlot cattle. *Vet Clin North Am Food Anim Pract* 31(3):351-65.

Wongpakaran, N., T. Wongpakaran, D. Wedding, and K. L. Gwet. 2013. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. *BMC Med Res Methodol* 13(1):61.

Woolums, A. R. 2015. Lower Respiratory Tract Disease. p583-617 In: Large animal internal medicine. Fifth edition. ed. B. P. Smith, ed. Elsevier Mosby, St. Louis, Missouri, USA.

Yates, W. 1982. A review of infectious bovine rhinotracheitis, shipping fever pneumonia and viral-bacterial synergism in respiratory disease of cattle. *Can J Comp Med* 46(3):225.

Ye, X., H. Xiao, B. Chen, and S. Zhang. 2015. Accuracy of Lung Ultrasonography versus Chest Radiography for the Diagnosis of Adult Community-Acquired Pneumonia: Review of the Literature and Meta-Analysis. *PLoS One* 10(6):e0130066.

Zapf, A., S. Castell, L. Morawietz, and A. Karch. 2016. Measuring inter-rater reliability for nominal data— which coefficients and confidence intervals are appropriate? *BMC Med Res Methodol* 16(1):93.

Zeineldin, M., J. Lowe, and B. Aldridge. 2019. Contribution of the mucosal microbiota to bovine respiratory health. *Trends Microbiol* 27(9):753-770.

Zeineldin, M. M., Y. El-Raof, H. A. El-attar, and M. M. Ghanem. 2016. Lung ultrasonography and computer-aided scoring system as a diagnostic aid for bovine respiratory disease in feedlot cattle. *Glob Vet* 17:588-594.

Zhang, M., J. E. Hill, C. Fernando, T. W. Alexander, E. Timsit, F. van der Meer, and Y. Huang. 2019. Respiratory viruses identified in western Canadian beef cattle by metagenomic sequencing and their association with bovine respiratory disease. *Transbound Emerg Dis* 66(3):1379-1386.

Annexes

Annexe 1: Facteurs de virulence des agents microbiens isolés et considérés comme pathogènes lors de bronchopneumonie infectieuse chez les veaux de race laitière., tableaux adaptés de Panciera et Confer (2010) pour les bactéries et le Smith (2015) pour les virus.













VIRUS

Virus	Matériel génétique	Enveloppe	Proteine virale	Tropisme	Pathogénie
Herpesvirus Bovin type-1 (BHV-1)	ADN double brin.	Double donc plus fragile dans l'environnement.	◆ Glycoprotéine C, D et B permettant l'attachement aux cellules épithéliales	Voies respiratoires supérieures	◆ Destruction des cellules épithéliales infectées ◆ Immunosuppression
Virus Parainfluenza type-3 (PI-3)	ARN- avec besoin d'une protéine de réplication donc période d'incubation élargie (3 à 5 jours).	Simple donc plus fragile dans l'environnement.	◆ Protéine F permettant la fusion cellulaire	Voies respiratoires supérieures	◆ Dommages à l'appareil muco-ciliaire ◆ Immunosuppression
Virus Syncytial Respiratoire Bovin (BRSV)	ARN- avec besoin d'une protéine de réplication donc période d'incubation élargie (3 à 5 jours).	Oui donc plus fragile dans l'environnement.	◆ Protéine F permettant la fusion cellulaire	Voies respiratoires inférieures	◆ Fusion des alvéoles et formation de syncytia ◆ Immunosuppression ◆ dégranulation des mastocytes de l'hôte
Virus de la Diarrhée Virale Bovine type 1 et 2 (BVD-1, BVD-2)	ARN+ avec besoin d'une protéine de réplication donc période d'incubation élargie (3 à 5 jours).	Oui donc plus fragile dans l'environnement.	◆ Protéine E2 permettant l'attachement aux cellules	Principales les lymphocytes	Immunosuppression
Adénovirus	ADN double brin- peut se lier au génome.	Non donc résistant dans l'environnement.	◆ Héماغglutinine pour l'agglutination ◆ Protéine de transformation aux propriétés oncogènes ◆ <i>Internal Terminal Repeat</i> aux propriétés oncogènes	Voies respiratoires supérieurs	Infection bénigne
Le Coronavirus type-A (Co)	ARN+ avec besoin d'une protéine de réplication donc période d'incubation élargie (3 à 5 jours).	Oui donc plus fragile dans l'environnement.	◆ Protéine M et E permettant l'attachement aux cellules	Voies respiratoires supérieurs	Infection bénigne
Virus Influenza D	ARN – avec besoin d'une protéine de réplication donc période d'incubation élargie (3 à 5 jours).	Oui donc plus fragile dans l'environnement.	◆ Neuraminidase qui facilite le mouvement du virus ◆ Héماغglutinine qui permet de fusionner avec les cellules et relargé son matériel génétique	Voies respiratoires inférieurs	Infection bénigne

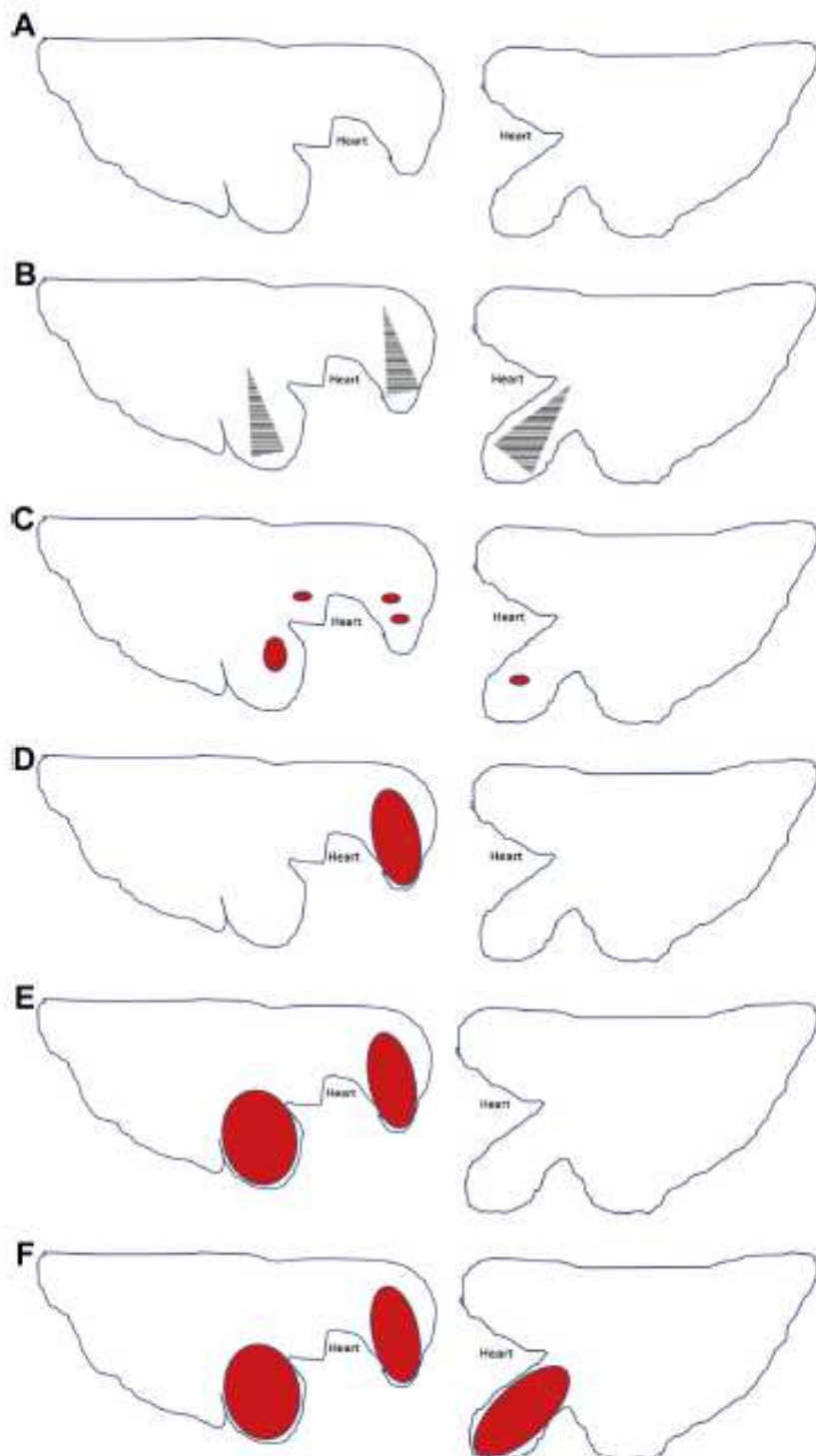
BACTÉRIES

Bactéries	Capsule	Endotoxine	Exotoxine	Protéines adhésions	Enzymes secrétées	Autres facteurs
<i>Mycoplasma bovis</i>	Non	Non	Toxine polysaccharide	VSPs		<ul style="list-style-type: none"> ◆ Biofilm ◆ Peroxide d'hydrogène
<i>Histophilus somni</i>	Non	LOS	-			<ul style="list-style-type: none"> ◆ Biofilm ◆ IgBPs ◆ Histamine
<i>Mannheimia haemolytica</i>	Oui	LPS	LKT	<ul style="list-style-type: none"> ◆ OmpA ◆ Lipoprotéine I ◆ N-acetyl-α-glucosamine ◆ Protéine d'adhésion du fibrinogène 	<ul style="list-style-type: none"> ◆ Neuraminidase ◆ Sialoglycoprotease 	Biofilm
<i>Pasteurella multocida</i>	Oui	LPS	-	<ul style="list-style-type: none"> ◆ OmpA ◆ Fimbriae Type IV ◆ FHA 	<ul style="list-style-type: none"> ◆ Neurominidase 	Biofilm
<i>Trueperella pyogenes</i>	Non	Non	Pyolysin	Protéine d'adhésions du collagène	<ul style="list-style-type: none"> ◆ Proteases ◆ ANDase 	Biofilm
<i>Bibersteinia trehalosi</i>	Oui	Oui	LKT selon les souches	<ul style="list-style-type: none"> ◆ 2 OmpA ◆ Protéine d'adhésion du fibrinogène 	<ul style="list-style-type: none"> ◆ Protease ◆ Dismutase superoxide 	
<i>Chlamydia psittaci</i>	Non	Non	-	Hemagglutinine qui facilite l'attachement des cellules		La réponse immunitaire à médiation cellulaire
<i>Moraxella bovis ou ovis</i>	Non	LPS	Cytotoxine	Pillin qui facilite l'attachement		

Annexe 2: Le score du Wisconsin, adapté de Mc Guirk (2008)

Score de santé des veaux			
0	1	2	3
Température rectale			
100-100.9 (37.3-38.2 C)	101-101.9 (38.3-38.8 C)	102-102.9 (38.9-39.4 C)	>ou = à 103 F (39.5 C)
Toux			
Aucune	Toux simple induite	Toux induite répétée	Toux répétée continue
Décharge nasale			
Normale	Petite quantité unilatérale trouble	Bilatérale trouble ou mucus excessif	Beaucoup d'écoulement mucopurulent
			
Yeux			
Normaux	Peu d'écoulement	Écoulement modéré bilatéral	Écoulement sévère
			
Oreilles			
Normales	Tremblement de la tête ou battement d'oreille	Oreille légèrement tombante d'un côté	Tête sur le côté ou 2 oreilles tombantes
			

Annexe 3 : Le score échographique de Ollivett et Buczinski (2016)



Score 0 = poumon normal sans consolidation et aucune ou quelque artefact de queue de comète

Score 1 = présence de multiples artefacts de queue de comète diffus sans consolidation







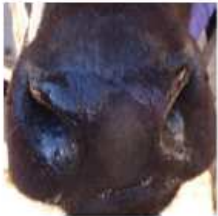





Score 2 = pneumonie lobulaire ou irrégulière. Les petites lésions lobulaires sont la plupart virale et ne nécessite pas nécessairement de traitement

Score 3 = pneumonie lobaire affectant seulement un lobe

Score 4 = pneumonie lobaire affectant deux lobes. La portion crâniale et caudale du lobe crânial droit est scoré individuellement







Score 5 = pneumonie lobaire affectant 3 lobes ou plus

Annexe 4 : Le score Californien, tiré de l'article de Aly et al. (2014)

Clinical sign	Score if normal	Score if abnormal (any severity) ²
Eye discharge	0	2
		 Or  Or 
Nasal discharge	0	4
		 Or  Or 
Ear droop/Head tilt	0	5
		 Or  Or 
Cough	0	2
Breathing	0	2
Temperature	0	2
	No cough	Spontaneous cough
	Normal	Rapid or difficult breathing
	<39.2°C	≥39.2°C (102.5°F)

Annexe 5 : Le score de Maier, tiré de l'article de Maier et al. (2019)

California bovine respiratory disease scoring system for post-weaned dairy calves

Clinical sign	Score if normal	Score if abnormal (any severity) ¹	
Sunken eyes	0 	4  Or 	
Low body condition	0 	5  Or 	
Cough	0 No cough	2	Spontaneous cough
Breathing	0 Normal	1	Rapid or difficult breathing
Diurnal temp fluctuation	0 ≤27° F (≤15° C)	1	> 27° F (>15° C)

With diurnal temperature data:
calf is score positive² if total score ≥ 2

Without diurnal temperature data:
calf is score positive² if total score ≥ 1

Confirmatory step for score positive ³	Do not treat	Treat
Rectal temperature	< 102.5° F (< 39.2° C)	≥ 102.5° F (≥ 39.2° C)

- Any abnormality including, but not limited to, the examples shown in the above pictures. For body condition, a score of ≤ 2 on a scale of 5 would be considered abnormal.
- With diurnal temperature fluctuation: screening sensitivity 77%, diagnostic sensitivity 100%, specificity 62%, without diurnal temperature fluctuation: screening sensitivity 84%, diagnostic sensitivity 100%, specificity 48%
- With diurnal temperature fluctuation: confirmatory step improves specificity to 77% (screening sensitivity 65%, diagnostic sensitivity 77%), without diurnal temperature fluctuation: confirmatory step improves specificity to 63% (screening sensitivity 71%, diagnostic sensitivity 77%)

Annexe 6 : Photos de la technique d'auscultation au sac. Un gant de fouille est placé sur le nez du veau par un manipulateur. En respirant à travers le gant, la quantité d'air inspiré est restreinte, favorisant une augmentation de l'amplitude respiratoire. Un deuxième manipulateur ausculte afin de tecter l'apparition de bruits respiratoires anormaux durant la respiration avec le gant puis quelques minutes après le retrait du gant.



Annexe 7 : Photos de la technique d'échographie thoracique, tiré de l'article de Ollivett et Buczinski (2016)



Photo A : Photo illustrant la technique d'échographie thoracique chez un veau de race laitière.

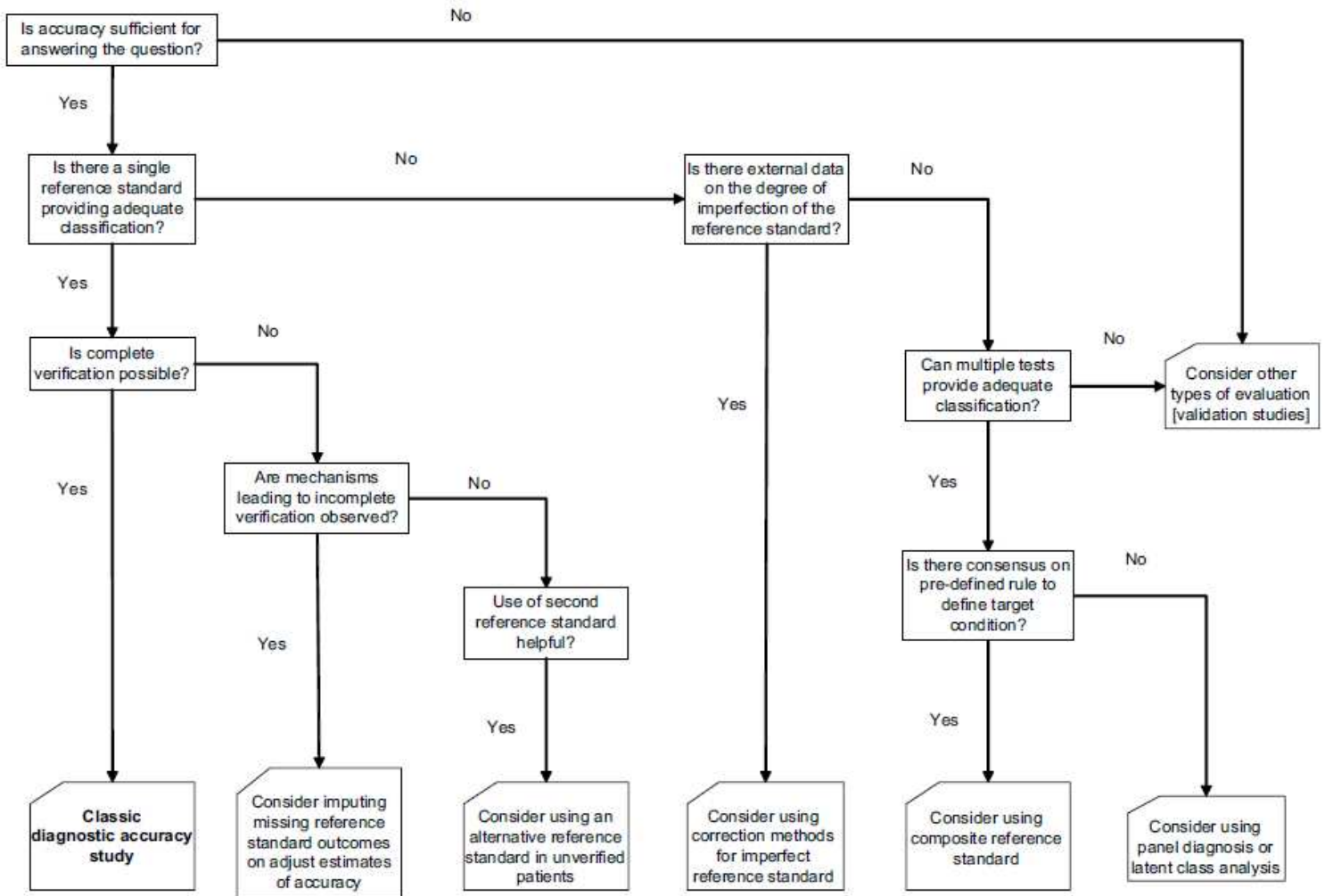
Photo B : Photo illustrant la technique d'échographie thoracique du lobe crânial droit chez un veau de race laitière.



Annexe 8 : Photo de la technique de radiographie thoracique chez un veau de race laitière en condition de terrain, tiré de l'article de Shimbo et al (2018c). Cette vue crâniale est obtenue en positionnant un des deux membres thoraciques en avant chez un veau debout. Le membre thoraciques controlatéral et les membres postérieures sont attachés à une chute à l'aide de sangle.



Annexe 9 : Algorithme décisionnel pour déterminer la méthode d'élaboration d'un test standard de référence, tiré de Reitsma et al. (2009)



Annexe 10 : Les performances de l'utilisation de l'auscultation thoracique seule *versus* la combinaison de l'auscultation thoracique et de l'échographie thoracique pour détecter la bronchopneumonie infectieuse (BPI) chez les veaux de race laitière, tiré de Buczinski et al. (2016).

(A) Probabilité de BPI dans une population, (B) après une auscultation thoracique (AUSC) (positive ou négative) et (C) après AUSC combiné à une échographie thoracique (TUS).

