Journal of Veterinary Diagnostic Investigation

Bayesian Estimation of the Performance of Using Clinical Observations and Harvest Lung Lesions for Diagnosing Bovine Respiratory Disease in Post-weaned Beef Calves

Brad J. White and David G. Renter

The online version of this article can be found at:

Published by:
⑤SAGE

On behalf of:

AAVLD

Official Publication of the American Association of Veterinary Laboratory Diagnosticians, Inc.

Additional services and information for *Journal of Veterinary Diagnostic Investigation* can be found at:

Email Alerts: http://vdi.sagepub.com/cgi/alerts

Subscriptions: http://vdi.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

>> Version of Record - Jul 1, 2009

What is This?

Bayesian estimation of the performance of using clinical observations and harvest lung lesions for diagnosing bovine respiratory disease in post-weaned beef calves

Brad J. White,[1] David G. Renter

Abstract. Bovine respiratory disease (BRD) diagnosis during the postweaning phase of beef production is an important component of effective preventive health and treatment programs. Although identification of diseased animals based on signs of clinical illness (CI) is a common method in the beef industry for identifying BRD, very little information is available on the accuracy of this method. Previous investigators hypothesized that monitoring pulmonary lesions at harvest (LU) could be a more reliable indicator of disease status during the postweaning phase. A structured literature review was conducted to identify research that compared CI and LU. Because there is no true gold standard for diagnosing BRD, Bayesian methods were used to estimate the sensitivity and specificity of each diagnostic method relative to a BRD diagnosis at any time during the postweaning phase. Results from the current study indicate that the estimated diagnostic sensitivity and specificity of CI were 61.8% (97.5% probability interval [PI]: 55.7, 68.4) and 62.8% (97.5% PI: 60.0, 65.7), respectively. Use of LU for a BRD diagnosis was estimated to have a sensitivity of 77.4% (97.5% PI: 66.2, 87.3) and a specificity of 89.7% (97.5% PI: 86.0, 93.8). Further analysis revealed that the probabilities of LU having higher sensitivity and specificity than CI were 99.4% and 100%, respectively. The present research indicates that neither method was perfect, and both methods were relatively poor at correctly classifying truly diseased animals (sensitivity) but that LU was more accurate than CI for BRD diagnosis. Results from the present study should be considered when these diagnostic methods are used to evaluate BRD outcomes in clinical and research settings.

Key words: Bayesian analysis; bovine respiratory disease; diagnostic methods.

Introduction

Accurate diagnosis of bovine respiratory disease (BRD) is an important component of an effective treatment program and is critical for objective evaluation of preventative health procedures. Respiratory disease in cattle is a common and costly ailment during the postweaning phase.[4,19] Multiple etiologies and risk factors contribute to the disease complex, and, often, the diagnosis is based on a syndrome, including a combination of clinical signs, rather than identifying a single pathogen. This disease complex affects the growth performance,[8] final carcass traits,[11] and profitability[7,17] of the affected cattle; therefore, accurate diagnosis and subsequent treatment is important to mitigate negative impacts.

A clinical BRD diagnosis often relies on observation of clinical illness (CI), based on signs, including depression, anorexia, nasal discharge, coughing, and lack of rumen fill,[20] with depression described as the most common clinical sign.[12] The final treatment decision often combines visual observations with a determination of rectal temperature.[22] Although CI is a common BRD diagnostic method, several investigators have suggested that pulmonary lesions at harvest (LU) may provide a more accurate means for assessing which animals experienced BRD during the feeding phase.[3,8,21,26] However, results of trials indicated that up to 68% of animals untreated based on clinical signs may possess LU[26] and that not all calves treated for BRD will have LU.[8]

Published data that compared CI and LU are sparse, and there is no scientific literature that describes the diagnostic performance of using these diagnostic methods to diagnose BRD during the postweaning phase. Estimating diagnostic characteristics of the selected methodologies can affect interpretation of diagnostic results. The accuracy of these diagnostic methods influences estimates of disease frequency, measures of treatment program efficacy, measures of associations between performance parameters and disease conditions, and the ability to evaluate preventative health programs. Neither CI nor LU is perfectly accurate at identification of BRD during the postweaning phase, and,

From the Departments of Clinical Sciences (White) and Diagnostic Medicine/Pathobiology (Renter), Kansas State University, Manhattan, KS.

[1] Corresponding Author: Brad J. White, Department of Clinical Sciences, Kansas State University, Manhattan, KS 66506. bwhite@vet.ksu.edu

thus, analytical procedures for evaluating the accuracy of these diagnostic methods in the absence of a gold standard are necessary.[2,6] Bayesian analysis has been used to estimate the sensitivity (Se) and specificity (Sp) of diagnostic methods for a particular syndrome or pathogen when no gold standard is present.[5,14–16] The aim of the current research was to use Bayesian methods to estimate Se and Sp of CI and harvest lung lesions for diagnosis of BRD in postweaned beef calves. This research was unique in that it provided estimates of diagnostic characteristics that can be applied to evaluating disease frequency and treatment outcomes in a more robust manner.

Materials and methods

Data selection

A structured veterinary literature search was conducted to identify relevant published articles with data that allowed comparisons of the diagnostic characteristics of CI and LU. Required information included the ability to cross-classify individual animals based on the results of both diagnostic methods during the postweaning phase. Broad search terms were used to maximize the number of hits for potential inclusion. A PubMed search (1966 to June 2008) was performed and included search terms: cattle (medical subject heading [MeSH]), lung (MeSH), lesions, treatment, and feedlot. The search was also completed on 9 Agricola databases (Agricola 2008/04–2008/06, Agricola 2006–2008/03, Agricola 2002–2005, Agricola 1996–2001, Agricola 1991–1995, Agricola 1984–1990, Agricola 1980–1983, Agricola 1975–1979, Agricola 1970–1974) by using the same search terms. Article titles and abstracts were reviewed to select published articles with data relevant to the current objective. Selected articles were then reviewed to ensure that required data were available and that the studies were performed on the desired population (postweaned beef calves).

Statistical analysis

After studies with required data were identified, kappa statistics were calculated[a] for each study to assess the agreement between the 2 diagnostic methods beyond that expected from chance alone. Data from the identified studies were then analyzed with Bayesian confidence profile models by using WinBUGS statistical software.[13] The Bayesian approach was chosen because this method of modeling allows incorporation of uncertainty in the values of Se and Sp, and makes no assumption of the true status of the sampling unit.[23] This methodology can be used to calculate the Se and Sp of diagnostic methods by applying the methods to 2 different populations with different disease prevalences.[23] Bayesian theory uses direct probabilities for inference about scientific hypothesis rather than P values to account for the chance aspect of data-driven inferences, so no P values are reported. Instead, the probability of the scientific hypotheses that lung lesions at harvest were superior in Se or Sp to clinical observations during the feeding phase was calculated. The model-generated estimates for each parameter and the 97.5% probability interval (PI) around each estimate were also reported.

Prior distributions were specified by using beta distributions to model-subjective probability about the unknown parameters.[6] There is no published information on the Se or Sp of either CI or LU; therefore, priors in the model were determined by expert panel opinion[b] by using methods described by previous investigators.[6] Commercial software[c] was used to generate beta distributions for the estimated Se and Sp of each diagnostic method and the prevalence in each population. For Se of LU, the expert panel estimated a most probable value of 75% with 90% confidence that the Se was greater than 65%. This would be a conservative estimate stating that the investigators were 90% confident that no more than 35% of the animals that truly had BRD during the postweaning phase will be free of pulmonary lesions at harvest. The expert panel estimated a most probable value of 85% for Sp of the pulmonary lesions with 90% confidence that the value was greater than 80%. This allowed for the possibility that animals had lung lesions before the induction of the feeding phase that were still evident at harvest, but the expert panel was 90% confident that this did not occur more than 20% of the time. The experts believed that both the Se and Sp were likely lower for CI than LU, but, because of uncertainty, a lower degree of confidence in the expected values for CI was suggested. The experts estimated the Se of CI to have a most probable value of 60%, with 90% certainty that Se was less than 70%. The Sp of CI was estimated as most likely 65%, with 90% certainty that it was less than 80%.

Prior values for the prevalence in each identified study were generated based on existing published data. The U.S. Department of Agriculture reported an average prevalence of BRD at 14.4% for cattle in U.S. feedlots[24]; however, this prevalence is highly variable because of multiple factors that contribute to the syndrome.[20] Although preventative health programs may play a role in disease risk, cumulative incidence of BRD during the first 60 days after arrival at a feeding facility was reported to be as high as 17% in preconditioned calves.[25] Therefore, prevalence priors were generated based on information from a 2008 article that described morbidity counts in a comparable population of postweaned beef calves fed in a common management scheme.[18] The 2008 study only followed the morbidity counts for the initial 12 weeks of the feeding phase; however, the investigators of the present study believe that this is still a reasonable estimate, because most BRD morbidity occurs within the first 45 days of the feeding period.[4] The 2008 study reported a mean cumulative incidence of 5.9%, with an individual pen range from 0% to 80%. Although the study populations in the current project were very similar (initial weight, calf type, and days on feed), cattle could be denoted as originating from single or multiple sources. Cattle that originated from multiple sources were identified to have up to twice the risk of initial BRD morbidity compared with a pen originating from a single source.[18] Therefore, the prior prevalence for single-source cattle was estimated with a most probable value of 6%, with 95% certainty that the true prevalence was less

Table 1. Cross-classification of the results from 2 independent studies, showing animals exhibiting clinical signs of bovine respiratory disease (BRD) at any time during the postweaning phase, pulmonary lesions at harvest, or both.

	Pulmonary lesions at harvest					
	Gardner et al. 1999 (n = 204)*			Thompson et al. 2006 (n = 2,036)†		
	Present	Absent	Total	Present	Absent	Total
BRD clinical signs						
Present	49	53	102	265	196	461
Absent	38	64	102	606	969	1,575
Total	87	117		871	1,165	

* Kappa (95% confidence interval [CI]) 0.11 (−0.03, 0.24).

† Kappa (95% CI) 0.14 (0.10, 0.18).

than 80%. The most probable prevalence for cattle originating from multiple sources was estimated to be 12%, with 95% certainty that the true prevalence was less than 80%.

Prior values were included in a conditional dependence model to evaluate the relationship between the 2 diagnostic methods, then a conditional independence Bayesian model for 2 tests in 2 populations was used as previously described.[2] To test the effect of the priors on the posterior distributions, the assumptions regarding Se and Sp of each diagnostic method were tested through a Se analysis that requires construction of additional models. The distributions of prior assumptions on Se and Sp of CI were changed to uniform (0,1) distributions, and results were compared with the previously selected levels. The process was also completed with the attributes of the LU set to noninformative priors while holding all other priors at previous levels. The impact of the assumptions regarding prevalence was tested by creating models using a uniform (0,1) distribution for prevalence in each study. A model was also completed with all priors modeled as noninformative, uniform (0,1) distributions.

Each model was assessed after a burn-in of 5,000 iterations, and posterior inferences were based on 100,000 iterations. Convergence was assessed by visual evaluation of kernel density and trace plots for each parameter.[6] The potential effect of the initial values and further assessment of convergence were evaluated by running multiple chains with divergent starting values and visual evaluation of a Gelman and Rubin statistic.[10]

Results

The structured search revealed several articles (10 PubMed, 2 Agricola) of potential relevance to the research. Two articles[8,21] were identified on both searches. A review of the abstracts narrowed the list to 3 articles with relevant information.[8,21,26] One of the articles, published in 1996,[26] followed calves from birth to harvest, and it was not possible from the data presented in the article to identify the cattle diagnosed with BRD only during the postweaning phase. A review of the remaining 2 articles, published in 1999 and 2006, respectively, revealed that the necessary

data were present in both[8,21]; therefore, they were selected for inclusion in the current study.[8,21]

Both the 1999[8] and 2006[21] articles used a similar case definition of BRD in the postweaning phase based on clinical signs of the BRD syndrome, including depression, a lack of rumen fill, slow-moving nasal or ocular discharge, or cough. An evaluation of rectal temperature was performed in both studies on cattle were presented with signs of BRD, and cattle that exhibited rectal temperatures higher than 40°C were treated (were considered cases). One minor difference in case definition is that, in the 2006 study, cattle that had a rectal temperature <40°C could be treated at the discretion of the treatment provider if they were showing signs of BRD.

The cattle in the study populations were relatively similar in both trials. The 1999 study consisted of 204 head, with a mean arrival weight of 291 kg, that were fed for an average of 150 days before harvest. Two thousand and thirty-six cattle were enrolled in the 2006 study; the cattle had a mean arrival weight of 233 kg and were fed for an average of 137 days before harvest. One difference between the study populations was cattle origin; cattle in the 1999 study were from a single location, whereas cattle in the 2006 study were procured from multiple sources. Cattle in both studies were fed in a commercial system by using similar, common management practices.

Cross-classified CI and LU results for the 2 studies are presented in Table 1. When using CI, 50% of the animals in the 1999 study were positive for BRD compared with 43% that presented with pulmonary lesions. The kappa value for the 2 diagnostic methods in this population showed low agreement (0.11). In the 2006 research, only 23% of the animals were diagnosed with BRD based on CI, yet 43% still had evidence of pulmonary lesions at harvest. Again, a comparison of the 2 diagnostic methods resulted in a low kappa value (0.14), which indicated a low level of agreement.

Table 2. Results from Bayesian models used to estimate mean and 97.5% probability intervals for the prevalence of bovine respiratory disease (BRD) in the 2 populations, and for sensitivity (Se) and specificity (Sp) of diagnosing BRD in calves by using pulmonary lesions at harvest (LU) or signs of clinical illness (CI) identified during the postweaning phase.*

	Mean % (97.5% probability interval)				
	Model 1†	Model 2‡	Model 3§	Model 4‖	Model 5¶
SeCI	61.8 (55.7, 68.4)	57.5 (52.3, 64.0)	62.2 (55.6, 69.6)	57.8 (52.8, 64.0)	61.7 (55.6, 68.2)
SpCI	62.8 (60.0, 65.7)	62.9 (59.9, 66.9)	62.8 (60.0, 65.7)	62.7 (59.8, 66.3)	62.8 (60.1, 65.8)
SeLU	77.4 (66.2, 87.3)	82.4 (59.1, 99.1)	77.5 (66.3, 87.3)	84.7 (62.9, 99.3)	76.9 (65.5, 87.0)
SpLU	89.7 (86.0, 93.8)	96.6 (90.2, 99.9)	89.6 (58.8, 93.1)	96.4 (90.1, 99.9)	89.8 (86.1, 93.2)
SeLUCI	99.4 (100.0, 100.0)	98.7 (100.0, 100.0)	99.3 (100.0, 100.0)	99.5 (100.0, 100.0)	99.3 (100.0, 100.0)
SpLUCI	100.0 (100.0, 100.0)	100.0 (100.0, 100.0)	100.0 (100.0, 100.0)	100.0 (100.0, 100.0)	100.0 (100.0, 100.0)
PrevGAR	56.3 (43.9, 69.9)	59.0 (49.4, 80.9)	56.0 (43.5, 69.7)	56.8 (43.2, 75.9)	57.3 (44.6, 71.4)
PrevTHO	19.1 (13.6, 25.3)	25.0 (16.2, 36.5)	19.0 (13.4, 25.2)	23.9 (15.9, 34.0)	19.4 (13.8, 25.7)

* Models were based on conditional independence between tests; CI used to abbreviate animals that met the treatment criteria outlined in each study by using CI, and LU represents the dichotomous classification of each animal at harvest based on the presence of pulmonary lesions. SeLUCI is the probability that the Se of LU is greater than the Se of CI, and SpLUCI is the probability that the Sp of LU is greater than CI. Prev = prevalence; GAR = Gardner et al. 1999; THO = Thompson et al. 2006.

† Model 1 used the following informative priors for beta distributions: SeCI: β (20.29,18.86), SpCI: β (6.97,4.21), SeLU: β (31.36,11.12), SpLU: β (98.71,18.24), PrevTHO: β (1.13,1.96), and PrevGAR: β (1.06,1.91).

‡ Model 2 used noninformative (uniform [0,1]) priors for all Se, Sp, and prevalence values.

§ Model 3 used noninformative (uniform [0,1]) priors for SeCI and SpCI and the priors described in model 1 for the remaining values.

‖ Model 4 used noninformative (uniform [0,1]) priors for SeLU and SpLU and the priors described in model 1 for the remaining values.

¶ Model 4 used noninformative (uniform [0,1]) priors for PrevGAR and PrevTHO and the priors described in model 1 for the remaining values.

Utilization of the conditional dependence model allowed evaluation of the potential dependence among accuracy measures for the 2 diagnostic methods. The estimated correlation of the Se (95% PI) was 0.05 (−0.29 to 0.35), whereas the estimated correlation of Sp (95% PI) between the methods was 0.14 (−0.07 to 0.30). The PIs for both values overlapped zero; therefore, models for independent methods were used.[2]

Bayesian estimates of the posterior distributions of Se and Sp for each diagnostic method in the different conditional independence models are presented in Table 2. All models exhibited convergence after 100,000 iterations, after a 5,000-iteration burn-in. Results of the Se analysis (models 2–5) revealed minimal differences in Se, Sp, or prevalence estimates for either study. The Sp of LU was slightly higher when noninformative priors were used for Se of LU and Sp of LU (models 2 and 4); however, the 97.5% PIs from these models broadly overlapped the PI for Sp of LU in model 1.

The CI method had relatively low Se and Sp, 61.8% (97.5% PI: 55.7, 68.4) and 62.8% (97.5% PI: 60.0, 65.7), respectively. In the current study, the model estimate for Se of using LU for diagnosing BRD was 77.4% (97.5% PI: 66.2, 87.3), and the estimated Sp of LU was relatively high, 89.7% (97.5% PI: 86.0, 93.8). When using a direct comparison between estimates for CI and LU, there was a 99.4% probability (97.5% PI: 100, 100) that LU had a higher diagnostic Se comp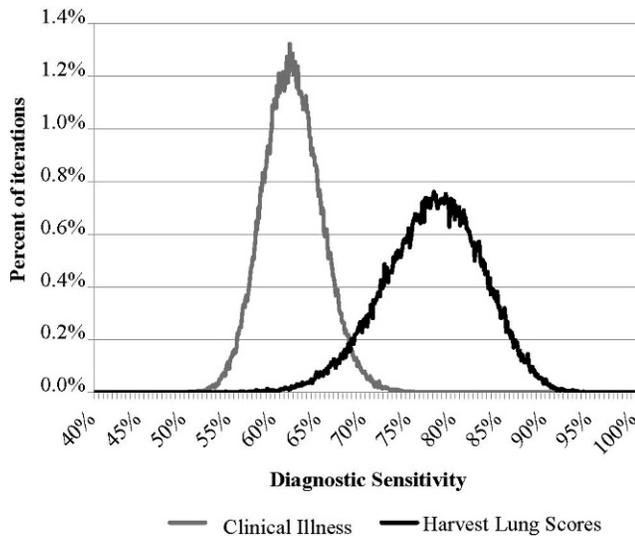ared with CI. The distribution of model-estimated Se values for each method are presented in Figure 1. The distribution of Sp findings (presented in Fig. 2) illustrate that the Sp for LU was also higher (100, 97.5% PI: 100, 100) than CI Sp.

Discussion

The present study used existing data and Bayesian models to evaluate the diagnostic performance of using clinical observation and lung lesions at harvest for diagnosing BRD. Although these methods have been advocated, little information exists on the diagnostic performance because of the absence of a true gold standard diagnostic method. The current study illustrates that LU has higher Se and Sp compared with CI and that the diagnostic characteristics of these methods may influence interpretation of health outcomes from clinical trials and field investigations.
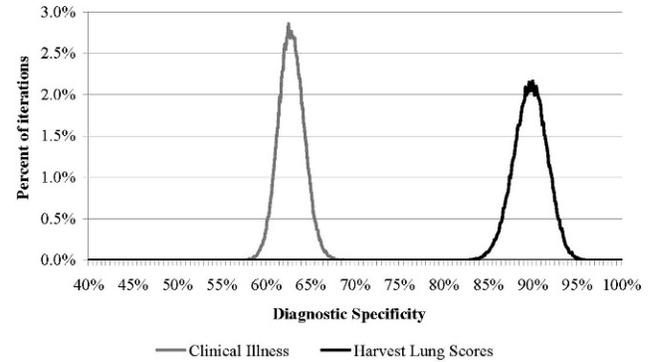
Results of the kappa analysis reveal very little agreement between the selected diagnostic methods in these trials and may indicate an inherent difference in diagnostic ability of the 2 methods. The low agreement further emphasizes the importance of estimating the Se and Sp of each diagnostic method, and, to accomplish this, a conditional independence Bayesian model for 2 tests in 2 populations was applied as previously described.[2]

The implied assumptions for the Bayesian model, also known as the Hui-Walter paradigm, state that 1) the population is divided into 2 or more populations with differing prevalences, 2) 2 tests are performed in

Figure 1. Distributions of model estimates (100,000 iterations) for the diagnostic sensitivity (Se) of using observed clinical illness (CI) and lung lesions at harvest (LU) for diagnosing bovine respiratory disease in postweaned cattle. Informative priors for beta distributions in the model were: SeCI: β (20.29, 18.86), specificity (Sp) CI: β (6.97, 4.21), SeLU: β (31.36, 11.12), SpLU: β (98.71, 18.24), PrevTHO: β (1.13, 1.96), and PrevGAR: β (1.06, 1.91). Data represent 100,000 iterations of one chain, with start values near the priors for each parameter. Prev = prevalence; THO = Thompson et al. 2006; GAR = Gardiner et al. 1999.

Figure 2. Distribution of model estimates (100,000 iterations) for the diagnostic specificity (Sp) of using observed clinical illness (CI) and lung lesions at harvest (LU) for diagnosing bovine respiratory disease in postweaned cattle. Informative priors used in the model: sensitivity (Se) CI: β (20.29, 18.86), SpCI: β (6.97, 4.21), SeLU: β (31.36, 11.12), SpLU: β (98.71, 18.24), PrevTHO: β (1.13, 1.96), and PrevGAR: β (1.06, 1.91). Data represent 100,000 iterations of one chain, with start values near the priors for each variable. Prev = prevalence; THO = Thompson et al. 2006; GAR = Gardiner et al. 1999.

both populations, and the Se and Sp are the same between populations, and 3) the 2 tests are conditionally independent.[23] This model was selected because the data were from 2 populations. Also, because of the high variability associated with BRD occurrence in postweaning situations, it was very probable that the prevalences differed, and the analysis revealed that prevalence in the 2 study populations did, indeed, differ (Table 2). The second assumption may be affected by the fact that cases were identified based on clinical signs, which can be subjective, and the studies were conducted in multiple areas, which meant that different observers were responsible for diagnosis. However, the alternative to this assumption was used, and the estimates of the Se and Sp were interpreted as average values across both populations.[2] Therefore, the model-generated Se and Sp values reported herein represent single estimates for both populations and cannot be interpreted separately for the 2 study populations that may have differed based on numerous factors, including cattle type, clinical observers, study locations, and treatment protocols. The assumption of conditional independence of the 2 diagnostic methods required assessment before use of the model. In theory, identification of illness based on both CI and LU are diagnostic methods that measure a similar biological phenomenon, and the relationship between

the 2 techniques could be dependent (correlated) within each subpopulation.[9] Therefore, it was assumed that the diagnostic methods were correlated, and a conditional dependence model for the 2 tests was used to estimate the correlation between Se and Sp for each test.[2] Conditional dependence was rejected, because the 95% PI for the correlations of Se and Sp between the tests included zero.[9]

Analysis was augmented by running models to evaluate Se to prior assumptions (Table 2). The similarity of findings among the models indicates the robustness of the initial model and the relative insensitivity to the specified prior distributions. Therefore, the final results were not influenced by these estimations, and results from model 1 (by using informative priors) will be discussed.

The relatively low Se and Sp of the CI method were somewhat expected, because diagnosis of BRD based on clinical signs is a subjective process, even when observations are combined with rectal temperature. Identifying a single sick animal in a population can be very challenging because of the prey behavior of cattle, which may cause them to hide signs of illness. The number of animals deemed sick in any population may also be influenced by the judgment of the animal caretaker viewing the calf. Because this is ultimately a very subjective evaluation, the accuracy of the diagnostic method may be highly variable, with the magnitude of error varying among individuals determining the diagnosis. Although this could affect the accuracy of diagnostic estimates for an individual observer, the findings in the present study may be applied as an estimate of overall diagnostic accuracy across all observers in the included studies. The Sp of

CI also may be influenced by the case definition of BRD. The disease is a syndrome and, as such, presents in a variety of manners with similar clinical signs as other diseases. Respiratory disease is the most common ailment in postweaned beef calves,[19] and, when clinical signs are unclear, animal health providers may attribute illness to the most common syndrome (BRD). Misdiagnosing other conditions as BRD based on clinical signs would lead to a decrease in the Sp of CI for BRD.

Clinical trials and preventive health programs are often evaluated by using BRD morbidity (measured by clinical observation) as an outcome to compare treatments or preventive protocols. The implications of the results of the current study are that the standard methods that use clinical observation combined with rectal temperature may not provide a very accurate representation of the disease frequency in populations. The relatively low Se (approximately 62%) of this diagnostic method means up to 38% of truly diseased animals may go undiagnosed in any given population. Perhaps more troubling is the relatively low Sp of CI (approximately 63%), which indicates that up to 37% of calves diagnosed with BRD and subsequently treated may not, in fact, be suffering from the syndrome. The poor Sp may explain why some treatments or interventions only appear to work on a segment of the population. Combined, the low Se and Sp of CI would increase the error associated with measures of morbidity in the postweaning phase, thereby making it more challenging to determine the efficacy and potential differences among treatment or prevention programs. The accuracy, or lack thereof, of this diagnostic method should be accounted for when determining measures of disease frequency and evaluating morbidity as an outcome in clinical trials and health program evaluations.

The model estimate for Se in the current study of using harvest lung lesions for diagnosing BRD was 77.4% (97.5% PI: 66.2, 87.3). This was lower than expected, but several hypotheses have been forwarded to explain apparent discrepancies between the presence of BRD and LU. Results of 2 previous studies suggested that the timing of the BRD episode during the feeding phase and differences in healing rates may influence the extent of the pulmonary lesions at harvest.[3,21] If cattle were sick early in the feeding phase and lesions were healed or minimal, they may not be detected at harvest, which would result in a low Se of the diagnostic method. This theory is supported by recent evidence that calves treated farther from harvest have improved performance compared with calves treated closer to harvest.[1] The ability of LU to retrospectively identify true BRD

cases may be limited by both the length of time from lesion appearance to evaluation and overall lesion magnitude (degree of long-term damage). Another potential explanation for the relatively low Se of LU is the potential response to treatment. If an animal is diagnosed early in the disease process and is given an effective treatment, then it is possible that the calf truly had BRD during the feeding phase, yet, because of the treatment, failed to possess pulmonary lesions at harvest. The 1999 study used for the analysis did not list specific components of the protocol to treat BRD cases; therefore, it is impossible to compare potential treatment effects between the 2 studies in this analysis.

The estimated Sp of LU for the current study was relatively high, 89.7% (97.5% PI: 86.0, 93.8). Pulmonary lesions are truly indicative of respiratory disease, yet, one reason LU Sp was not 100% in the present study may be associated to the case definition. The target case was BRD that occurred in the postweaning phase, and it is possible that a respiratory episode occurred before feedlot entrance and that lesions were still present at harvest. Because most BRD events occur during the feeding phase, it was not surprising that the Sp of LU was still relatively high, despite this potential mechanism for misclassification.

Models that directly compared the 2 diagnostic methods revealed higher Se and Sp for LU when compared with CI. The Se for LU illustrated a relatively wide 97.5% PI (66.2, 87.3), and model-generated Sp estimates for both methods displayed relatively narrow distributions. A portion of the difference in Sp between the diagnostic methods may be attributed to the fact that CI relies almost wholly on subjective observations of clinical signs, and these symptoms may be attributed to other disease syndromes in addition to BRD. Evaluation of pulmonary lesions at harvest provides insight to lesions attributed, by definition, to respiratory disease. Overall, the results of the current study indicate that LU was the superior diagnostic method for correctly classifying true BRD cases. However, it is not possible to use this technique in the field at the time of illness to make real-time treatment decisions. The choice of diagnostic method applied in a given situation depends on the objective of the testing (clinical research vs. real-time therapeutic decisions), but, with either diagnostic modality, the results should be interpreted in light of the limitations of each method.

The discrepancy between the diagnostic accuracy and relatively low agreement of the 2 methods may indicate that the tests are measuring unique biological processes, and the independence of the methods was supported by the present study. Another factor may

be the time lag between when the 2 diagnostic evaluations occur: CI is a near real-time measure of disease occurrence and LU is a retrospective measure of animal wellness status. However, the analysis was constructed to evaluate the diagnostic ability of the methods to determine if a calf had BRD at any time point during the feeding phase. Therefore, the findings should not be interpreted to imply that the estimated Se and Sp for CI are applicable at the time of diagnosis but rather they should be used in the final evaluation of the level of morbidity during the entire feeding phase based on repeated, daily evaluations of CI. These findings provide a framework for applying the 2 diagnostic methods on retrospective data and appropriately interpreting results to determine a better estimate of the true level of BRD in the population.

One limitation of the current research is the fact that data from only 2 previous studies were included in the analysis. Unfortunately, published literature that compared the observation of CI and pulmonary lesions at harvest for BRD diagnosis during the feeding phase is very sparse, as evidenced by the results of the structured literature search. Inclusion of only 2 study populations in the current analysis limits the external validity of the findings and subjects the current findings to any potential biases inherent in the previous research. Both reference populations had relatively high rates of clinical disease, and findings from the current study may not be as applicable to populations with low disease rates if diagnostic Se and Sp estimates are affected by disease rates in the populations. Therefore, extrapolating Se and Sp estimates for these diagnostic methods to other feedlot populations should be done with extreme caution. Despite potential limitations in external validity, the current research provides important initial estimates for accuracy of these diagnostic methods and provides an impetus for further research on the accuracy of these BRD diagnostics in different populations of feedlot cattle.

Research findings of the present study confirmed a common assumption that Se of clinical observations may be relatively low. In fact, both diagnostic methods displayed relatively low Se. The Sp estimates for the 2 methods indicated that Sp was relatively high for LU but low for CI. This finding is important, because use of the defined criteria for evaluation based on LU will result in relatively few false positives. Findings also indicate that, if only CI is used to monitor morbidity, then a high degree of diagnostic error should be expected. Although LU is a valuable postmortem diagnostic method with a high Sp, Se was relatively low, and this diagnostic method is inherently limited, because it can only be used after the feeding phase and not for real-time decision making. The accuracy of diagnostic methods must be considered when evaluating BRD morbidity in the postweaning phase, and more research should be conducted to identify methods to improve the ability to identify BRD in populations of cattle.

Acknowledgements

Sources and manufacturers

a. PEPI version 4.0 (http://www.brixtonhealth.com/).
b. Expert panel consisted of the following: Drs. Mike Apley, associate professor; Robert Larson, professor, Coleman Chair of Food Animal Production Medicine; Mike Sanderson, associate professor; and Dan Thomson, associate professor and Jones Professor of Production Medicine and Epidemiology, all from Clinical Sciences Department, College of Veterinary Medicine, Kansas State University, Manhattan, KS.
c. BetaBuster version 1.0 (www.epi.ucdavis.edu/diagnostictests/betabuster.html).

References

1. Babcock AH, White BJ, Dritz SS, et al.: 2009, Feedlot health and performance effects associated with the timing of respiratory disease treatment. J Anim Sci 87:314–327.
2. Branscum AJ, Gardner IA, Johnson WO: 2005, Estimation of diagnostic-test sensitivity and specificity through Bayesian modeling. Prev Vet Med 68:145–163.
3. Bryant LK, Perino LJ, Griffin D, et al.: 1999, A method for recording pulmonary lesions of beef calves at slaughter, and the association of lesions with average daily gain. Bovine Pract 33:163–173.
4. Edwards AJ: 1996, Respiratory diseases of feedlot cattle in the central USA. Bov Pract 30:5–7.
5. Elvinger F, Akey BL, Senne DA, et al.: 2007, Characteristics of diagnostic tests used in the 2002 low-pathogenicity avian influenza H7N2 outbreak in Virginia. J Vet Diagn Invest 19:341–348.
6. Enoe C, Georgiadis MP, Johnson WO: 2000, Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. Prev Vet Med 45:61–81.
7. Fulton RW, Cook BJ, Step DL, et al.: 2002, Evaluation of health status of calves and the impact on feedlot performance: assessment of a retained ownership program for postweaning calves. Can J Vet Res 66:173–180.
8. Gardner BA, Dolezal HG, Bryant LK, et al.: 1999, Health of finishing steers: effects on performance, carcass traits, and meat tenderness. J Anim Sci 77:3168–3175.
9. Gardner IA, Stryhn H, Lind P, Collins MT: 2000, Conditional dependence between tests affects the diagnosis and surveillance of animal diseases. Prev Vet Med 45:107–122.

10. Gelman A, Rubin DB: 1992, Inference from iterative simulation using multiple sequences. Stat Sci 7:457–511.
11. Larson RL: 2005, Effect of cattle disease on carcass traits. J Anim Sci 83:E37–43.
12. Lechtenberg KF, Smith RA, Stokka GL: 1998, Feedlot health and management. Vet Clin North Am Food Anim Pract 14:177–197.
13. Lunn D, Thomas A, Best N, Spiegelhalter D: 2000, WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput 10:325–337.
14. Mainar-Jaime RC, Barberan M: 2007, Evaluation of the diagnostic accuracy of the modified agglutination test (MAT) and an indirect ELISA for the detection of serum antibodies against Toxoplasma gondii in sheep through Bayesian approaches. Vet Parasitol 148:122–129.
15. Monti GE, Frankena K, Engel B, et al.: 2005, Evaluation of a new antibody-based enzyme-linked immunosorbent assay for the detection of bovine leukemia virus infection in dairy cattle. J Vet Diagn Invest 17:451–457.
16. Perez A, Cobo E, Martinez A, et al.: 2006, Bayesian estimation of *Tritrichomonas foetus* diagnostic test sensitivity and specificity in range beef bulls. Vet Parasitol 142:159–162.
17. Pinchak WE, Tolleson DR, McCloy M, et al.: 2004, Morbidity effects on productivity and profitability of stocker cattle grazing in the Southern Plains. J Anim Sci 82:2773–2779.
18. Sanderson MW, Dargatz DA, Wagner B: 2008, Risk factors for initial respiratory disease in United States' feedlots based on producer-collected daily morbidity counts. Can Vet J 49:373–378.
19. Smith RA: 1998, Impact of disease on feedlot performance: a review. J Anim Sci 76:272–274.
20. Smith RA, Stokka GL, Radostits O, Griffin D: 2001, Health and production management in beef feedlots. *In:* Herd health: food animal production medicine, ed. Radostits O, 3rd ed., pp. 592–595. WB Saunders, Philadelphia, PA.
21. Thompson PN, Stone A, Schultheiss WA: 2006, Use of treatment records and lung lesion scoring to estimate the effect of respiratory disease on growth during early and late finishing periods in South African feedlot cattle. J Anim Sci 84:488–498.
22. Thomson DU, White BJ: 2006, Backgrounding beef cattle. Vet Clin North Am Food Anim Pract 22:373–398.
23. Toft N, Jorgensen E, Hojsgaard S: 2005, Diagnosing diagnostic tests: evaluating the assumptions underlying the estimation of sensitivity and specificity in the absence of a gold standard. Prev Vet Med 68:19–33.
24. U.S. Department of Agriculture (USDA): 1998, National Health Monitoring System beef '97 part III: reference of 1997 beef cow-calf production management and disease control, ed. Animal and Plant Health Inspection Service, Veterinary Services, pp. 1–48. USDA, Fort Collins, CO.
25. White BJ, McReynolds S, Goehl DR, Renter DG: 2008, Effect of vaccination and weaning timing on backgrounding morbidity in preconditioned beef feeder calves. Bov Pract 42:1–5.
26. Wittum TE, Woollen NE, Perino LJ, Littledike ET: 1996, Relationships among treatment for respiratory tract disease, pulmonary lesions evident at slaughter, and rate of weight gain in feedlot cattle. J Am Vet Med Assoc 209:814–818.