# Determining associations and predictability of pen level management factors and health outcomes related to respiratory disease in the first 45 days of the feeding phase

by

Hector Anthony Rojas

B.S., Kansas State University, 2019

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Clinical Sciences
College of Veterinary Medicine

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2022

Approved by:

Major Professor
Dr. Bradley White

# Copyright

# Abstract

Bovine respiratory disease (BRD) is an important health issue in the beef cattle industry. This syndrome is a multifactorial disease and continues to be a leading cause of morbidity and economic loss in feedlot cattle operations. Research has identified risk factors associated with BRD; however, knowledge gaps remain. Three studies were conducted to evaluate potential associations and predictability of pen level management factors and their impacts on BRD-related health outcomes during the first 45 days on feed (DOF) in the feeding phase. The first study evaluated potential associations between pen level management factors, such as area per head and bunk space per head, combined with cattle attributes and their impact on BRD morbidity incidence in the first 45 DOF. Our results showed pen housing characteristics related to pen area per animal and bunk space per head were significantly associated ($P < 0.05$) with BRD incidence in the first 45 DOF; however, effects of these associations were modified by cattle attributes. The second study had two objectives. The first objective was to evaluate the diagnostic performance of five predictive models using area under the curve (AUC) to classify incoming groups of cattle into high- (≥15% BRD incidence 45 DOF) and low-risk (<15% BRD incidence 45 DOF) groups based on the BRD morbidity within the first 45 DOF. The types of models used for this analysis were logistic regression, decision tree, random forest, discriminant linear, and naïve Bayes models. The second objective was to evaluate the models using an economic analysis to determine if the models were economically advantageous compared to a control scenario that represents feedlot personnel classifying expected risk. AUC of the models ranged from .682 to .789 for with a random forest model having the highest AUC. The model with the best economic performance compared to the person classifying expected risk was determinant on the proportion of high-risk cohorts in the population. The decision tree model

displayed the highest potential economic advantage when the proportion of high-risk cohorts reaches approximately 40% or less compared to the person. The third study evaluated potential associations between pen level management factors related to the number of water sources, shared water sources, and shared fence lines combined with cattle attributes and their impact on BRD morbidity incidence in the first 45 DOF. Shared fence lines, number of water sources, and their interactions with cattle demographics included in the models, were significantly associated with BRD morbidity in the first 45 DOF ($P < 0.05$). The interaction between shared water sources and total cohort size at arrival was found to have statistical significance ($P < 0.05$), but did not display biological significance. Several pen level management factors were tested for associations and predictability related to BRD morbidity in the first 45 DOF. Understanding how these factors and their impact on BRD can potentially lead to increased animal welfare, increased sustainability, and lessen the burden of economic losses in the feedlot cattle industry.

# Table of Contents


List of Figures ... viii
List of Tables ... x
Acknowledgements ... xi
Chapter 1 - A literature review of risk factors that impact the risk of bovine respiratory disease in feedlot cattle ... 1
Introduction ... 1
Cohort-level cattle demographic risk factors associated with Bovine Respiratory Disease ... 2
Arrival Weight ... 2
Cohort Size ... 3
Sex ... 5
Time of Entry ... 6
Weather ... 8
Pen-level management factors associated with bovine respiratory disease incidence ... 9
Pen Area ... 9
Feed Bunk Space and Feeding Behavior ... 12
Water Sources and Management ... 14
Conclusions ... 16
References ... 16
Chapter 2 - Associations between pen management characteristics and bovine respiratory disease incidence in the first 45 days post-arrival in feedlot cattle ... 20
Abstract ... 20
Introduction ... 21
Materials and methods ... 22
Statistical Methods ... 25
Results ... 26
Discussion ... 28
Conclusion ... 33
Figures ... 34
Tables ... 38

Chapter 3 - Predicting bovine respiratory disease risk in feedlot cattle in the first 45 days post-arrival ... 41
Abstract ... 41
Introduction ... 42
Materials and methods ... 43
Data ... 43
Data Preparation ... 45
Data Partitioning ... 45
Model Recipe Creation ... 46
Classification Algorithms ... 46
Logistic Regression ... 46
Decision Tree ... 47
Random Forest ... 47
Naïve Bayes ... 47
Linear Discriminant ... 47
Resampling – cross-validation ... 48
Model optimization/tuning ... 48
Model evaluation ... 49
Economic analysis ... 50
Results ... 53
Descriptive Statistics ... 53
Model Performance Diagnostics ... 53
Area Under the Curve and Classification Accuracy ... 53
Sensitivity and Specificity ... 54
Positive/Negative Predictive Value ... 54
Economic Results ... 54
Discussion ... 55
Conclusions ... 61
Figures ... 63
Tables ... 66

Chapter 4 - Impact of water sources and shared fence lines on bovine respiratory disease incidence in the first 45 days on feed ........ 70
Abstract ........ 70
Introduction ........ 71
Materials and Methods ........ 72
Water Sources ........ 74
Shared Fence Lines ........ 74
Statistical Methods ........ 74
Results ........ 76
Discussion ........ 77
Conclusion ........ 81
Figures ........ 82
Tables ........ 85
Chapter 5 - Conclusions ........ 88
References ........ 91
Appendix A - Supplementary material ........ 96

# List of Figures

Figure 2.1. Histogram of the distribution of the level of morbidity of BRD incidence in the first 45 days on feed (DOF) in the study population. The x-axis displays the percentage of BRD incidence in a cohort during the first 45 DOF. The y-axis displays the count of cohorts for each value. ........ 34

Figure 2.2. Model estimated mean probability of BRD incidence by area per animal and arrival weight category in commercial feedlot cattle during the first 45 DOF. Error bars represent SE of least square means. ........ 35

Figure 2.3. Model estimated mean probability of BRD incidence by area per animal and size of the cohort at arrival in commercial feedlot cattle during the first 45 DOF. Error bars represent SE of least square means. ........ 36

Figure 2.4. Model estimated mean probability of BRD incidence by bunk space per animal and arrival weight category in commercial feedlot cattle during the first 45 DOF. Error bars represent SE of least square means. ........ 37

Figure 3.1. Flowchart of data refinement, data partitioning, algorithm training, and classification model algorithm evaluation. ........ 63

Figure 3.2. Flowchart of diagnostic outcomes and calculations generated from predictive classification models using cutoff of 15% BRD incidence in the first 45 DOF. ........ 64

Figure 3.3. Estimated economic results ($/animal) of the five classification model compared to a person (control) classifying expected BRD morbidity risk of incoming cohorts of cattle in the first 45 DOF across different proportions of high-risk cohorts to low-risk cohorts (0-100%). Gray line represents the control. At any prevalence, if the difference from the control for a model is above the control line ($0/animal) then it has a potential economic advantage relative to the control. If the difference from the control for a model is below the control line ($0/animal) then it has a potential economic disadvantage relative to the control. ........ 65

Figure 4.1. Histogram of the distribution of the level of BRD incidence in the first 45 days on feed (DOF) in the study population. The study population contained 1,563 cohorts which included 168,482 individual animals from 2018-2020. The x-axis displays the percentage of

BRD incidence in a cohort during the first 45 DOF. The y-axis displays the count of cohorts for each value. ........................................................................................................ 82

Figure 4.2. Model estimated probabilities for BRD incidence in the first 45 DOF by total water sources (one source/two sources) and cohort size at arrival category (25-99, 100-175, >175 animals in each cohort) in commercial feedlot cattle during the first 45 DOF using the NWmod results. Error bars represent SE of the probability. ............................................... 83

Figure 4.3. Model estimated probabilities for BRD incidence in the first 45 DOF by total water sources (one source/two sources) and average arrival weight category (500-599 lb., 600-699 lb., 700-799 lb., 800-899 lb., and 900-1,000 lb.) in commercial feedlot cattle during the first 45 DOF using the NWmod results. Error bars represent SE of the probability................... 84

# List of Tables

Table 2.1. Descriptive statistics for the study population cohorts (n = 1,733) from 10 feedlots. 38

Table 2.2. Distribution of variables used for analysis from 10 feedlots from 2015-2020 ........... 39

Table 2.3. Final generalized linear mixed-model demonstrating housing characteristics and cattle demographic factors and their association with bovine respiratory disease incidence during the first 45 DOF ........................................................................................... 40

Table 3.1. Predictor and Outcome variables used in analysis of classification algorithms ......... 66

Table 3.2. Variables included in the economic analysis to compare the cost-benefit of using one of the predictive models compared to the control scenarios ................................................ 67

Table 3.3. Classification model results and the financial consequence associated with following model predictions .............................................................................................................. 68

Table 3.4. Final diagnostic performance estimates utilizing the test dataset for BRD morbidity risk during the first 45 days on feed to classify cohorts as high- or low- risk for BRD development within the first 45 days post-arrival using a 15% cutoff at the optimum cutoff where 10.85% of the cohorts had >15% BRD morbidity ..................................................... 69

Table 4.1. Distribution of variables from the study population used for analysis from 10 feedlots from 2018-2020. The study population consisted of 1,563 cohorts which represented 168,482 animals. ............................................................................................................... 85

Table 4.2. Descriptive statistics for continuous variables in the study population cohorts (n = 1,563) from 10 feedlots. Data in study population consisted of 1,563 cohorts which represented 168,482 animals. ........................................................................................... 86

Table 4.3. Final generalized linear mixed-models demonstrating housing characteristics and cattle demographic factors and their association with bovine respiratory disease (BRD) incidence during the first 45 days on feed (DOF). The level of significance was set a P < 0.05 for each model. Three models (NWmod, SWmod, SFmod) for each hosing factor of interest (number of waters, shared waters, and shared fence lines) were evaluated. ............ 87

Acknowledgements

The opportunity to complete my graduate program and be a part of the Beef Cattle Institute for the past two years would not be possible without the guidance and support of Dr. Brad White. I am thankful to have had the privilege to work with Dr. White throughout my program. I would also like to Dr. David Amrine and Dr. Robert Larson for serving on my graduate committee. The advice and teaching that all three individuals have provided are invaluable and have made me a better professional and an even greater person.

I would like to thank the members of the Beef Cattle Institute for making this program more memorable thank I could ever imagine. From research meetings, traveling to conferences, and even just down time in the office, I thank you all for pushing me each day to improve. As I move on to the next part of my life, I know that I when I reflect back on my time with the Beef Cattle Institute that it will always be accompanied by a smile.

Finally, completion of my program would not be possible without the constant support of my family and friends both in Kansas and in Illinois. Balancing two lives that are over 600 miles apart has been difficult, but the everyday encouragement of those close to me made it much easier to handle.

# Chapter 1 - A literature review of risk factors that impact the risk of bovine respiratory disease in feedlot cattle

## Introduction

Despite advancements in management and treatment protocols over the years bovine respiratory disease (BRD) remains the primary cause of morbidity and mortality in feedlot cattle (Griffin 1997; Edwards, 2010). It is also one of the costliest diseases attributing approximately between $800 and $900 million in losses annually to the American feedlot industry (Chirase and Greene 2001). The treatment costs associated for each individual case of BRD has nearly doubled from $12.59 per case to $23.60 per case in the United States from 1999 to 2011 (USDA, 2013b). In 2011, 16.2% of cattle in feedlots with a capacity of 1,000 head or more were affected by respiratory disease after feedlot arrival (USDA, 2013a).

The risk factors that predispose feedlot cattle to BRD are complex as it is a multifactorial disease and is the result of a combination of complex interactions related to host, environment, and agent factors that predispose cattle to pneumonia (Cusack et al., 2003). The epidemiology of BRD has been studied previously and led to an improved understanding of the risk factors associated with BRD in feedlot cattle operations (Snowder et al., 2006; Sanderson et al., 2008; Taylor et al., 2010; Avra et al., 2017). The majority of BRD cases in feedlot cattle occur within the first 40 days post-arrival, although the timing of disease is variable dependent on numerous factors (Babcock et al., 2009). While several risk factors have been identified related to BRD there are still gaps in knowledge about risk factors related to BRD incidence in feedlot cattle as much of the variability associated with the disease and health outcomes is not fully understood. The objective of this literature review is to evaluate previous research of risk factors and related to BRD health outcomes in feedlot cattle.

## Cohort-level cattle demographic risk factors associated with Bovine Respiratory Disease

### Arrival Weight

Studies have identified cohorts with lower body weight entering a feedlot are at higher risk for onset of BRD compared to cohorts of higher body weights. (Sanderson et al., 2008; Reinhardt et al., 2009; Cernicchiaro et al., 2012a; Babcock et al., 2013). Arrival weight is determined as a cohort level variable and is often reported as the average weight of all animals within the cohort. Reinhardt et al. (2009) found that the average number of respiratory treatments per animal for steers and heifers weighing <226kg (0.52 treatments, 0.34 treatments), and steers and heifers weighing between 227-272kg (0.31 treatments, 0.24 treatments) were significantly higher than cattle in heavier weight categories: 272-317kg (0.22 treatments, 0.14 treatments), 318-362kg (0.21 treatments, 0.13 treatments), ≥363kg (0.16 treatments, 0.06 treatments). Cernicchiaro et al (2012a) used three different sizes of cattle (227kg to 271.9kg, 272kg to 317.9kg, and 318kg to 363kg) with 318kg – 363kg used as the referent category to assess if there was an association between BW class and BRD incidence in the first 45 DOF. The two lightest categories had a higher risk of BRD morbidity. However, the middle category displayed a higher risk (OR = 1.43, OR 95%CI = (1.23 to 1.68)) than the lightest category (OR = 1.27, OR 95%CI = (1.07 to 1.50)).

Babcock et al. (2013) investigated the effects of cohort-level risk factors and their associations towards total morality. Lightweight males and females (<182 kg) had greater mortality than middleweight classes (182 to 272 kg) and heavyweight classes (>271 kg) although the magnitude of difference depended on the month of the year. Sanderson et al. (2008) categorized arrival weight into three categories (< 250 kg, 250 to318 kg, >318 kg) using <250kg as the referent category. There was a decreased BRD incidence risk for cattle in the heavier arrival weight category (>318 kg; IRR = 0.18; 95% CI = (0.08 to 0.38) and no difference in BRD incidence risk in the middleweight category (250 to 318kg; IRR = 0.52; 95% CI = (0.25 to 1.10). These research reports display that lightweight calves in the feedlot industry are at higher risk for BRD morbidity compared to heavier weighted cattle and that feedlots should expect to see increased morbid calves due to respiratory disease.

## Cohort Size

The cohort size represents the number of cattle grouped together at feedlot arrival. An Australian report looked at the effect that the number of animals in a group at arrival (< 50, 50 to99, and ≥ 100) had on BRD incidence during the first 50 days on feed (Hay et al. 2014). Results showed that the group size was significantly associated with BRD incidence in the first 50 DOF. Animals in small groups displayed a greater cumulative BRD incidence in the first 50 DOF (24.1%) than animals in middle-sized groups (21.3%) and large-sized groups (6.9%). The large-sized groups appeared to have much lower BRD incidence compared to animals in middle- and small-sized groups which may be a result of how these cattle were managed depending on the expected risk of these cattle.

An additional study investigated cohort size as a risk factors towards BRD incidence in the first 45 DOF. (Cernicchiaro et al., 2012a). Categories for cohort size in this were <91 cattle, 91 to 138 cattle, 139 to 202 cattle, and >202 cattle, with <91 cattle used as the referent category. Results from this study varied as BRD incidence risk increased as cohort size increased to 91 to 138 cattle (OR = 1.31, OR 95%CI = (1.21 to 1,41)) and decreased when cohort size increased to >202 cattle (OR = 0.73, OR 95%CI = (0.67 to 0.80)). There was no difference in BRD incidence risk when the cohort size increased to 139 to 202 cattle (OR = 0.94, OR 95%CI = (0.87 to 1.02)). Another study also looked at the cohort size (categorized as 20 to 100, 101 to 150, 151 to 200, ≥ 200) and found that cohort size was significantly associated with BRD morbidity risk in feedlot cattle (Cernicchiaro et al., 2012b). While this study used different categories, it also found that the risk of BRD morbidity continued to decrease at a steady rate as the cohort size increased. Using a cohort size of 20 to 100 as the referent category, 101 to 150 cattle (IRR = 0.77, 95% CI = (0.75 to 0.80)), 151 to 200 cattle (IRR = 0.64, CI 95% = (0.61 to 0.66)), and >200 cattle (IRR = 0.43, CI 95% = (0.42 to 0.45)) all had reduced risk of BRD incidence compared to the smallest cohort size analyzed in the study. An additional report from Australia reported that the size of the cohort did not affect the risk of BRD incidence in the first 50 DOF (Hay et al., 2016). The categories for cohort size that were investigated in this study were <175, 175 to <325, and ≥325 cattle with <175 as the referent category. Cohort sizes with 175 to <325 did not display a difference in BRD incidence in the first 50 DOF (OR = 1.6; 95% CI = (0.9-2.6)). Cohort sizes with ≥325 cattle also did not display an increased risk in BRD incidence in the first 50 DOF (OR =1.3; 95% CI = (0.6 to 2.4)). The size of the cohort could be a proxy for other characteristics related to feedlot management such as commingling. Several studies have shown that increased commingling plays a role in increased BRD incidence in feedlot cattle as it is a major stressor experienced by receiving cattle when cattle from multiple sources are commingled (Martin et al., 1982, Ribble et al., 1998; Step et al., 2008; Taylor et al., 2010). As group size decreases this may potentially lead to less commingling which may lead to decreased risk of respiratory disease. Commingling may play a role in the effect of large cohort sizes decreasing the risk of BRD incidence as larger cohort sizes may originate from a single large ranch whereas a smaller group may consist of cattle that originated from a larger number of farms which increased commingling. As cattle become increasingly commingled and are crowded into a pen space the risk of increased transmission of pathogens that are related to respiratory disease (*Mannheimia haemolytica*, *Bovine Viral Diarrhea Virus*, etc.) may also increase. Overall the effect of cohort size differs between studies and it may be possible that the effect of cohort size is modified by additional factors that were not investigated in previous studies.

### Sex

Sex is a cohort-level risk factor that has been investigated towards BRD. Sanderson et al., (2008) investigated the effect of sex on BRD incidence by comparing the effect of steers, heifers, and mixed heifers/steers pens on BRD incidence, with steers as the referent category. The results reported that mixed pens were associated with a higher incidence of BRD (IRR = 3.7, 95% CI = (1.9, 7.2)) when compared to steer pens and heifer pens displayed no difference compared to steer pens (IRR = 1.3, 95% CI = (0.8, 2.0)). Cernicchiaro et al. (2012b) utilized multivariable models and determined that sex was associated with BRD incidence in feedlot cattle and that males (IRR = 1.18, 95% CI = (1.12 to 1.24)) had an increased risk of BRD when compared to females. Another study by Cernicchiaro et al. (2012a) also demonstrated that heifers had a decreased risk of BRD incidence (OR = 0.87, 95% CI = (0.83 to 0.92)) compared to steers. This contrasts from the previous report by Sanderson et al (2008) as the previous study found that steers and heifers had similar risk of BRD incidence; however, the report by Sanderson et al. were different in that it also evaluated pens that were mixed with heifers which is an effect on sex that was not evaluated in the other studies. An additional study also reported that steers (20% BRD incidence) had a higher BRD incidence compared to heifers (14% BRD incidence) (Snowder et al., 2006). Avra et al., (2017) reported that sex was significantly associated with

BRD treatment failure, but did not find any differences in BRD risk between heifers and steer cohorts which is consistent with results from Sanderson et al. (2008). Lonergan et al. (2001) used feedlot data from the late 1990's and reported that heifers had a higher risk of BRD compared to steers which is different than results from more recent studies.

The different results between studies could be related to other factors that impact risk of BRD in feedlot cattle. Interactions between sex and other variables were assessed in the previous work, but only an interaction between sex and shrink was found to be the only interaction significantly associated with sex included (Cernicchiaro et al., 2012b). An additional reason as to why some studies found higher BRD risk in steers compared to heifers is castration. Males may be at greater risk for BRD due to castration before or after arrival to a feedlot (Taylor et al., 2010). Increased risk of BRD morbidity may also be attributed to "buller" syndrome where a steer can be mounted and ridden by pen mates repeatedly and can potentially lead to increased BRD morbidity. Although sex has been found to be associated with an increase in respiratory disease, there are other factors that modify this effect and should always be considered when evaluating the association of sex and BRD.

### Time of Entry

Several studies have determined that cattle entering the feedlot in the fall are at higher risk for BRD morbidity compared to cattle entering at other seasons of the year. Ribble et al. (1995) used feedlot data from 58,885 spring born calves from 1985 to 1988 and found a higher risk of BRD for feedlot cattle arriving in the fall compared to other seasons. A retrospective cohort study using United States feedlot data between the years 1994-1999 found a higher monthly mortality for all causes for animals that entered the feedlot in the fall months. (Loneragan et al., 2001).

Cernicchiaro et al (2012b) used a multivariable mixed-effects regression model to assess the effects of body weight loss during transit (shrink), on health and performance using retrospective data from cattle arriving to feedlots during the years 2000 to 2008 and included season of arrival in the model. Results showed the effect of shrink on BRD morbidity was dependent on season of arrival. Four seasons (winter, spring, summer, fall) were evaluated with winter used as the referent category. Cattle arriving in the fall and summer months showed a significantly greater BRD morbidity across all shrink categories (< 0%, 0 to 2.5%, 2.6 to 5.0%, and >5.0%) compared to cattle who arrived in the winter. Spring cattle displayed a greater risk for BRD compared to cattle arriving in the winter when shrinkage was <0% and 0 to 2.5%. This is consistent with other studies showing that cattle arriving in the fall are at higher risk for BRD, but also displays that summer cattle are at higher risk for respiratory disease as well. In addition, this study also shows that other risk factors such as shrinkage may interact with the impact of season of arrival on BRD risk. Wisnieski et al. (2021) investigated BRD mortality incidence in the first 60 DOF using a mixed-effects negative binomial regression model and reported that summer and fall had more BRD mortality incidence compared to spring and winter.

Overall, the impact time of entry on BRD risk appears to be influenced by other factors related to incoming cattle and management. One impact is cattle age as freshly weaned calves are reported to be at higher risk of morbidity than yearling cattle. Weight is normally a proxy for age and freshly weaned lightweight cattle often arrive to feedlots in the fall season while heavier yearling animals are more likely to arrive during spring months. The season of entry could also be a proxy for many other factors such as low temperature weather in the winter and high heat and prolonged dehydration in the summer. High numbers of cattle entering the feedlot and long work hours during certain seasons of the year could also lead to human factors, such as delays in processing that lead to increased cattle stress which ultimately can impact the risk of BRD in cattle arriving to feedlots. Time of entry is a factor that has been documented to have an impact on the risk of BRD in feedlot cattle. Therefore, it must be taken into consideration when evaluating the expected risk of BRD in feedlot cattle.

## Weather

Weather has also been investigated as a factor that may be associated with BRD incidence and can be modified by time of entry. An Australian study found that the lower the minimum daily ambient temperature, and greater the range of the ambient temperature, the higher the incidence of daily BRD treatments (Cusack et al. 2007). However, Cusack et al. (2007) did not adjust for other covariates when reporting the associations between weather and BRD incidence. Another study utilized a multivariable model to assess the impacts of weather on BRD incidence of 1,904 cohorts of feedlot cattle (Cernicchiaro et al., 2012a). The model found that several weather factors were significantly associated with BRD incidence, but were modified by several cattle demographic factors including month of arrival. The arrival month modified the effect of maximum wind speed recorded in a 5 to 7 day lag on BRD incidence with cattle arriving in September and October had higher odds of BRD incidence than cattle arriving in November with a wind speed of ≥ 32.2 kph. Wisnieski et al. (2020) also reported that absolute air temperature change was significantly associated with BRD incidence and cattle that experienced a temperature change between 1.9°C to < 5.3 °C had the greatest BRD mortality incidence compared to cattle that experienced a temperature change of <1.9°C and > 9.4 °C. This is inconsistent with the findings from Cusack et al., (2007) who reported the greater the range on ambient temperature, the higher the incidence of BRD treatments.

It cannot always be concluded that weather conditions at the time of treatment are the primary cause of increased BRD incidence. For example, weather has been linked to increased BRD in the fall season, but the weather at that time is not the only predominate factors that leads to respiratory disease at that moment. Other factors such as an increased marketing of cattle and increase in commingling and transport of high-risk calves can add stress to animals and increase the risk of BRD in addition to weather factors. Acclimation to a new environment, social order, and processing procedures can also can also modify the association of weather on BRD. Weather can also be highly variable and can be dependent on large changes in air temperature in multiple geographic. Weather has been found to be associated with BRD risk, but other factors related to cattle demographics and management should be taken into consideration first when determining BRD morbidity risk in feedlot cattle.

## Pen-level management factors associated with bovine respiratory disease incidence

### Pen Area

Pen area is the total amount of space allocated to cattle within a pen. This factor can also be represented as the pen density which is defined as the pen area available per individual animal. Sanderson et al. (2008) using a multivariable model found no significant (*P* < 0.05) association of pen density and BRD morbidity. An Australian study further investigated pen density (m2/standard cattle unit (equivalent to an animal with a live-weight of 600kg)) (Hay et al., 2016). This variable was tested through multilevel mixed effects logistic regression models to evaluate if pen density was associated and a modifiable variable towards risk BRD morbidity in the first 50 DOF. Pen density was also categorized into two separate categories (11 to < 15 $m^2$/standard cattle unit, ≥ 15 $m^2$/standard cattle unit). No differences in BRD 50-day incidence risk were found between the two categories which is a similar finding to Sanderson et al. (2008).

A prospective study out of Brazil aimed to determine if space allowance per animal was a management factor that may be associated with respiratory disease symptoms (Macitelli et al., 2020). 1,350 bulls were evaluated for approximately 3 months where the first 6 weeks were defined as a dry period and the last 6 weeks as a rainy period. These animals were analyzed for health indicators (nasal and ocular discharge). Three different sized pen spaces were the treatment for this study (6$m^2$/animal, 12$m^2$/animal, and 24 $m^2$/animal) with 450 animals per treatment and three pens of each treatment (9 pens total). Each pen had a similar length of feed bunk space (33cm/animal). A mixed-effects model was used to assess the effects of each treatment and found a greater number of sneezes, in the treatment with the smallest area per animal at 6$m^2$/animal pens (1.2 sneezes/minute) compared to the 12$m^2$/animal (0.5 sneezes/minute) and 24 $m^2$/animal pens (0.2 sneezes/minute) in the dry period, but no differences in health indicators were found in the rainy period across all categories of pen area per animal. In addition, 6$m^2$/animal pens had a greater prevalence of coughing (0.3 coughs/minute) compared to pens with 12$m^2$/animal (0.1 coughs/minute) and animals in pens with 24$m^2$/animal (0.0 coughs/minute). Nasal discharge was also higher in animals with 6$m^2$/animal (12.2%) compared to animals with 24$m^2$/animal (3.4%), but was not statistically different compared to animals with 12$m^2$/animal (7.3%). No difference in difficulty breathing was found among any of the treatment groups. This study contrasts from other studies as it reports significant differences between different pen area sizes, but it only reports health indicator symptoms and not respiratory disease as Sanderson et al. (2008) and Hay et al. (2016) were evaluating. They also used different category cutoffs for the pen area sizes compared to the other studies.

Rojas et al. (2022) evaluated the association of pen area per head, and how cattle demographics impacted its effect, on BRD morbidity in feedlot cattle in the first 45 DOF in a feedlot setting. Three different sized pen spaces were the treatment for this study (≤23.22sq. m, 23.23-32.52 sq. m, and >32.52 sq. m) in accordance with management recommendations published from Boyer et al. (1997). Utilizing generalized linear mixed models an association was found between pen area per head and BRD incidence. Several interactions between pen area per head and cattle demographics were also significantly associated with BRD incidence in the first 45 DOF. The average arrival weight of the cohort modified the effect of area per animal on BRD incidence where cohorts with an average arrival weight of 409-453 kg had a higher probability of BRD incidence 3.6% ± 0.64) if they were placed in pens that had ≤ 23.22 sq. m of area per animal available compared to similar weight cattle placed in pens that had 23.23-32.52 sq. m (25-99, 2.05% ± 0.39), or > 32.52 sq. m (25-99, 2.47% ± 0.47) of area per animal available. Total cohort size at arrival also modified the effect of area per animal on BRD incidence. The probability of BRD incidence was higher in cohort sizes greater than 175 animal when given less than or equal to 23.22 sq. m of area per animal (4.46% ± 0.83) compared to similar cohort sizes with 23.23-32.52 sq. m per animal (2.83% ± 0.54) or greater than 32.52 sq. m per animal available (2.40% ± 0.46). Previous studies have found that pen area/pen density was not associated with risk of BRD incidence, but results from this study show that it impacted the risk slightly when it is evaluated with interactions with cattle demographics. Overall, pen area has been shown to not impact BRD risk in feedlot cattle, but when it is evaluated with other risk factors commonly associated with BRD incidence there are small effects that impact the risk when considering cattle weight and cohort size.

## Feed Bunk Space and Feeding Behavior

The feed bunk is a long trough used to feed cattle throughout the feeding phase. Feed bunk management is defined as delivering consistent feed ration designed to optimize dry matter intake, reduce feed waste, improve feeding behavior, and ensure adequate feed bunk space. Several studies have attempted to investigate feed bunk behavior as a potential factor associated with BRD morbidity. Two separate 32 day feeding trials were performed to determine if monitoring feeding behavior (total time spent within 50 cm of the feed bunk) of incoming steers with cumulative summation (CUSUM) procedures could detect morbidity in a commercial feedlot (Quimby et al., 2001). The results displayed that using the CUSUM feeding behavior detected BRD morbidity compared with detection of animals deemed morbid by experienced pen riders approximately 4.1 days earlier with an average sensitivity, positive predictive value, and accuracy of 90%, 91%, and 86%, respectively, between the two studies and feeding behavior can be useful in detecting BRD morbidity. Kayser et al. (2018) also evaluated the effectiveness and accuracy of CUSUM procedures to monitor feeding behavior patterns to predict onset of BRD. A full multivariate model was analyzed with the variables dry matter intake, bunk visit frequency, bunk visit duration, head down duration, eating rate, maximal nonfeeding interval, SD of nonfeeding interval, and time to bunk. Results indicated that the model had an accuracy of 75% and sensitivity of 70% which is lower than the accuracy and sensitivity reported by Quimby et al. (2001). Another study elaborated feeding behavior towards BRD and used 213 steers to investigate if daily feeding behavior (i.e., feeding times, meal intake and frequency) could be used to detect BRD earlier than through visual observation (Wolfger et al., 2015). The study found that several feeding behavior variables were significantly associated with a decreased hazard for developing BRD 7 days before visual identification. BRD hazard decreased as mean intake per meal, mean feeding time, frequency, and inter-meal interval increased. Each of these studies found value in using feeding behavior as a tool to detect BRD early instead of utilizing visual signs normally assessed by pen riders. Benefits of using feeding behaviors and technologies to detect feed behaviors may include detecting morbidity sooner and more accurately and saving costs towards treatment and performance loss. However, the ability of pen-riders to identify morbid animals varies between different feedlot operations and the cost of installing and utilizing technologies to detect feeding patterns may not offer an economic advantage compared to the feedlot staff at for producers at certain operations.

Since feeding behaviors have been found to be associated with BRD morbidity, it has been hypothesized that the amount of bunk space allocated to cattle may contribute to risk of BRD. Research has been conducted to assess if the amount of bunk space given has an effect on BRD incidence as the amount of bunk space per head would inherently be tied to the cohort level. Hay et al. (2016) tested the potential association of bunk space (m/head) using a multilevel mixed effects logistic regression models on BRD incidence in the first 50 DOF. Bunk space was categorized into two categories (<0.18m, ≥0.18m) and no significant differences were found between the two categories. Rojas et al. (2021) evaluated the association of bunk space per head at the cohort-level, and its interactions with demographics, on BRD morbidity in feedlot cattle in the first 45 DOF. Three different categories of bunk space were investigated (≤0.3 m, 0.31-0.46 m, >0.46 m) in accordance with management recommendations published from Boyer et al. (1997). Utilizing generalized linear mixed models, it was determined that bunk space was associated with BRD morbidity risk in the first 45 DOF, but only the interaction between bunk space and average arrival weight displayed a meaningful difference. A small difference was seen between cohorts with an average arrival weight between 409-453 kg as cohorts in this weight category had a lower probability of BRD incidence if they were provided less than or equal to 0.3 m of bunk space per animal (1.91% ± 0.37) compared to similar weight cattle placed in pens that had 0.31-0.46 m (3.34% ± 0.63) or greater than 0.46 m of bunk space per animal available (2.86% ± 0.54). All other interactions and the main effect of bunk space per animal were significantly associated from the model results but displayed no meaningful differences.

Overall, feeding behavior has been found to be associated with BRD morbidity in previous studies and can be used as a factor to monitor BRD risk. In addition, the amount of bunk space allocated to animals has been determined to not have much of a significant impact on BRD risk at both the individual- and cohort-level. As a result, the amount of bunk space allocated to cattle does not appear to be useful in determining the risk of BRD in feedlot cattle operations.

## Water Sources and Management

Water management and the number of drinking water troughs are factors that many feedlots pay attention to in order to ensure cattle have access to and receive adequate quantities of drinking water. Previous work has hypothesized that the number of water troughs could be a factor that can be related to feedlot management and risk of BRD. Hay et al. (2016) attempted to quantify the effects of numerous factors related to water sources and determine if they were significantly associated with risk of BRD. Shared pen water (Yes/No) and pen shade (Yes/No) were binomial categorical variables included in a multilevel mixed effects logistic regression model to determine if these variables were associated with the risk of BRD in the first 50 DOF. No significant differences were seen between the pen shade categories; however, a significant difference was seen between the shared pen water categories as sharing pen water appeared to increase BRD incidence compared to not sharing (OR = 4.3, 95% CI = 1.4 to 10.3). The difference in risk between the shared pen water categories was hypothesized that the shared pen water will increase the risk of BRD as animals in a pen share the same water source and thus are more susceptible to spread of respiratory pathogens. Rojas et al. (2022) utilized general linearized mixed models to evaluate the impact of two binomial variables: total water sources per pen (0 = one total water source; 1 = two water sources) and shared pen water sources (0 = no shared water sources; 1 = shared water sources) with BRD incidence in the first 45 DOF. The interactions between the water variables and how cattle demographics impact their influence on BRD morbidity were also assessed. Cohort size (25-99, 100-175, >175) was included as a variable in the model and the interaction between cohort size and total water sources per pen was found to be significantly associated with BRD incidence in the first 45 DOF. Differences were seen in the 100-175 category as cattle with one water source had a higher probability of BRD incidence in the first 45 DOF (5.50% ± 0.97) compared to cattle with two water sources (3.11% ± 0.05). Average arrival weight (227-272kg, 273-318kg, 319-363kg, 364-408kg, 409-453kg) was also included and the interaction between average arrival weight and total water sources per pen was significantly associated with BRD incidence in the first 45 DOF. Cattle arriving between 227-273kg displayed a higher risk for BRD incidence in the first 45 DOF when given one water source (8.8% ± 1.5) compared to similar weighted cattle given two water sources (5.5% ± 0.10). Cattle arriving between 273-318kg displayed a lower risk for BRD incidence in the first 45 DOF when given one water source (8.17% ± 1.4) to similar weighted cattle given two water sources

(11.60% ± 1.92). This differs with the results from Hay et al. (2016) as this study reports the interactions between total water sources per pen and factors related to cattle demographics. Rojas et al. (2022) reports that the total number of water sources may have an impact on BRD incidence, but it is dependent on its interactions with other cattle demographic risk factors. Further research must be conducted in order to determine the usefulness of modifying water sources as a management strategy in order to manage BRD risk in feedlot cattle.

## Conclusions

Several studies have and identified risk factors related to BRD in feedlot cattle related to cattle demographics and management factors. The pen-level factors of interest in our studies that we investigated were pen area per head, bunk space per head, number of total/shared waters sources, and shared fence lines. Our goal was to identify potential associations between these factors and cattle demographics that may influence BRD risk in feedlot cattle. The knowledge gained from these studies will allow further quantification of the effects of pen-level factors on health outcomes related to BRD morbidity within the first 45 DOF in the feedlot. In addition to analyzing the associations of these variables towards BRD, the predictive power of the variables included in the analyses will also be assessed. Identifying novel information and results regarding pen-level management variables will allow further knowledge towards management strategies that may be valuable to managing BRD risk during the feedlot period.

## References

Avra, T. D., Abell, K. M., Shane, D. D., Theurer, M. E., Larson, R. L., & White, B. J. (2017). A retrospective analysis of risk factors associated with bovine respiratory disease treatment failure in feedlot cattle1. *Journal of Animal Science*, *95*(4), 1521–1527. https://doi.org/10.2527/jas.2016.1254

Babcock, A. H., Cernicchiaro, N., White, B. J., Dubnicka, S. R., Thomson, D. U., Ives, S. E., Scott, H. M., Milliken, G. A., & Renter, D. G. (2013). A multivariable assessment quantifying effects of cohort-level factors associated with combined mortality and culling risk in cohorts of U.S. commercial feedlot cattle. *Preventive Veterinary Medicine*, *108*(1), 38–46. https://doi.org/10.1016/j.prevetmed.2012.07.008

Babcock, A. H., White, B. J., Dritz, S. S., Thomson, D. U., & Renter, D. G. (2009). Feedlot health and performance effects associated with the timing of respiratory disease treatment1. *Journal of Animal Science*, *87*(1), 314–327. https://doi.org/10.2527/jas.2008-1201

Boyer, W., Davidson, J., George, H., Graber, R., Murphy, J. P., & Harner, J. (2017). Guidelines for Planning Cattle Feedlots. *Kansas State University Research & Exetension*. https://bookstore.ksre.ksu.edu/pubs/MF3392.pdf

Cernicchiaro, N., Renter, D. G., White, B. J., Babcock, A. H., & Fox, J. T. (2012a). Associations between weather conditions during the first 45 days after feedlot arrival and daily respiratory disease risks in autumn-placed feeder cattle in the United States1. *Journal of Animal Science*, *90*(4), 1328–1337. https://doi.org/10.2527/jas.2011-4657

Cernicchiaro, N., White, B. J., Renter, D. G., Babcock, A. H., Kelly, L., & Slattery, R. (2012b). Associations between the distance traveled from sale barns to commercial feedlots in the United States and overall performance, risk of respiratory disease, and cumulative mortality in feeder cattle during 1997 to 20091. *Journal of Animal Science*, *90*(6), 1929–1939. https://doi.org/10.2527/jas.2011-4599

Chirase, N. K., & Greene, L. W. (2001). Dietary zinc and manganese sources administered from the fetal stage onwards affect immune response of transit stressed and virus infected offspring steer calves. *Animal Feed Science and Technology*, *93*(3–4), 217–228. https://doi.org/10.1016/S0377-8401(01)00277-2

Cusack, P., McMeniman, N., & Lean, I. (2003). The medicine and epidemiology of bovine respiratory disease in feedlots. *Australian Veterinary Journal*, *81*(8), 480–487. https://doi.org/10.1111/j.1751-0813.2003.tb13367.x

Cusack, P., McMeniman, N., & Lean, I. (2007). Feedlot entry characteristics and climate: Their relationship with cattle growth rate, bovine respiratory disease and mortality. *Australian Veterinary Journal*, *85*(8), 311–316. https://doi.org/10.1111/j.1751-0813.2007.00184.x

Griffin, D. (1997). Economic impact associated with respiratory disease in beef cattle. *The Veterinary Clinics of North America. Food Animal Practice*, *13*(3), 367–377. https://doi.org/10.1016/s0749-0720(15)30302-9

Griffin, D., Chengappa, M. M., Kuszak, J., & McVey, D. S. (2010). Bacterial Pathogens of the Bovine Respiratory Disease Complex. *Veterinary Clinics of North America: Food Animal Practice*, *26*(2), 381–394. https://doi.org/10.1016/j.cvfa.2010.04.004

Hay, K. E., Barnes, T. S., Morton, J. M., Clements, A. C. A., & Mahony, T. J. (2014). Risk factors for bovine respiratory disease in Australian feedlot cattle: Use of a causal diagram-informed approach to estimate effects of animal mixing and movements before feedlot entry. *Preventive Veterinary Medicine*, *117*(1), 160–169. https://doi.org/10.1016/j.prevetmed.2014.07.001

Hay, K. E., Morton, J. M., Clements, A. C. A., Mahony, T. J., & Barnes, T. S. (2016). Associations between feedlot management practices and bovine respiratory disease in Australian feedlot cattle. *Preventive Veterinary Medicine*, *128*, 23–32. https://doi.org/10.1016/j.prevetmed.2016.03.017

Kayser, W. C., Carstens, G. E., Jackson, K. S., Pinchak, W. E., Banerjee, A., & Fu, Y. (2019). Evaluation of statistical process control procedures to monitor feeding behavior patterns and detect onset of bovine respiratory disease in growing bulls. *Journal of Animal Science*, *97*(3), 1158–1170. https://doi.org/10.1093/jas/sky486

Loneragan, G. H., Gould, D. H., Mason, G. L., Garry, F. B., Yost, G. S., Miles, D. G., Hoffman, B. W., & Mills, L. J. (2001). Involvement of microbial respiratory pathogens in acute interstitial pneumonia in feedlot cattle. *American Journal of Veterinary Research*, *62*(10), 1519–1524. https://doi.org/10.2460/ajvr.2001.62.1519

Macitelli, F., Braga, J. S., Gellatly, D., & Paranhos da Costa, M. J. R. (2020). Reduced space in outdoor feedlot impacts beef cattle welfare. *Animal*, *14*(12), 2588–2597. https://doi.org/10.1017/S1751731120001652

Martin, S. W., Meek, A. H., Davis, D. G., Johnson, J. A., & Curtis, R. A. (1998.). Factors Associated with Mortality and Treatment Costs in Feedlot Calves: The Bruce County Beef Project, Years 1978, 1979, 1980. *Canadian Journal of Comparative Medicine*, *46*(4), 341–349.

Quimby, W. F., Sowell, B. F., Bowman, J. G. P., Branine, M. E., Hubbert, M. E., & Sherwood, H. W. (2001). Application of feeding behaviour to predict morbidity of newly received calves in a commercial feedlot. *Canadian Journal of Animal Science*, *81*(3), 315–320. https://doi.org/10.4141/A00-098

Reinhardt, C. D., Busby, W. D., & Corah, L. R. (2009). Relationship of various incoming cattle traits with feedlot performance and carcass traits1. *Journal of Animal Science*, *87*(9), 3030–3042. https://doi.org/10.2527/jas.2008-1293

Ribble, C. S., Meek, A. H., Shewen, P. E., Jim, G. K., & Guichon, P. T. (1995). Effect of transportation on fatal fibrinous pneumonia and shrinkage in calves arriving at a large feedlot. *Journal of the American Veterinary Medical Association*, *207*(5), 612–615.

Ribble, C. S., Meek, A. H., Shoukri, M. M., Guichon, P. T., & Jim, G. K. (1998). Risk Factors Associated With Fatal Fibrinous Pneumonia (Shipping Fever) in Feedlot Calves. *American Association of Bovine Practitioners Proceedings of the Annual Conference*, *31*, 104–109.

Rojas, H. A., White, B. J., Amrine, D. E., Larson, R. L., Capik, S.F. (2022). Associations between pen management characteristics and bovine respiratory disease incidence in the first 45 days post-arrival in feedlot cattle. Unpublished manuscript.

Rojas, H. A., White, B. J., Amrine, D. E., Larson, R. L., Capik, S.F. (2022). Impact of Water sources and shared fence lines on bovine respiratory disease incidence in the first 45 days on feed. Unpublished manuscript.

Sanderson, M. W., Dargatz, D. A., & Wagner, B. A. (2008). Risk factors for initial respiratory disease in United States' feedlots based on producer-collected daily morbidity counts. *The Canadian Veterinary Journal*, *49*(4), 373–378.

Snowder, G. D., Van Vleck, L. D., Cundiff, L. V., & Bennett, G. L. (2006). Bovine respiratory disease in feedlot cattle: Environmental, genetic, and economic factors. *Journal of Animal Science*, *84*(8), 1999–2008. https://doi.org/10.2527/jas.2006-046

Step, D. L., Krehbiel, C. R., DePra, H. A., Cranston, J. J., Fulton, R. W., Kirkpatrick, J. G., Gill, D. R., Payton, M. E., Montelongo, M. A., & Confer, A. W. (2008). Effects of commingling beef calves from different sources and weaning protocols during a forty-two-day receiving period on performance and bovine respiratory disease1,2. *Journal of Animal Science*, *86*(11), 3146–3158. https://doi.org/10.2527/jas.2008-0883

Taylor, J. D., Fulton, R. W., Lehenbauer, T. W., Step, D. L., & Confer, A. W. (2010). The epidemiology of bovine respiratory disease: What is the evidence for predisposing factors? *The Canadian Veterinary Journal = La Revue Veterinaire Canadienne*, *51*(10), 1095–1102.

United States Department of Agriculture (2013b) Feedlot 2011. Part IV: health and health management on U.S. feedlots with capacity of 1,000 or more head: https://www.aphis.usda.gov/animal_health/nahms/feedlot/downloads/feedlot2011/Feed11_dr_Pa rtIV_1.pdf.

United States Department of Agriculture (2013b Types and Costs of Respiratory Disease Treatments in U.S. Feedlots. https://www.aphis.usda.gov/animal_health/nahms/feedlot/downloads/feedlot2011/Feed11_is_RespDis_1.pdf

Wisnieski, L. C., Amrine, D. E., Cernicchiaro, N., Sanderson, M. W., & Renter, D. G. (2021). Weather conditions associated with death attributed to bovine respiratory disease complex in high-risk auction market–sourced male beef calves. *American Journal of Veterinary Research*, *82*(8), 644–652. https://doi.org/10.2460/ajvr.82.8.644

Wolfger, B., Schwartzkopf-Genswein, K. S., Barkema, H. W., Pajor, E. A., Levy, M., & Orsel, K. (2014). Feeding behavior as an early predictor of bovine respiratory disease in North American feedlot systems. *Journal of Animal Science*, jas.2014-8030. https://doi.org/10.2527/jas.2014-8030

# Chapter 2 - Associations between pen management characteristics and bovine respiratory disease incidence in the first 45 days post-arrival in feedlot cattle


## Abstract

Associations between pen housing characteristics and bovine respiratory disease (BRD) in feedlot cattle have not been well studied. We aimed to evaluate potential relationships between pen area per animal and bunk space per animal, and first treatment BRD incidence during the first 45 days on feed (DOF) during the feedlot phase of production using retrospective data from 10 commercial feedlots. Pen area per animal and bunk space per animal were combined with retrospective feedlot data which included cattle attributes such as sex, arrival weight, cohort size at arrival, and arrival date. Generalized linear mixed models were used to evaluate potential associations between pen characteristics, cattle attributes, and BRD incidence in the first 45 DOF. Overall, cohorts had a mean 0.42 m (standard deviation 0.23 m) and a median 0.34 m of linear bunk space per animal and a mean 28.64 m2 (standard deviation of 55.40 m2) and a median of 24.68 m2 of pen area per animal. Our results demonstrated that pen area per animal and bunk space per animal were associated ($P < 0.05$) with BRD incidence in the first 45 DOF. The effects of these associations were also modified by cattle attributes. For example, cattle weighing between 409–453 kg at arrival had higher BRD incidence in the first 45 DOF in pens with less than 23.33 m2 of pen area per animal compared to cattle housed in pens with of 23.23-32.52 m2 of pen area per animal or over 32.52 m2 of pen area per animal. Results also demonstrated that the interactions between bunk space per head and average arrival weight were associated with BRD incidence in the first 45 DOF ($P < 0.05$). Cohorts of cattle with an

average arrival between 409-453kg displayed a lower risk of BRD risk when given less than or equal to 0.3m of bunk space per animal compared to groups given 0.31m-0.46m, and greater than 0.46m of bunk space per animal. Further research is needed to explore if pen- and yard-level characteristics are associated with BRD incidence in the first 45 DOF in other feedlot settings.

## Introduction

Bovine respiratory disease (BRD) is the most common disease that contributes to morbidity and mortality in feedlot cattle (United States Department of Agriculture, 2013b). The average antimicrobial treatment cost for a single case of BRD has also increased from $12.59 to $23.60 between 1999-2013 (United States Department of Agriculture, 2013). Annual costs for BRD were estimated to be greater than $500 million per year which includes costs of treatment and loss of production from the disease (Miles 2009). Management practices towards controlling BRD morbidity in feedlot cattle may be difficult as the disease etiology of BRD multifactorial disease meaning that there are several host-, environment-, and pathogen-level factors that contribute to risk of BRD.

Several factors are associated with increased BRD risk such as transport, commingling, body weight, weather, sex, and more (Snowder et al., 2006; Sanderson et al., 2008; Cernicchiaro et al., 2012; Taylor et al., 2010; Hay et al., 2014). While several risk factors towards BRD are documented, knowledge gaps still remain regarding potential associations that could influence BRD morbidity risk in feedlot cattle. Associations of risk factors related to cattle housing and feeding conditions such as pen area and bunk space have been previously investigated. A recent Australian report evaluated the effects of similar management factors and their associations with

BRD (Hay et al., 2016). However, the report did not look at the potential interactions between pen housing conditions and other known risk factors on BRD risk.

The objective of this retrospective study was to evaluate the potential associations between pen-level housing factors (pen area and bunk space), cattle attributes, and BRD incidence in the first 45 days on feed (DOF) in commercial feedlots. Quantifying the effects of housing parameters relative to BRD incidence is useful for cattle managers in order to be able to calculate potential cost-benefits and modify their current management techniques. For example, if a pen area per head is associated with increased BRD incidence, managers could estimate the cost of increasing/decreasing the pen area available to cattle on operational efficiency compared to the expected benefit of BRD reduction. Our goal was to find information regarding pen housing conditions that would fill important knowledge gaps and enhance understanding of management strategies that can be utilized by commercial feedlot operations to reduce BRD incidence.

## Materials and methods

Animal Care and Use Committee approval was not obtained for this study due to obtaining retrospective data from pre-existing commercial feedlot data.

Retrospective data from 10 midwestern (Kansas and Nebraska) feedlots in the US between January 1, 2015 and April 16, 2020 were collected for this study. These data included information routinely collected at the cohort (lot and pen) and individual animal levels. As some lots were made up of cattle housed in multiple physical pens, a variable representing a lot-pen combination was created and a cohort was defined as animals managed at this lot-pen level. Data that were not able to create a lot-pen combination were treated as missing data and were removed from analysis. Data available at the cohort level included demographic characteristics of the cattle such as sex, arrival date, average arrival weight, and cohort size at arrival (Table 2.1.). Individual animal data contained information on animal treatment events and was joined to cohort level data using the yard, lot, and pen where the event occurred. Bovine respiratory disease incidence, our outcome, was defined as the number of cattle that were treated at least once for BRD based on feedlot diagnosis within the first 45 DOF divided by the size of the cohort. The case definition for a BRD treatment was any animal that received an antimicrobial treatment for BRD during the first 45 DOF. Cases were limited to 1st treatments only, and any additional treatments were excluded from analysis. If an animal was treated more than once, the first treatment record was utilized. As cattle within a cohort could be moved between pens, the DOF for cohorts in each distinct pen were calculated. To minimize cattle movement in our data only cohorts that were housed in 2 or fewer pens within the first 45 DOF were included for analysis. If a cohort was housed in one pen for the entirety of the 45 DOF period then the dimensions of the one pen were used for analysis. If a cohort was housed in 2 pens during the 45 DOF period then the dimensions of the second pen were used for analysis, but only when the cohort was limited to <7 DOF in the first pen. Any cohorts that were moved between 3 or more pens during the first 45 DOF were excluded from analysis.

The dataset was refined to remove potentially sparse data and enhance external validity. Cohorts that contained at least 25 animals at arrival were included in the analysis. The mean arrival weight was confined to cohorts with an average weight 227-453 kg as this weight range contained sufficient data for analysis. To avoid violating the assumption of

, total cohort size at arrival (25-99, 100-175, >175) and average arrival weight (227 – 272 kg, 273 – 318 kg, 319 – 363 kg, 364 – 408 kg, 409 – 453 kg) were categorized similarly to previous literature (Babcock et al. 2009). Heifers, steers, and mixed sex cohorts were included in the dataset. Arrival dates were included and were categorized into quartiles based on the arrival month to determine which quarter of the year the cohort entered the feedlot for cohorts that arrived in January through March (1), April through June (2), July through September (3), and October through December (4). Cohorts with missing or incomplete data for any of these variables and criteria were excluded from the study population.

Collected data were aligned with inclusion criteria, validated, categorized, and limited to only those with BRD-specific treatments. Inclusion criteria included cohort size restricted to ≥25 animals; only include cattle between 226-453kg; include heifer, steer, and mixed cohorts. Pen dimensions were added and tied to each cohort. Dimensions of each pen were measured utilizing Google Earth Pro (Google Earth Pro version 7.3.3.7786). Pen dimension characteristics were total pen area (square meters) and bunk space available (meters). Pen area was calculated by multiplying the length of the pen by width of the pen if the shape of the pen was a square or rectangular shape. If the pen had an irregular polygonal structure, then the 'polygon tool' was utilized to estimate the area of the pen in square meters. Bunk space was measured linearly by measuring the length of the bunk in each pen. These new measurements were added as variables to the dataset and tied to their corresponding pen number so each cohort had the pen dimensions during the first 45 DOF. Pen area available per animal was calculated by dividing pen area (square meters) by the number of cattle housed in each pen. Pen bunk space available per animal was calculated by dividing pen bunk space available by number of cattle housed in each pen.

The primary covariates of interest (pen space per animal and bunk space per animal) were categorized based on expected non-linear relationships between these variables and the outcome of interest (BRD incidence 45 DOF). Recommendations from Kansas State Research and

Extension were used to create three categories for both variables of interest. (Boyer et al., 2017). The first category represented values that were below recommendations, the second category represented values that met recommendations, and the third category represented values that were above recommendations. The categories for area per animal were ≤23.22m2, 23.23-32.52 m2, and >32.52 m2. The categories for bunk space per animal were ≤0.3 m, 0.31-0.46 m, and >0.46 m.

The data originated from a retrospective dataset provided from several commercial feedlot operations. As a result, vaccination programs from each operation were unavailable for this analysis. In addition, variables such as the distance cattle traveled to the feedlot, preconditioning status, and dietary profile were also unavailable for inclusion in the analysis.

**Statistical Methods**

A generalized linear mixed-model was fitted with the 'lme4' package (Bates et al. 2015) in R Studio version 3.6.2. (R Core Team, 2020) to assess potential associations of pen housing factors with BRD within the first 45 DOF. The outcome variable of interest was BRD incidence in the first 45 DOF and was calculated as the total number of first BRD treatments in the first 45 DOF (events) / total animal in the pen (trials). Covariates included average arrival weight, cohort size at arrival, arrival date quarter, sex, area per animal, and bunk space per animal (Table 2.2.). Several interaction terms were incorporated based on previous research which included: sex with average arrival weight; average arrival weight with arrival date quarter; area per animal with sex; arrival weight; cohort size at arrival; and arrival date quarter; bunk space per animal with sex; average arrival weight; total animal received; and arrival date quarter. (Cernicchiaro et al., 2012; Babcock et al., 2013a; Avra et al., 2017). A random intercept for feedlot was included to account for data hierarchical structure. Variables that have been previously determined to be associated with BRD (quarter of arrival, arrival weight, sex, cohort size at arrival) were retained in the model regardless of statistical significance. Remaining variables (including interactions) were retained only if they were significantly associated ($P < 0.05$) with the outcome or were part of a significant interaction term. All main effects were included regardless of significance if they were part of a significant ($P < 0.05$) interaction.

## Results

The final dataset used for analysis consisted of 1,733 cohorts representing a study population of 188,118 individual animals over 10 feedlots in 2 states. There were a total of 11,028 (5.9% of study population) cases of BRD within the 45 day evaluation period, with a mean of 3 and a range of 0 to 112 cases per cohort. Figure 2.1. shows the distribution of the total number of BRD treatments per cohort. Table 2.2. displays the distribution of cohorts in each area per animal and bunk space per animal category.

Variables significantly associated ($P < 0.05$) with BRD incidence in the first 45 DOF included sex, cohort size at arrival, average arrival weight, arrival date quarter, area per animal, bunk space per animal (Table 2.3.). All interactions between pen housing characteristics (area per animal, bunk space per animal) and cattle attributes (sex, cohort size at arrival, average arrival weight, arrival date quarter) were significantly associated ($P < 0.05$) with BRD incidence in the first 45 DOF. Estimated probabilities and standard error for BRD incidence in first 45 DOF were calculated for all interactions. All interactions included in the model were significant ($P < 0.05$) towards BRD incidence in the first 45 DOF.

The average arrival weight of the cohort modified the effect of area per animal on BRD incidence (Figure 2.2.). For example, a significant contrast was seen as cohorts with an average arrival weight of 409-453 kg had a higher probability of BRD incidence (3.6% ± 0.64) if they were placed in pens that had less than or equal to 23.22 m2 of area per animal available compared to similar weight cattle placed in pens that had 23.23-32.52 m2 (2.05% ± 0.39) or >32.52 m2 (2.47% ± 0.47) of area per animal available.

Total cohort size at arrival also modified the effect of area per animal on BRD incidence (Figure 2.3.). Cohort sizes with 25-99 or 100-175 showed no differences in the effect of area per animal on BRD incidence. However, the probability of BRD incidence was higher in cohort sizes >175 animal when given less than or equal to 23.22 m2 (4.46% ± 0.83) of pen area per animal compared to similar cohort sizes with 23.23-32.52 m2 pen area per animal (2.83% ± 0.54) or >32.52 m2 (2.40% ± 0.46) pen area per animal available.

There was an association between area per animal and BRD incidence in the first 45 DOF; however, this effect was modified by both the sex and arrival date quarter of the cohort. Although these two covariates (sex and arrival date quarter) were statistically significant there was no apparent pattern for their effect on BRD incidence.

Sex, cohort size at arrival, average arrival weight, and arrival date quarter were significantly associated ($P < 0.05$) with the effect of bunk space per animal on BRD incidence. Average arrival weight was the only covariate that modified the effect of bunk space per animal on BRD incidence (Figure 2.4.). A difference was seen between cohorts with an average arrival weight between 409-453 kg as cohorts in this weight category had a lower probability of BRD incidence if they were provided less than or equal to 0.3 m of bunk space per animal compared to similar weight cattle placed in pens that had 0.31-0.46 m or greater than 0.46 m of bunk space per animal available.

## Discussion

This study was conducted to estimate the relationship between feedlot pen-level housing conditions (pen area per head and bunk space per head) and cohort-level probability for BRD incidence within the first 45 DOF. Analysis of this relationship is important to determine the potential associations between pen-level housing conditions and whether changes in these conditions could be used to mitigate BRD cases in feedlot cattle. Previous research investigating this relationship is limited (Sanderson et al., 2008; Hay et al., 2016;). Both studies did not find identify associations between variables involving to housing conditions related to pen area and bunk space; however, the interactions between housing variables, and other cattle demographic risk factors commonly associated with BRD, were not assessed. Our categories for area per animal (≤23.22m2, 23.23-32.52 m2, >32.52 m2) and bunk space per animal (≤0.3 m, 0.31-0.46 m, >0.46 m) were not utilized in other studies that investigated pen housing conditions as risk factors. We choose these cutoffs for area/bunk space per animal to represent a category for below recommendations (≤23.22m2; ≤0.3 m), meets recommendations (23.23-32.52 m2; 0.31-0.46 m), and exceeds recommendations (>32.52 m2; >0.46 m) according to feedlot guidelines (Boyer et al., 2017). This was implemented in an attempt to improve external validity and evaluate potential differences in BRD incidence in the first 45 DOF between cohorts placed in pens that were below, met, or were above the recommendations for pen area per head and bunk space per head according to published guidelines. Our data encompassed several years from multiple commercial feedlots, and data structure allowed quantification of both effects of pen housing characteristics and interactions that have not been previously described at the cohort-level. Results of this study suggest pen housing factors related to pen area per animal and bunk space per animal are associated with BRD incidence in the first 45 DOF of the feeding period, but this impact is modified by cattle demographics. While several interactions were statistically significant, limited biological significance (or meaningful differences in BRD incidence) were present in some interactions. Our findings are unique and provide novel explanations on how pen housing conditions, combined with cattle demographic factors, may potentially influence animal health.

Previous research reported an association between the number of animals in a cohort and risk of BRD. (Cernicchiaro et al., 2012a; Cernicchiaro et al., 2012b; Hay et al., 2014;). Results from these studies vary on whether smaller or larger sized cohorts were associated with an increase in BRD risk. Our results demonstrated that the effect of cohort size on BRD incidence in the first 45 DOF was influenced by the amount of area per animal (sq. ft.) provided in each pen. Larger cohort sizes (>175) were associated with higher BRD incidence in the lowest area per animal category (<23.22 m2) when compared to smaller cohort sizes given a similar amount of pen area per head. There are several reasons as to why larger cohort sizes displayed a higher risk for BRD when given a lower amount of pen area per head. The impact of cohort size on BRD risk may be influenced to additional characteristics that are related to feedlot management and infrastructure. For example, increased commingling may have influenced the effect of cohort size as mixing cattle from different sources into a large-sized cohort may increase stress and leave cattle more prone to infection when having access to less pen area. (Step et al., 2008; Wiegand et al., 2020). Increased commingling in larger cohort sizes can also increase the transmission of communicable BRD pathogens (Mannheimia haemolytica, Bovine Viral Diarrhea Virus, etc.) that can potentially suppress the host immune system and impact the risk of BRD (Griffin et al., 2010). The retrospective data we collected was missing or did not have sufficient data to measure these potential factors that could be related to management and infrastructure and as a result our data cannot properly evaluate additional factors and their effects on cohort size beyond what is included in our model.

The interaction between average arrival weight and pen area per head were significant towards BRD incidence in the first 45 DOF. Across all categories of pen area per head and bunk space per head, categories of light weight cattle had an increased probability of BRD incidence in the first 45 DOF when compared to categories of heavier weighted cattle. The probability of BRD incidence was the highest in the lowest arrival weight categories (227kg-272kg and 273kg-318kg) and lowest in the highest arrival weight categories (364kg-408kg and 409kg-453kg). These findings are consistent with previous research that determined light weight cattle are more susceptible to BRD compared to heavier cattle (Reinhardt et al., 2009; Taylor et al., 2010; Babcock et al., 2013a). However, our study results demonstrated that cattle in the heaviest weight category (409-453 kg) had an increased BRD incidence when given less area per animal (23.22 m2) compared to similar weighted cattle given more area per animal. This was the only average arrival weight category that displayed a difference in BRD risk across the pen area per head categories. Heavy weight cattle arriving at 409kg-453kg are likely to already be started on feed and are likely to be less susceptible to BRD incidence Heavy weight cattle given fewer square meters per animal may be at greater risk of BRD incidence as less than 23.22 square meters per animal may have different effects on cohorts with light weight cattle compared to heavy weight cattle. Cohorts of heavy weight cattle will have less pen area per animal to utilize compared to a lighter weight animal given the same pen area with equal cohort sizes. Age may also be a proxy for BRD morbidity and can be related to the average arrival weight for cattle arriving to a feedlot (Taylor et al., 2010). Differences in our data towards BRD risk were not observed unless cattle had an average arrival weight at 409-453kg. Differences in this heavy weight category could be attributable to cattle age and these animals have likely been started on feed, received vaccinations, and overall be atypical of an animal that would be at an elevated risk of BRD morbidity (light weight animals).

Previously feeding behavior related to feeding timing has been investigated as a potential factor that may be associated with BRD morbidity. (Kayser et al., 2018). We hypothesized that in addition to feeding behavior, the amount of bunk space allocated to cattle may contribute to BRD risk. All interactions between bunk space per animal and the risk factors included in the model were significantly associated with the probability of BRD incidence. However, average arrival weight was the only risk factor that modified the effect of bunk space per animal on BRD incidence in the first 45 DOF. Our data indicated heavier weighted cohorts (409-453 kg) were associated with lower BRD incidence) in the smallest bunk space per animal category (≤0.3 m) when compared to cohorts in larger bunk space per animal categories (.31m-.46m, >.46m). The finding that heavier cattle with a smaller bunk space had lower BRD incidence could be explained by the type of cattle placed in this pen configuration. For example, these heavier cattle are likely older and could have already started on feed prior to the feedlot and thus are already past the high risk BRD period. Results of another study indicate that BRD risk was not different between different bunk space per animal categories (Hay et al., 2016); However, our bunk space per animal categories were categorized according to feedlot guidelines (≤0.3 m, 0.31-0.46 m, >0.46 m) while their categories for bunk space per animal were less than 0.18m and greater than or equal to 0.18m. We used these cutoffs as they were United States feedlot guidelines as opposed to a cutoff intended for Australian feedlot cattle. We also used three categories for our cutoffs whereas the previous study used two cutoffs.

A potential limitation of our study is that it is a retrospective analysis looking at pre-existing observational data. Retrospective studies may be subject to confounding and cannot determine causation, only associations. The results are also confined to the feedlots that were included in the dataset utilized and may not be applicable to feedlots outside of our dataset due to differences in management, dates of data recorded, geography, cattle types, different case definitions for BRD, and many other potential differences. Our data is was from commercial feedlots which is inherently "messy" and may contain unknown biases or errors. In addition, the data came from a retrospective analysis from commercial feedlot data that consisted of information collected from multiple feedlot operations. As a result, vaccination programs from each operation were unavailable for this analysis. In addition, the distance cattle traveled to the feedlot, and preconditioning status were also unavailable. The time frame from which we evaluated BRD treatments was only during the first 45 DOF, so we only incorporated 1st pull BRD treatments during the first 45 DOF and not all total BRD treatments throughout the entire feeding phase. It is possible that some of the risk factors we explored related to BRD incidence are significantly associated with risk of retreatment, risk of becoming a chronic animal, or risk of dying; our analysis did not include those outcomes due to limitations of the data provided in our dataset. Additional studies will be needed to evaluate those other important outcomes and their relationship with management-related risk factors. Several cohorts were also removed from the dataset if they were housed in more than 2 pens throughout the first 45 DOF as it was difficult to track which pen they were in when they developed BRD. There are many possible reasons related to cattle flow and management decisions that may have caused cohorts to move several times throughout the first 45 DOF. The risk status of cohorts entering the feedlot was not known in our dataset, so we could not determine which cohorts of cattle were classified as high risk for

BRD at arrival. A future, well-controlled prospective study examining the risk of BRD during the first 45 DOF in association with pen housing conditions should be conducted to help determine the differences in BRD incidence.

## Conclusion

Our results provide further initial estimates of how pen housing characteristics related to pen area per animal and bunk space per animal affect BRD incidence in the first 45 DOF. Our retrospective study determined that the probability of BRD incidence in the first 45 DOF is associated with pen housing conditions related to area per animal, bunk space per animal, and their interactions with cattle demographics. Pen housing conditions and their estimated impact on BRD incidence have not been thoroughly evaluated, and our results suggests that the associations of those pen housing condition variables with BRD incidence in the first 45 DOF is modified by other well-known risk factors. Further research in this area will lead to a better understanding of the impacts of housing conditions for feedlot cattle and how these conditions can potentially be modified to reduce the risk of BRD in the feedlot industry.

# Figures

**Figure 2.1. Histogram of the distribution of the level of morbidity of BRD incidence in the first 45 days on feed (DOF) in the study population. The x-axis displays the percentage of BRD incidence in a cohort during the first 45 DOF. The y-axis displays the count of cohorts for each value.**

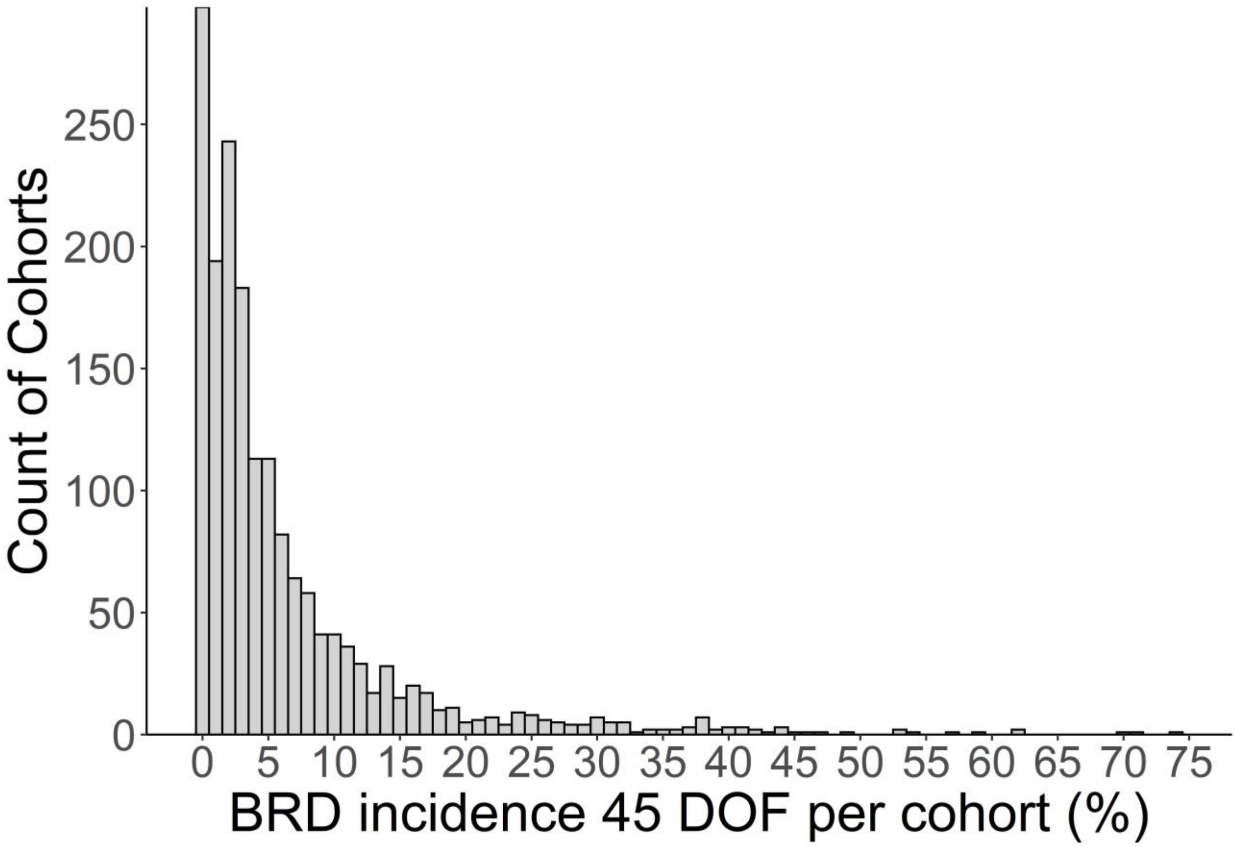


**Figure 2.2. Model estimated mean probability of BRD incidence by area per animal and arrival weight category in commercial feedlot cattle during the first 45 DOF. Error bars represent SE of least square means.**

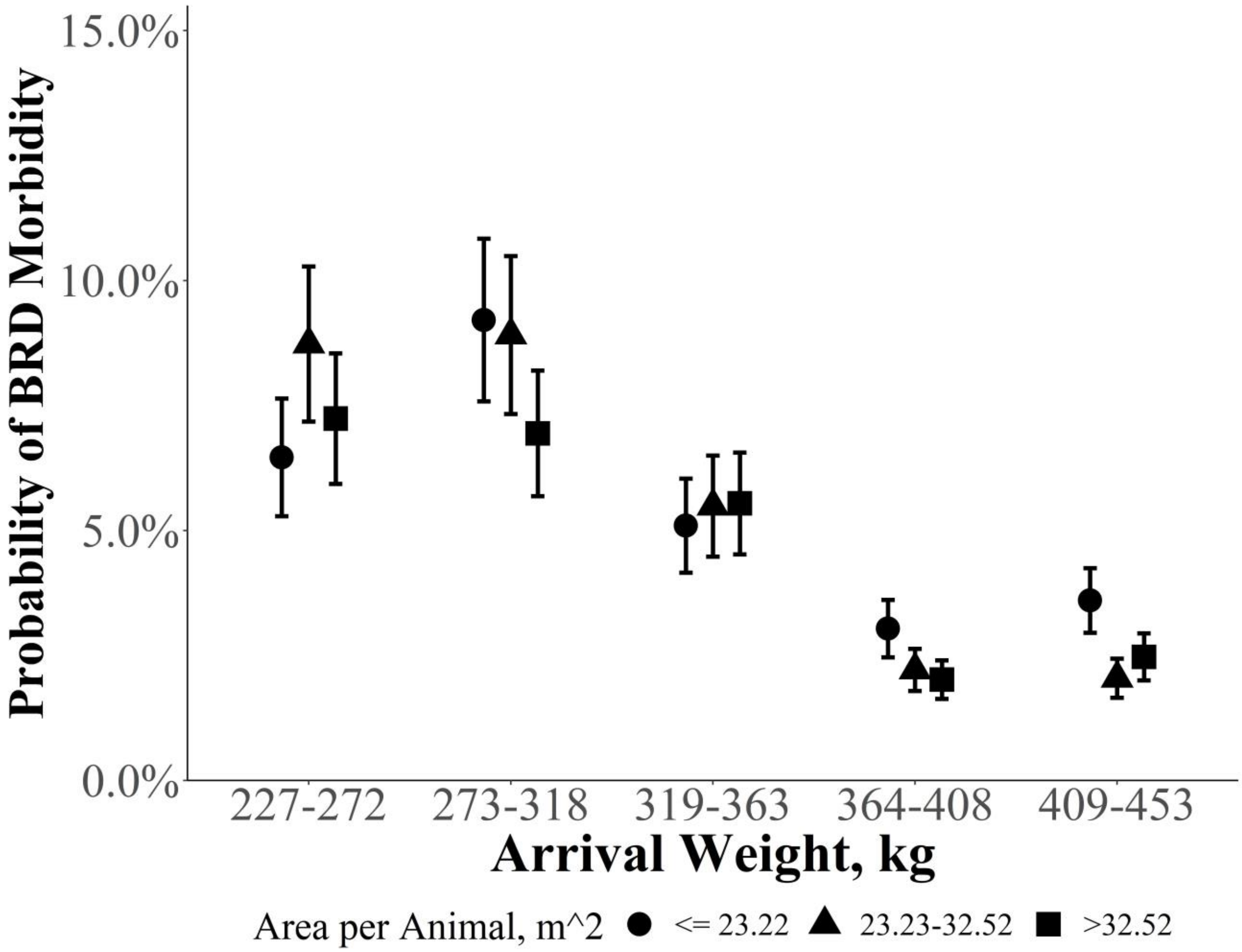


**Figure 2.3. Model estimated mean probability of BRD incidence by area per animal and size of the cohort at arrival in commercial feedlot cattle during the first 45 DOF. Error bars represent SE of least square means.**

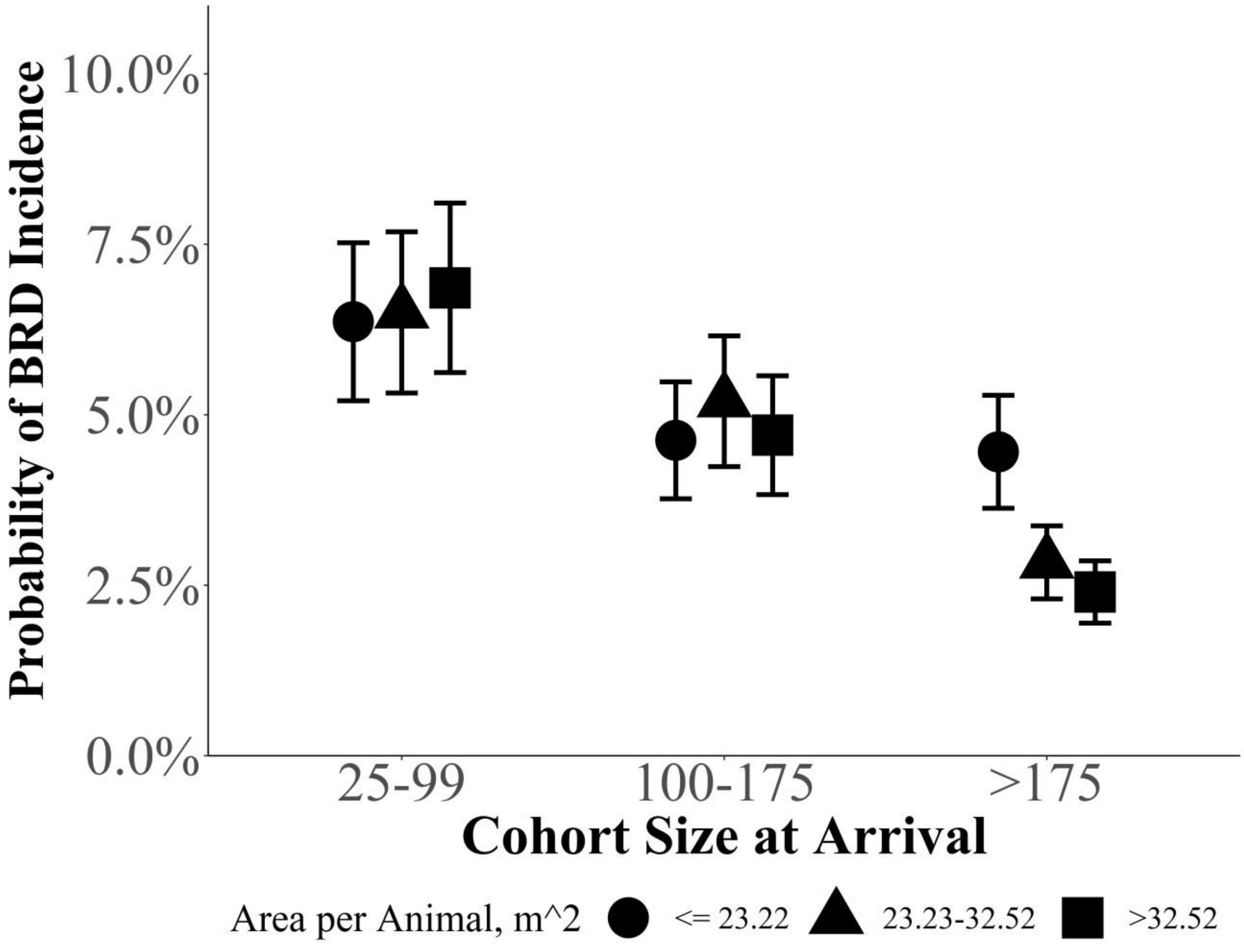
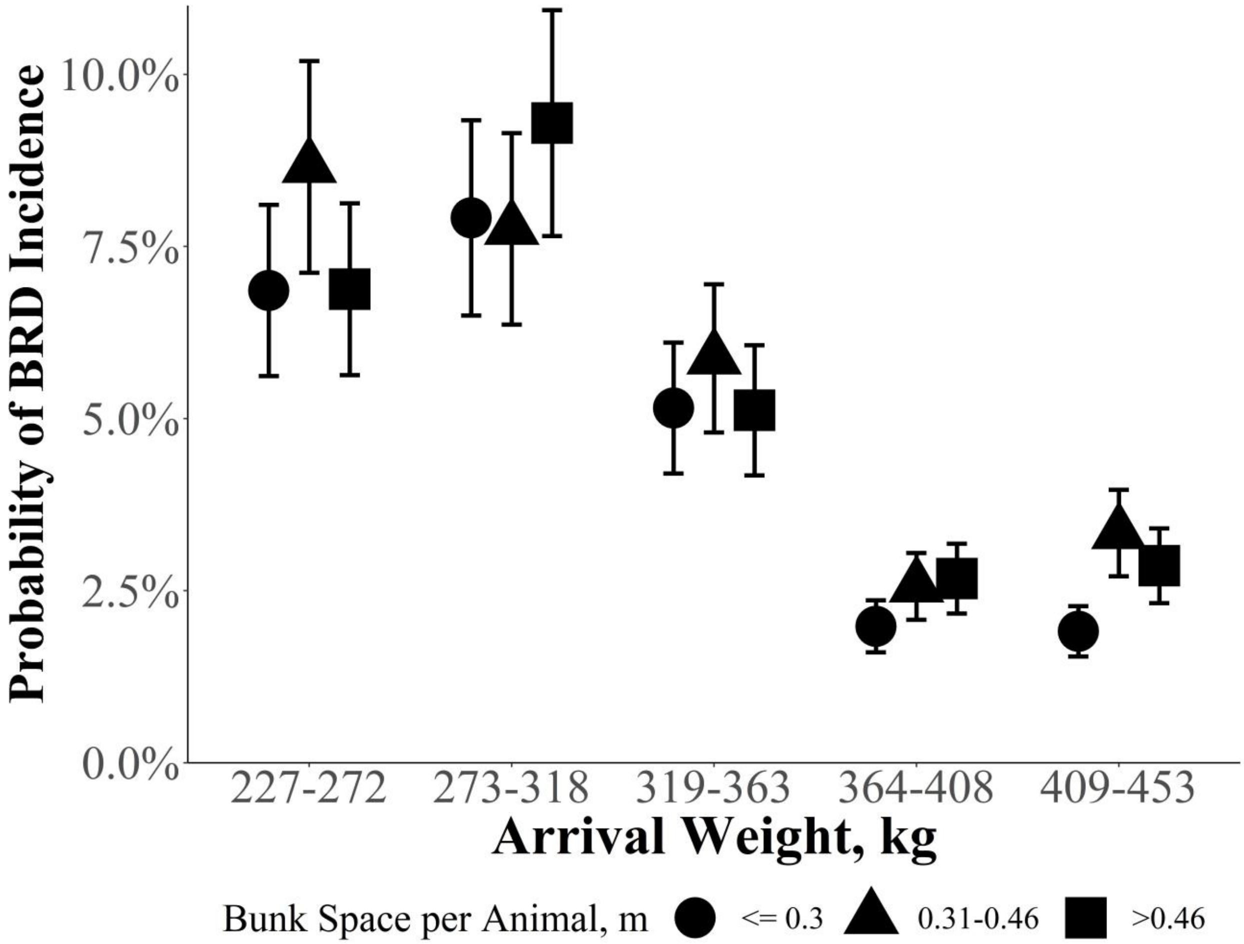


**Figure 2.4. Model estimated mean probability of BRD incidence by bunk space per animal and arrival weight category in commercial feedlot cattle during the first 45 DOF. Error bars represent SE of least square means.**



# Tables

**Table 2.1. Descriptive statistics for the study population cohorts (n = 1,733) from 10 feedlots**

| Variable | Mean | SD[2] | Median | Range | IQR[3] |
|---|---|---|---|---|---|
| Cohort Size at arrival | 108.69 | 55.88 | 87 | 25 to 324 | 64 to 144 |
| Average arrival weight, kg | 344.96 | 101.68 | 346.09 | 228.16 – 453.59 | 314.79 – 377.39 |
| Area per animal, $m^2$ | 28.64 | 16.88 | 24.68 | 4.77- 249.49 | 20.32 – 31.02 |
| Bunk space per animal, m | 0.42 | 0.24 | 0.34 | 0.13 – 2.06 | 0.29 – 0.43 |
| BRD incidence[1], % | 6.44 | 9.00 | 3.26 | 0 to 74.07 | 1.28 to 7.84 |

[1] First treatment bovine respiratory disease (BRD) incidence was our outcome variable and was calculated only for the initial 45 days on feed

[2] SD = standard deviation

[3] IQR= Interquartile Range

**Table 2.2. Distribution of variables used for analysis from 10 feedlots from 2015-2020**

| Variable & Category | Number (%) of cohorts |
|---|---|
| Cohort size at arrival | |
| 25-99 | 946 (54.58) |
| 100-175 | 537 (30.99) |
| >175 | 250 (14.43) |
| Average Arrival Weight. kg | |
| 227 – 272 | 131 (7.56) |
| 273 – 318 | 351 (20.25) |
| 319 – 363 | 650 (37.51) |
| 364 – 408 | 455 (26.26) |
| 409 – 453 | 146 (8.42) |
| Sex | |
| Heifers | 922 (53.20) |
| Steers | 678 (39.12) |
| Mixed | 133 (7.68) |
| Arrival Date Quarter | |
| Jan-March (1) | 473 (27.29) |
| April-June (2) | 471 (27.18) |
| July-September (3) | 481 (27.76) |
| October-December (4) | 308 (17.77) |
| Area available per animal, $m^2$ | |
| ≤23.22 | 733 (42.30) |
| 23.23-32.52 | 540 (31.16) |
| >32.52 | 460 (26.54) |
| Bunk space available per animal, m | |
| ≤0.3 | 539 (31.10) |
| 0.31-0.46 | 808 (46.62) |
| >0.46 | 386 (22.28) |

**Table 2.3. Final generalized linear mixed-model demonstrating housing characteristics and cattle demographic factors and their association with bovine respiratory disease incidence during the first 45 DOF**

| Variable | P-value |
|---|---|
| Sex | <0.01 |
| Cohort size at arrival | <0.01 |
| Average arrival weight | <0.01 |
| Arrival date quarter | <0.01 |
| Area per animal | <0.01 |
| Bunk space per animal | <0.01 |
| Sex x Average arrival weight | <0.01 |
| Average arrival weight x Arrival date quarter | <0.01 |
| Sex x Area per animal | <0.01 |
| Average arrival weight x Area per animal | <0.01 |
| Cohort size at arrival x Area per animal | <0.01 |
| Arrival date quarter x Area per animal | <0.01 |
| Sex x Bunk space per animal | <0.01 |
| Average arrival weight x Bunk space per animal | <0.01 |
| Cohort size at arrival x Bunk space per animal | <0.01 |
| Arrival Date Quarter x Bunk space per animal | <0.01 |

# Chapter 3 - Predicting bovine respiratory disease risk in feedlot cattle in the first 45 days post-arrival

## Abstract

Bovine respiratory disease (BRD) is the leading cause of morbidity and mortality in feedlot cattle. At arrival feedlot cattle groups are assigned into high- or low- risk groups based on their expected probability of cumulative BRD risk based on several criteria. Improved predictions of cohort-level BRD risk at arrival to a feedlot setting facilitates appropriate and judicious intervention application. Our goal was to build and evaluate the diagnostic performance of five different classification models to predict the expected risk of BRD morbidity within the first 45 days on feed (DOF). In addition, an economic analysis was performed to determine if there was a potential health cost advantage when using a predictive model compared to standard methods of predicting expected BRD risk.

Data from 10 U.S. feedlots containing 1,733 cohorts representing 188,188 cattle with known health outcomes were classified by predictive models into high (≥ 15% BRD morbidity) or low (<15%) BRD risk in the first 45 DOF. Classification models including logistic regression, decision tree, random forest, discriminant linear, and naïve Bayes models were trained to identify expected BRD risk. Each models' performance was evaluated compared to actual health outcomes using receiver operating characteristic (ROC) curves to determine the area under the curve (AUC). In addition, a net health cost benefit (NHCB) was calculated for each model using estimated costs for diagnostic outcomes (TP, FN, FP, TN) to determine if there was a health cost advantage at different proportions of high-risk cohorts for each model and a control scenario that represented a human classifying expected risk without a model.

Area under the curve was calculated using each model-generated ROC curve from the test dataset and ranged from .682 to .789. The economic performance for each model was dependent on the true proportion of high-risk cohorts in the population. The decision tree model had a greater potential economic advantage compared to the control scenario when the proportion of high-risk cohorts was ≤ 45%. Results illustrate that predictive models can be useful at delineating cattle at high or low risk for disease and may provide more economic value than standard methods. The amount of value provided by each model varied by the prevalence of high risk cohorts in the population.

## Introduction

Bovine respiratory disease (BRD) remains the costliest disease in the American feedlot industry costing between approximately $800 and $900 million annually (Chirase and Greene 2001). Cattle arriving to a feedlot are managed in groups and management decisions such as the administration of antimicrobial metaphylaxis upon arrival is determined on the perceived risk of a high percentage of cattle within a group developing BRD (Ives and Richeson, 2015). Feedlot decision-makers use BRD risk classification of high-risk versus low-risk to decide whether cohorts of cattle will or will not receive metaphylactic treatment upon arrival to the feedlot. Many factors can drive the perceived BRD risk of incoming cattle groups which can vary by personnel and organizational policies. As a result, misclassifications may occur when determining BRD risk classes. These misclassifications may negatively impact cattle health and ultimately lead to increased expenditures towards treating health-related events.

Previous work has encouraged the use of operational feedlot data to predict BRD health outcomes. (Babcock et al., 2013; Amrine et al., 2014). The ability to use feedlot data available at the time cattle arrive to correctly predict and classify incoming cohorts of cattle into high- or low- risk groups would allow for more judicial use of antimicrobials, potentially increase economic performance, and allow personnel to focus their efforts on those cohorts expected to classify as high-risk. Previously investigated cohort risk factors such as average body weight, sex, quarter of arrival, and cohort size have all been found to be associated with BRD morbidity risk. (Taylor et al., 2010; Cernicchiaro et al., 2012; Hay et al., 2014). Pen housing conditions such as pen area and bunk space per head have also been explored as risk factors for BRD. (Hay et al., 2016). The objective for this study was to assess the ability of five different classification algorithms to accurately predict an incoming group of cattle's risk classification (high/low) using commercial feedlot data during the first 45 days on feed (DOF). In addition, the economic performance of each model was evaluated to determine if a potential health-cost advantage was present with the use of each models' final model outputs.

## Materials and methods

### Data

Retrospective data from 10 Midwest feedlots were collected between January 2018 and April 2020 and utilized for this study. A cohort was defined as a group of cattle purchased and managed in a similar manner and housed together throughout the study period during the initial 45 days on feed post-arrival. Cohort- and individual-level variables were included in the dataset. Cohort-level variables included: average arrival weight (total weight of all animals within the cohort divided by the total head in that cohort), number of cattle in cohort at arrival, arrival date quarter, and sex (steers, heifers, mixed gender). Individual-level data included the total number of individual first treatments in each cohort for BRD within the first 45 DOF. Bovine respiratory disease incidence, our outcome, was defined as the number of cattle that were treated at least once for BRD based on feedlot diagnosis within the first 45 DOF divided by the size of the cohort. The case definition for a BRD treatment was any animal that received an antimicrobial treatment for BRD during the first 45 DOF. Cases were limited to first BRD treatments only, and any additional treatments were excluded from analysis. If an animal was treated more than once, the first treatment record was utilized. Cohorts with missing data for any of these variables were excluded from the study population.

Pen housing variables were calculated for each cohort including: pen area (m2), bunk space available (m), pen area per head (m2) and bunk space per head (m). Dimensions of each pen were measured utilizing the 'ruler tool' Google Earth Pro (Google Earth Pro version 7.3.3.7786). Pen area was calculated by measuring the square meters of each pen. This was done by multiplying the length of the pen by width of the pen if the pen shape was square or rectangular. If the pen had an irregular polygonal shape then the 'polygon tool' was utilized to measure the area of the geometric shape of the pen. Linear bunk space was recorded by measuring the length (m) of visible bunk in each pen. Pen area per head was calculated by dividing pen area (m2) by the cohort size at arrival for each cohort. Bunk space per head was calculated by dividing pen bunk space available (m) by cohort size at arrival for each individual cohort. Cohorts without available pen housing measurements were removed from the dataset. Cohorts that were housed in 2 or fewer pens within the first 45 DOF were included for analysis. If a cohort was housed in one pen for the entirety of the 45 DOF period then the dimensions of the one pen were used for analysis. If a cohort was housed in 2 pens during the 45 DOF period then the dimensions of the second pen were used for analysis, but only when the cohort was limited to <7 DOF in the first pen. Any cohorts that were moved between 3 or more pens during the first 45 DOF were excluded from analysis.

## Data Preparation

The cumulative percent of cattle receiving a first treatment for BRD within the first 45 DOF was calculated for each cohort. The primary study outcome was expected cohort-level BRD risk classification (high- or low-risk) based on a treatment cutoff of 15% total BRD morbidity within the first 45 days on feed that has previously been used in prior research (Theurer et al., 2015). If 15% or more animals in a cohort were treated for BRD at least once in the first 45 DOF the cohort was classified as a high-risk cohort. If less than 15% were treated for BRD during the first 45 DOF then the cohort was classified as a low-risk cohort. A new binary cohort-level variable was created to represent the cutoff and populated with a value of 1 if BRD morbidity was greater than or equal to 15%, or 0 if BRD morbidity was less than 15%.

## Data Partitioning

Models may become overfitted and provide inaccurate, biased estimates when utilizing a single dataset for training and testing the models. An overfitted model developed with a single dataset may fail to predict new data sets accurately (Cawley and Talbot, 2010). Multiple datasets are used to avoid biased estimates and improve each models' discrimination ability by evaluating final diagnostic performance in a dataset independent of data used for model building phase. Data were partitioned 75% into a training dataset (n = 1,300) and 25% (n = 433) into a testing dataset using the 'tidymodels' (Kuhn et al., 2020*)* R package. The data splitting process was stratified to ensure that the training and test dataset produced the same frequency distribution of high- and low-risk cohorts in each dataset. Each of the five individual models were created using the training dataset and the final metrics for each models' performance were obtained using the testing dataset only once. A flow diagram of data preparation, partitioning, and classification is shown in Fig. 3.1.

## Model Recipe Creation

The ‘tidymodels’ (Kuhn and Wickham 2020) R framework was utilized to create a recipe that defines a series of data preprocessing tasks and develops a model specification formula. Within the recipe, BRD morbidity risk (high/low) in the first 45 days on feed was selected as the outcome variable and predictor variables of interest were identified. (Table 3.1.). Variables that were not meaningful in external application of the model such as pen ID, cohort ID, and feedlot ID were excluded from analysis. An indicator (or dummy) variable was created for each qualitative variable and converted them into a matrix of dummy variables that are 0 or 1 for that categorical variable. This formula and training dataset were used across the five models tested in our analysis.

## Classification Algorithms

Five commonly used predictive models were used to predict the BRD risk class of each cohort of cattle. The models used were: logistic regression, decision tree, random forest, naïve Bayes, and linear discriminant. Each individual predictive model was trained with the training dataset. Evaluation of the model performance was performed using the test dataset with the pre-defined cutoff (15%) for BRD morbidity risk within the first 45 days after arrival as the outcome of interest.

### Logistic Regression

Logistic regression is a statistical model used when the outcome variable is binary. It describes the linear relationship between the outcome and explanatory variables using the logistic function to observe the effect of each variable on the probability of the observed event of interest (Petrie and Watson 2013). The predicted class selected is based on which class has the highest probability. The ‘glmnet’ function in R was used to create the logistic regression models (Friedman et al., 2010).

### Decision Tree

Decision tree is a hierarchical classification machine learning model composed of decision rules that recursively classify data from the training dataset through a series of questions (Myles et al., 2004). Each node in the tree contains a question regarding the predictor variables and question nodes are added incrementally to increase separation of the training data into their categories as effectively as possible (Kingsford and Salzberg, 2008). Decision tree models were built using the ‘rpart’ R package (Therneau and Atkinson 2019).

### Random Forest

Random forest is a classification machine learning algorithm that generates many classification models and aggregates their results (Liaw and Wiener., 2002). Random forest models operate as an ensemble that consists of many individual decision tress that arrive at a class prediction. The model’s prediction is determined by the most abundant class. The ‘ranger’ package was used to create the random forest models (Wright and Ziegler 2017).

### Naïve Bayes

Naïve Bayes is a classification algorithm that uses Bayes’ theorem of probability and assumes independence among predictors in a given class (Webb, 2016). Naïve Bayes models provide a mechanism that uses the training data to estimate the posterior probability of each class given a specific variable. The class with the highest posterior probability is the outcome of the prediction. Naïve Bayes models were built using the “naivebayes” package (Majka 2019).

### Linear Discriminant

Linear discriminant is a classification algorithm that determines a hyperplane to maximize the separation of the projected means of classification groups (Koehler and Erenguc, 1990). Groups are specified by the discriminant process and data points are classified by where they lie on the hyperplane. Linear Discriminant models were built using the 'mda' R package (Hastie and Tibshirani 2020).

## Resampling – cross-validation

A *k*-fold cross-validation resampling method was applied to the training dataset. The goal of using cross-validation was to generate different versions of the training dataset to estimate how well the models will perform with new data that was not used to train the models. This helps to avoid overfitting and selection bias within each model. (Cawley and Talbot, 2010). In this case, *k*-fold cross-validation splits the training dataset into *k* smaller subsets, or "folds," of the data. Each model is trained using $k - 1$ of the folds as training data and the model is validated on the remaining part of the data as a test set. The performance metrics reported by the *k*-fold cross-validation is the average of the values between all folds of the data (Kuhn and Johnson 2020).

The training data within all five classification models were evaluated with 10 distinct folds. For each iteration, data from the training dataset were randomly partitioned (75% - 1,300 cohorts out of the original 1,733) into 10 equally sized subsets (folds) of data. The remaining 25% (433 cohorts out of the original 1,733) were used as the test dataset. Stratified sampling was done to ensure that each fold had the same frequency distribution of the outcome.

## Model optimization/tuning

Model optimization/tuning is performed to find a combination of hyperparameters in a given machine learning algorithm that provides the best model performance. Hyperparameters have a direct impact with the model's learning process and act as model settings that can be adjusted to optimize the model's performance (Kuhn and Johnson, 2020). A grid search was performed to determine candidate tuning parameter values for each model. Some models have more than one tuning parameter, and in this case candidate parameters combinations values are created. The resampling data was used to evaluate each parameter value combination and obtain estimates of how well each candidate model performs. After evaluation, the hyperparameter values that produce the best results in the grid search were selected and used for analysis of the test dataset for final analysis of each model utilizing the cross-fold validation dataset.

## Model evaluation

Final evaluation of the models was performed by allowing each algorithm to classify predictions using the test dataset. Classifier predicted probabilities of BRD morbidity risk of low or high were created for each distinct cohort for each classification model. Receiver-operating characteristic (ROC) curves were created utilizing these probabilities compared to known actual health outcomes using the 'yardstick' package in R (Kuhn and Vaughan 2021). ROC curves show the diagnostic ability of binary classification models and the trade-off between sensitivity and specificity for every possible cutoff for a test (Gardner and Greiner, 2006). The cutoff point that was utilized from each generated ROC curve was the point where sensitivity and specificity were maximized by calculating Youden's index (Youden, 1950). Youden's index has a range from 0 and 1, with the value of 1 indicating the test has perfect sensitivity and specificity. Classification model performance was then evaluated using the final predicted classes based on the cut point selected. Our primary metric for initial model comparison is AUC because it is a measure of the degree of separability and how well the model can distinguish between classes using a range from 0 to 1 where a value of 0 indicates a perfectly inaccurate test and a value of 1 indicates a perfectly accurate test (Safari et al., 2016). Additional metrics calculated and evaluated were true positives (TP), true negatives (TN), false positive (FP), false negatives (FN), Positive Predictive Value (PPV), Negative Predictive Value (NPV), sensitivity (Se), specificity (Sp), and accuracy. Fig. 3.2 displays a flowchart describing how a model would arrive at each diagnostic outcome and the calculations for each metric.

### Economic analysis

An economic analysis was performed with the goal estimating a cohort-level Net Health Cost Benefit (NHCB) for each predictive model and a control scenario that represented a person classifying expected BRD risk without the use of a model. This NHCB was meant to represent the health costs associated with each model to predict expected BRD risk. These health costs include expenses associated with administration of BRD treatment and potential lost value from a morbid animal compared to a healthy animal. The values for these costs were determined based on previous reports and averages from the study population dataset (Table 3.2.). BRD morbidity was defined as the number of cattle within a cohort that were treated for respiratory disease at least once in the first 45 DOF. The cost of a morbid animal was considered as $151.18 per head based on data from a previous Texas A&M Ranch to Rail summary report (McNeil et al., 2001). This cost considered the return difference from healthy animals compared to sick animals in medicine costs and “lost value” due to reduced efficiency, lowered gain, and reduced sale value. The cost of a single metaphylaxis treatment ($23.60) was the average cost for a respiratory disease treatment reported in a USDA NAHMS report. (USDA 2013). The average cohort size (n = 109) was set as the average cohort size in our study population dataset and was calculated by taking the total number of animals in our study population and dividing it by the total number of cohorts in the population (n = 188,118/1,733). The proportion of high-risk cohorts served as a range that represented the proportion of high-risk cohorts that could have potentially been present in the data. For example, if this number was set as .25 then 25% of the cohorts in the dataset would be expected to be high-risk for BRD. The metaphylaxis efficacy was set at 0.5 (50%) to represent the reduced morbidity after metaphylaxis treatment (Avra et al., 2017).

In this analysis there was a decision on whether an incoming cohort is going to receive, or not receive, metaphylactic treatment and thus a cohort could be true positive (had ≥15% morbidity within the first 45 DOF and were predicted to be high risk) or a false positive (predicted to be high risk and had < 15% morbidity within the first 45 DOF). In our analysis we varied the prevalence of high-risk cohorts (0-100) to represent feedlots with different types of incoming cattle. The total number of diagnostic outcomes (TP, TN, FP, FN) were calculated at each level of prevalence for each predictive model. A cost incorporating the cost of metaphylaxis and lost value from a morbid animal was assigned to each diagnostic outcome (TP, TN, FP, FN) to calculate the NHCB. The NHCB was then subsequently divided by the average cohort size of animals in the study population (mean = 109) to produce an average cost/benefit per animal for use of each model. Table 3.3. describes the potential diagnostic outcomes generated from each model prediction with the associated metaphylaxis decision that would be decided by feedlot management and the anticipated financial result from the decision made. The NHCB was then subsequently divided by the average cohort size of animals in the study population (mean = 109) to produce an average cost per animal for use of each model. The formulas for the NHCB and cost of each diagnostic outcome are shown below:

$$NHCB = TPcost + TNcost + FPcost + FNcost$$

$$TNcost = \$0\ (baseline; no\ incurred\ costs)$$

$$TPcost = (MC * (BRD45 * ME) * TC) - (CS * TC)$$

$$FPcost = -(TC * CS)$$

$$FNcost = -((BRD45 * ME) * MC * CS)$$

TNcost, TPcost, FPcost, and FNcost represent the cost per animal of a true negative, true positive, false positive, and false negative outcome, respectively. MC represent the cost incurred from the lost value of a morbid animal compared to a healthy animal ($151.18). BRD45 represents the percent of cohorts in the population that are expected to be high-risk for BRD morbidity in the first 45 DOF and in our analysis could take a value from 0 to 100. ME represents the metaphylaxis efficacy, estimated to be a 50% reduction in morbidity. TC represents the average treatment cost for a single metaphylactic treatment ($23.60). CS represents the average size of a cohort (109).

A control scenario was included to represent to a human classifying expected BRD risk to incoming cohorts without the use of a predictive model. This was added in order to compare the cost of using a model against standard methods to predict expected BRD risk. A NHCB for a control scenario was also calculated to compare the economic output between the model results and the control scenario. To achieve this, Se and Sp were calculated from a subset of data (n = 177 cohorts) that included the actual risk status assigned to each cohort by feedlot management. This Se and Sp were calculated by comparing the feedlot's classifications to the actual health outcomes for each cohort based on the 15% BRD morbidity cutoff that was used in the modeling process. For example, if a feedlot classified an incoming cohort as high-risk, and the percent of the cohort that was treated for BRD once was 15% or greater, then it was called a TP. Diagnostic outcomes (TP, FP, TN, and FN) outcomes were calculated based on these criteria. The NHCB was formulated in the same manner as the costs for each model using the calculated sensitivity and specificity. A difference from the control ($/animal) was calculated at each proportion of high-risk cohorts to low-risk cohorts to compare the NHCB between the five models and the control. The control was set at $0/animal and all models were compared to the control. If a model displayed a value greater than $0/animal at any proportion of high-risk cohorts then that indicated that there was a potential economic advantage to use the model relative to the control. If a model displayed a value less than $0/animal at any proportion of high-risk cohorts then that indicated that there was a potential economic disadvantage to use the model relative to the control.

## Results

### Descriptive Statistics

The study population data included 1,733 distinct cohorts of cattle representing 188,118 individual animals. Each cohort was classified into a risk class category of high or low based on BRD morbidity risk in the first 45 DOF in relation to a 15% cutoff. In our training dataset there were 141 (10.85%) cohorts that were high-risk and 1,159 (89.15%) cohorts that were low-risk using the 15% cutoff. In the test dataset there were 47 cohorts that were high-risk (10.85%) and 386 cohorts that were low-risk (89.15%). The mean BRD morbidity in the first 45 days for all cohorts of cattle was 6.30%. The mean BRD morbidity in the first 45 DOF for high-risk cattle was 27.07% and mean morbidity for low-risk cattle was 3.77%. A total of 1,300 cohorts were partitioned into the training dataset while 433 cohorts were partitioned in the test dataset. The prevalence of high-risk cohorts in the training and test dataset was 10.85%, respectively.

### Model Performance Diagnostics

#### Area Under the Curve and Classification Accuracy

The accuracies and area under the curve (AUC) of the five classification models were evaluated using the test dataset (Table 3.4.). AUC was calculated using each model generated ROC curve from the test dataset. AUC of the models ranged from .682 to .789 for decision tree and random forest respectively. Accuracy of the models ranged from 10.9% to 79.4% for naïve Bayes and logistic regression, respectively.

### Sensitivity and Specificity

Sensitivity between the classification models ranged from 44.7% to 100% (Table 3.4.). The highest sensitivity, 100%, was achieved using the naïve Bayes model. The lowest sensitivity, 44.7%, was achieved using the decision tree model. The model with the highest specificity was the decision tree model at 83.7%. The model with the lowest specificity was the naïve Bayes model at 0%.

### Positive/Negative Predictive Value

The model positive predictive value and negative predictive value with a 10.85% prevalence of high-risk cohorts ranged from 10.9% to 25.0%, and 92.6% to 97.9%, respectively (Table 3.4.). The model with the highest positive predictive value was the decision tree model (25.0%); and the model with the lowest positive predictive value was the Naïve Bays (10.9%). The model with the highest negative predictive value was the random forest (97.9%); whereas the model with the lowest negative predictive value was the decision tree (92.6%).

### Economic Results

The derived sensitivity and specificity for the control scenario was 83.75% and 59.79%, respectively. The Net Health Cost Benefit and the difference from the control ($/animal) were calculated for each of the five models. The difference from the control for each model was variable depending on the proportion of high-risk cohorts to low-risk cohorts in the population. As a result, the potential economic advantage/disadvantage of using a model compared to not using a model (control) was volatile at different proportions of high-risk cohorts to low-risk cohorts (Fig. 3.3.). In our study, logistic regression and random forest models always offered a positive, but small, difference from the control (higher $/animal) than the control method across all possible prevalence of high-risk cohorts as they had a higher cost per head advantage at all proportions of high-risk cohorts to low-risk cohorts. Decision Tree models had a positive difference from the control when the proportion of high-risk cohorts to low-risk cohorts are below 45%. Naïve Bayes models had a positive difference from the control when the proportion of high-risk cohorts to low-risk cohorts was above approximately 83%. Linear Discriminant models had a positive difference from the control when the proportion of high-risk cohorts was above approximately 25%.

## Discussion

Protocols related to health management for cattle entering a feedlot are often based on the expected risk class of disease within the group. Accurately predicting the health outcome of incoming cohorts of cattle can serve to increase feedlot performance, efficiency, and economic performance. Previous studies have looked at incorporating cohort characteristics and risk factors at arrival into predictive algorithms to accurately classify cohorts into classes related to BRD morbidity. (Babcock et al., 2013; Amrine et al., 2018). However, these studies did not incorporate variables linked to pen housing conditions such as pen and bunk space that previously have been investigated. (Hay et al., 2016). In this study we evaluated the diagnostic ability of five predictive algorithms to predict BRD morbidity risk (high or low) for cattle arriving to a feedlot within the first 45 days post arrival using a 15% cutoff while incorporating previously associated risk factors for BRD. We evaluated the predictive ability towards the outcome of interest of high/low BRD morbidity risk in the first 45 DOF from the models that were produced from the use of the variables in our analysis. The amount of BRD morbidity risk within a cohort that is acceptable before management intervention differs among feedlot producers; therefore, an economic analysis was performed to better determine if the predictive performance of any of the five models would be economically beneficial compared to a person classifying incoming groups of cattle high- or low-risk.

The AUC of each model's receiver operating characteristic (ROC) curve was the metric used to rank the predictive performance of each model. In general, an AUC of 0.5 describes a model that has no discriminatory ability and serves as a model that has a 50% probability of correctly classifying an observation. An AUC between 0.7 to 0.8 is acceptable, 0.8 to 0.9 is excellent, and more than 0.9 is outstanding. (Mandrekar 2010). In our study, the model with the highest AUC is the random forest model at .79, with the range of AUC being .68 to .79 between all models. This indicates the models' performance ranged from poor to acceptable based on AUC. The overall accuracy of each model was also calculated, however evaluating accuracy alone may be misleading when interpreting the final results in an imbalanced dataset. (Menardi & Torelli, 2012). This was because the majority of the cohorts in the test data were classified in the low-risk category (<15% were treated for BRD during the first 45 DOF). If a predictive model classified every cohort as low-risk it would have an accuracy of 89.1%. On the surface this appears to be an acceptable accuracy, however, when a model doesn't have any discriminating ability and predicts all cohorts the same, the model is relatively useless. The naïve Bayes model appeared to be doing this, but instead of predicting all cohorts as low-risk, it is predicted all cohorts as high-risk resulting in an accuracy of 10.9%, even when the optimal threshold was selected for this model. As a result, AUC was used to rank the performance of each model as it avoids this bias and allows us to better understand the predictive ability of our models.

The prevalence of high-risk cohorts in our dataset was 10.85% (188 high-risk cohorts out of 1,733 total cohorts). The PPV represents the proportion of predicted high-risk cohorts that were truly high-risk. The negative predictive values report the proportion of predicted low-risk cohorts that are truly low-risk. The PPV from our models ranged from 10.9% (Naïve Bayes) to 25% (Decision Tree) demonstrating that our final models have a low probability of predicting positives (high-risk) that are actual positives. In contrast, our models' NPV ranged from 92.6% (Decision Tree) to 97.9% (Random Forest), with one algorithm generating a division by zero error (Naïve Bayes). The naïve Bayes model had this error because the model's specificity was 0%, so it predicted every cohort to be positive (TP or FP). As a result, 0 cohorts were predicted as negatives (TN or FN) and a NPV could not be calculated from the results of the naïve Bayes model. Four of the five models created from this data perform well at predicting negatives (low-risk) that are actual negatives. As prevalence decreases, positive predictive value will increase, and the negative predictive value will decrease, and vice versa. (Tenny and Hoffman 2021). When evaluating the PPV and NPV for each model the expected prevalence of the outcome of interest should be considered.

Determining the practicality of each model's usage in a feedlot setting requires both knowledge of cohort-level feedlot data characteristics and predictive modeling to bridge model outputs and how they apply to the business and economics of a feedlot operation. Feedlot producers generally use expected BRD risk to make a decision on whether to give an incoming group of cattle metaphylactic treatment to reduce cattle morbidity, increase production, and increase profitability. Costs are associated with correctly/incorrectly administering metaphylaxis to incoming cohorts of cattle. The ability to utilize a predictive model to predict the expected BRD risk of an incoming cohort may aid feedlot operations in correctly administering metaphylaxis treatment and consequently maximize profitability. To evaluate if there was a potential economic advantage with using a predictive model, an economic analysis was performed with the goal of calculating a Net Health Cost Benefit for each model to determine if a model would be economically advantageous in comparison to a person making a decision on the risk status of incoming cohorts.

Previous studies have utilized predictive analytic techniques to attempt to predict an outcome of interest, however these studies have not incorporated an economic component to their model analysis. (Babcock et al., 2013; Amrine et al., 2014). We included a deterministic economic approach to estimate a cost for use of each model when classifying cohorts into high- or low-BRD morbidity risk. As a deterministic approach was used to calculate costs, there is no randomness or variability in our results from potential factors that may impact the population. This was not meant to represent a full-scale economic analysis and only estimates the costs and potential benefits from associated treatments and use of metaphylaxis to mitigate BRD. In our economic analysis there was a decision to be made on whether an incoming cohort will receive metaphylaxis based on its expected risk for BRD. Cohorts called high-risk (positive) will receive metaphylaxis and cohorts identified as low-risk (negative) will not receive metaphylactic treatment. The cost of a true-negative, false-negative, true-positive, and false-positive at different proportions of high-risk cohorts were calculated to determine the costs associated with each outcome. These costs only take into consideration the associated health costs for a metaphylaxis treatment and the potential lost value of a morbid animal compared to a healthy animal. As a result, the calculations for these costs do not consider expenditures related to additional factors related to feed costs, management costs, and other potential costs that were not included.

The model that provided the greatest economic advantage was dependent on the prevalence of high-risk cohorts in the population, and the severity of BRD morbidity risk within those high-risk groups. A lower prevalence of high-risk cohorts will favor models that have a higher specificity. In contrast, a greater prevalence of high-risk cohorts will favor models that have a higher sensitivity. Depending on the expected prevalence of disease managers can determine which specific is the best option and if it is more beneficial to use a model compared to not using a model.

In our study, sensitivity represents the model's ability to correctly identify cohorts that are high-risk based on a selected cutoff; and specificity represents the model's ability to correctly identify cohorts that are low risk on a selected cutoff. Determining which metric to prioritize is dependent on the importance of minimizing false-positives or false-negatives and the cost of each outcome. False positives with this data would be a cohort that was truly low-risk (<15% BRD morbidity 45 DOF), but was predicted to be high-risk. False negatives would be a cohort that was truly high-risk (>= 15% BRD morbidity 45 DOF), but was predicted to be low-risk. The costs and consequences for each type of error are different. An increase in false positives may lead to additional unnecessary metaphylaxis treatment costs that are administered to cohorts that are at low expected risk for BRD. In our study the cost of a false-positive would be on average an extra $23.60 spent per animal for each false positive cohort that did not need metaphylactic treatment. False negatives may lead to negative health outcomes as cattle would not receive metaphylaxis for respiratory disease when they truly needed treatment. This misclassification can lead to increased health costs, losses in performance, and potentially increased mortality. In our study the cost of a false negative was the loss of value of a sick animal compared to a healthy animal, which was an estimated lost value of $151.18 per treated animal and the number of treated animals varies based on estimated morbidity.

The decision tree model has the highest potential economic advantage compared to other models when the proportion of high-risk cohorts present was ≤ 45%. For example, when 5% of the cohorts entering the feedlot are high-risk cohorts the decision tree model offers a potential $5/animal health cost advantage compared to human control. The decision tree model had the highest specificity which increases the model's ability to detect low-risk cohorts. Therefore, at this level of prevalence (5% high risk, 95% low risk) the increased specificity of the decision tree was more valuable ($5/animal health cost advantage) than the models with higher sensitivity. These results agree with a previous study that reported that increasing diagnostic test specificity increased economic net returns in comparison to increasing sensitivity (Theurer et al., 2015). However, in our study, the estimated health-cost advantage of the decision tree compared to the human control decreased as the proportion of high-risk cohorts increased. Once the prevalence of high-risk cohorts was 83% and above, the naïve Bayes model, which has the highest sensitivity, has an economic advantage over the control. However, at this level of high-risk cohorts feedlot managers would likely not distinguish between high- and low-risk cohorts, and would likely treat all groups with metaphylaxis, well before the proportion reaches 83%. This shows that the cost of using each model was dependent on the prevalence of high-risk cohorts that are arriving to the feedlot, which should be considered when determining the health-cost advantage to using these model compared to a human. The expected prevalence of incoming high-risk cohorts that will enter the feedlot is important to accurately estimate in order to determine the usefulness of utilizing a model to predict expected BRD morbidity risk.

A potential limitation may have been that the feedlots in our dataset only represented Midwest feedlots and the data may not represent cohorts of cattle from all feedlots in terms of the dates recorded, location, cattle types, and many other factors. Another limitation was that we did not have data indicating whether groups of cattle in our dataset had received metaphylactic treatment. This could have impacted our outcome of interest, BRD morbidity risk in the first 45 DOF, as we are not aware if cohorts in either the low or high risk category were previously mass treated. This could have affected what expected risk cohorts were placed into at the 15% treatment cutoff as we do not know if the percent of cattle treated for BRD in each cohort was affected by metaphylaxis. This could potentially impact our results by placing truly high-risk cattle into the incorrect classification. For example, if a truly high risk group of cattle were identified by feedlot personnel and received metaphylactic treatment then the overall percentage of cattle that were treated for BRD within the first 45 days on feed in this cohort could have potentially been lower than 15% due to metaphylaxis and they would not have been included in the high risk category in our study. As a result, there are likely cattle that are truly high-risk present in the low-risk category. The calculated sensitivity and specificity for our human control was representative of a subset of data that was available and does not represent the sensitivity and specificity for all feedlots. This was a small portion of data (n = 177 cohorts) and the generalizability of this data most likely does not reflect all feedlots.

## Conclusions

The objectives of this study were to evaluate the diagnostic performance of five classification models to classify incoming groups of cattle into high- and low-risk categories based on the BRD morbidity within the first 45 DOF and evaluate the models using an economic framework to determine if the models were advantageous to a person classifying expected risk.

We used AUC to evaluate model performance as this metric measures the models' degree of separability between high- and low- risk cohorts. Using area under the curve, the random forest model had the best performance with a value of .789 using the testing dataset. Although the random forest model had the highest AUC, it was not always the best model to use economically. The economic performance of each model was dependent on the prevalence of high-risk cohorts in the population. The decision tree provided the greatest estimated economic benefit when the proportion of high-risk cohorts was lower than 45% in the population. In addition to previously evaluated factors, this study provides a new outlook using arrival and pen housing factors to classify cohorts into risk categories. In order to further evaluate the true impact of these predictive models a prospective study should be considered to validate the true diagnostics and costs of using a predictive algorithm compared to current management strategies to determine expected BRD risk of incoming cohorts of cattle. In addition, more data, including new predictor variables and observations of data, are needed to continue to refine the algorithms and provide a better estimate of each models' predictive performance.

# Figures

**Figure 3.1. Flowchart of data refinement, data partitioning, algorithm training, and classification model algorithm evaluation.**

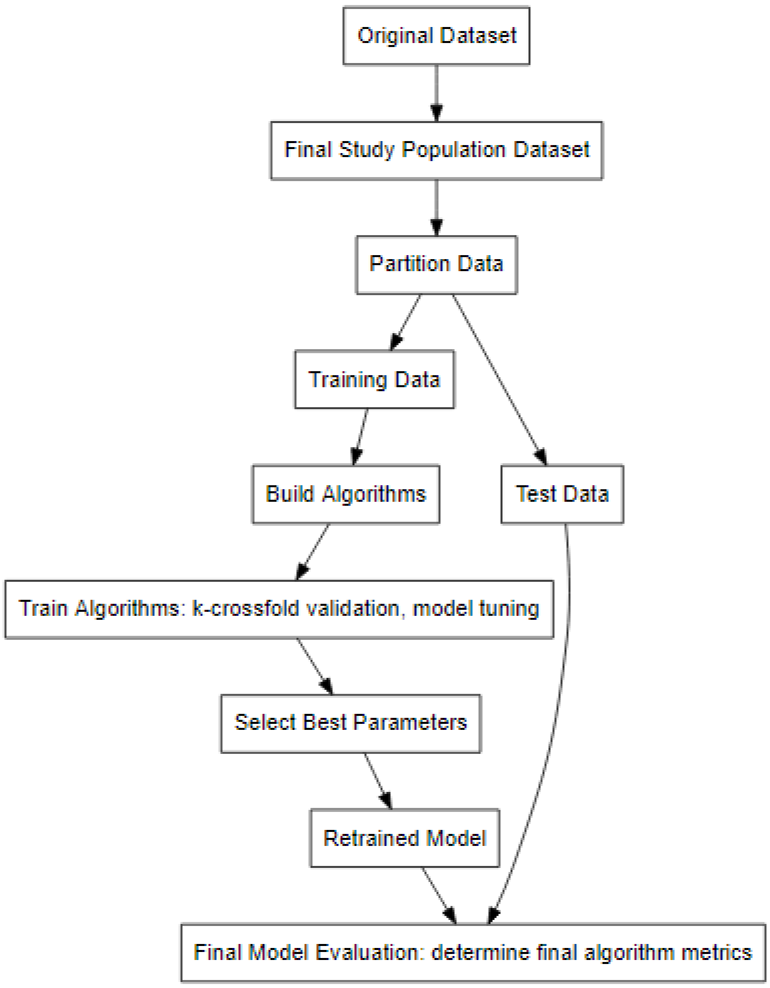


**Figure 3.2. Flowchart of diagnostic outcomes and calculations generated from predictive classification models using cutoff of 15% BRD incidence in the first 45 DOF.**

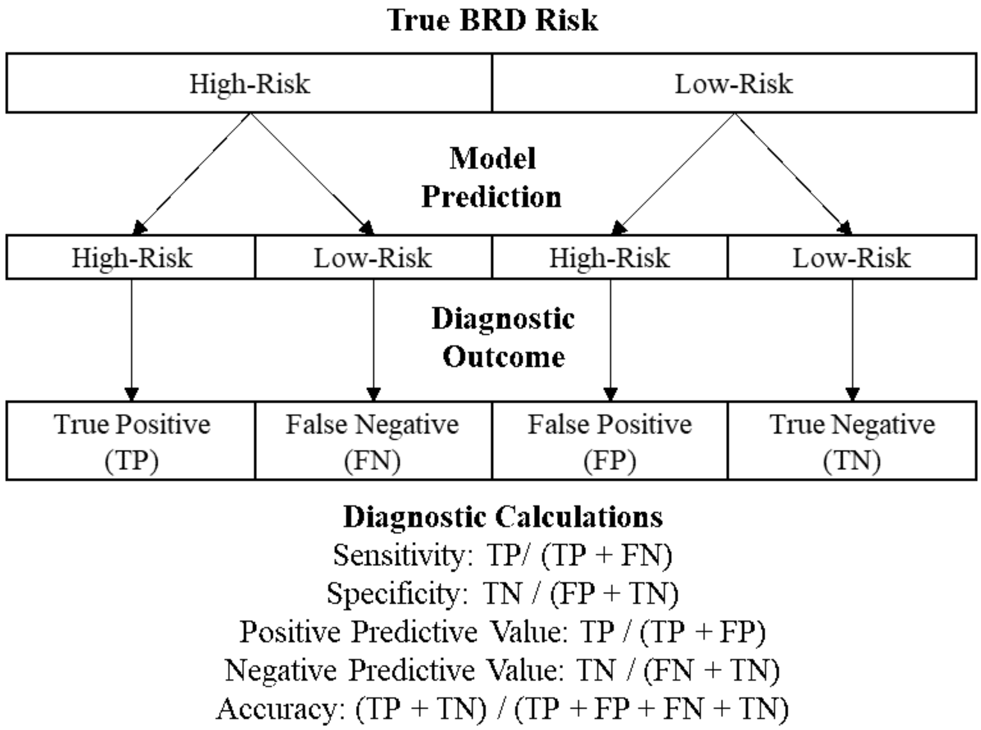
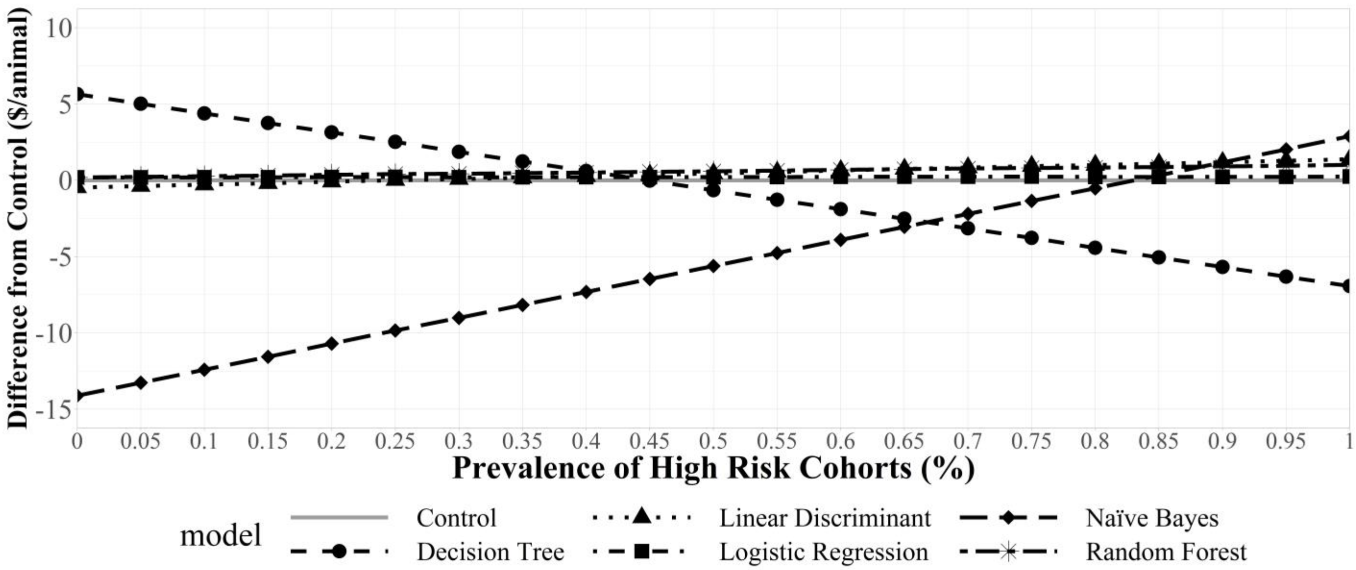


**Figure 3.3. Estimated economic results ($/animal) of the five classification model compared to a person (control) classifying expected BRD morbidity risk of incoming cohorts of cattle in the first 45 DOF across different proportions of high-risk cohorts to low-risk cohorts (0-100%). Gray line represents the control. At any prevalence, if the difference from the control for a model is above the control line ($0/animal) then it has a potential economic advantage relative to the control. If the difference from the control for a model is below the control line ($0/animal) then it has a potential economic disadvantage relative to the control.**



# Tables

**Table 3.1. Predictor and Outcome variables used in analysis of classification algorithms**

| Variable | Description |
|---|---|
| Cohort size at arrival | Total animals in cohort upon arrival to the feedlot |
| Average arrival weight at arrival | Total weight of all animals / cohort size at arrival |
| Arrival Date Quarter[1, 2] | Quarter of the year that cohort arrived (1,2,3,4) |
| Sex[1] | Gender of the cohort (steer, heifer, nixed gender) |
| Total pen area (sq. m) | Total area of the pen that cohorts were placed in |
| Bunk space length (m) | Total length of bunk available in pen |
| Pen area available per head (sq. m) | Total pen area / cohort size at arrival |
| Bunk space available per head (m) | Bunk space length / cohort size at arrival |
| BRD morbidity risk[3] | 1 = total cohort BRD morbidity risk >=15% |
| | 0 = total cohort BRD morbidity risk <15% |

[1] Qualitative variables that were converted to quantitative variables as dummy variables
[2] 1 (Jan, Feb, Mar), 2 (Apr, May, Jun), 3 (Jul, Aug, Sep), 4 (Oct, Nov, Dec)
[3] Binary outcome variable that we are predicting

**Table 3.2. Variables included in the economic analysis to compare the cost-benefit of using one of the predictive models compared to the control scenarios**

| Variable | Value |
|---|---|
| Total number of lots[1] | 1,733 |
| Average cohort size[2] | 109 |
| Cost of single metaphylactic treatment per animal[3] | $23.60 |
| Prevalence of High Risk cohorts (%) | 0-100 |
| Cost of morbid animal[4] | $151.18 |
| BRD morbidity % in true-positives[5] | 27% |
| BRD morbidity % in true-negatives[6] | 3% |
| Metaphylaxis efficacy | 0.5 |

[1] Total of number of cohorts in study population
[2] Average cohort size in study population
[3] Average cost per animal to administer metaphylaxis (United States Department of Agriculture, 2013a)
[4] Average cost of a sick animal (McNeil et al., 2001)
[5] Average morbidity in the true positive (high-risk) cohorts in study population
[6] Average morbidity in the true negative (low-risk) cohorts in study population

**Table 3.3. Classification model results and the financial consequence associated with following model predictions**

| Diagnostic Outcome | Truth | Model Prediction | Metaphylaxis Decision | Financial Consequence |
|---|---|---|---|---|
| TP | High-risk | High-risk | Treat | Animals that are truly high-risk are metaphylactically treated. Treatment costs are incurred, but expenses are saved by avoiding lost value from potential morbid animals. |
| FN | High-risk | Low-risk | Do not treat | Animals that are truly high-risk are not treated. These animals are expected to become morbid and provide lower value compared to healthy animals. The magnitude of financial loss is dependent on the prevalence of high-risk cohorts. |
| FP | Low-risk | High-risk | Treat | Animals that are truly low-risk are metaphylactically treated. These animals are expected to be healthy, but received treatment regardless, so the incurred costs are only the treatment cost of metaphylaxis for animals in each cohort. |
| TN | Low-risk | Low-risk | Do not treat | Animals that are truly low-risk are not metaphylactically treated. This is the baseline cost that is compared to all other outcomes. Since animals are expected to be healthy, and no treatment costs are incurred, the value for this outcome will always be $0. |

**Table 3.4. Final diagnostic performance estimates utilizing the test dataset for BRD morbidity risk during the first 45 days on feed to classify cohorts as high- or low- risk for BRD development within the first 45 days post-arrival using a 15% cutoff at the optimum cutoff where 10.85% of the cohorts had >15% BRD morbidity**

| Performance Metric | Logistic Regression | Decision Tree | Random Forest | Naïve Bayes | Linear Discriminant |
|---|---|---|---|---|---|
| AUC[1] | .785 | .682 | .789 | .743 | .760 |
| True Positives | 40 | 21 | 42 | 47 | 37 |
| False Positives | 152 | 63 | 153 | 386 | 141 |
| True Negatives | 234 | 323 | 233 | 0 | 245 |
| False Negatives | 7 | 26 | 5 | 0 | 10 |
| Accuracy % | 63.3 | 79.4 | 63.7 | 10.9 | 61.4 |
| Sensitivity % | 85.1 | 44.7 | 89.4 | 100.0 | 91.5 |
| Specificity % | 60.6 | 83.7 | 60.6 | 0.0 | 57.8 |
| PPV %[2] | 20.8 | 25.0 | 21.3 | 10.9 | 20.8 |
| NPV %[3] | 97.1 | 92.6 | 97.9 | DBZ[4] | 96.1 |

[1] AUC = Area under the curve
[2] PPV = Positive Predictive Value
[3] NPV = Negative Predictive Value
[4] DBZ = Division by zero (error)

# Chapter 4 - Impact of water sources and shared fence lines on bovine respiratory disease incidence in the first 45 days on feed


## Abstract

Bovine respiratory disease (BRD) is a major cause of morbidity in feedlot cattle. A retrospective analysis was performed using a population of U.S. feedlot cattle to determine potential associations between pen-level management factors related to number of water sources (one or two), shared water sources (yes/no), and shared fence lines (yes/no) with the risk of BRD incidence during the first 45 days on feed (DOF). Generalized linear mixed models were used to evaluate associations between management factors, cattle demographics, and BRD incidence.

Number of water sources, shared fence lines, and their interactions with cattle demographics were significantly associated with BRD incidence ($P < 0.05$). The effect of shared water sources was modified by cohort size. The effect of total number of water sources was modified by a cohort's average arrival weight. Shared fence lines and interactions between cattle demographics did not display biologically significant impacts on BRD incidence risk. The interaction between shared water sources and cohort size at arrival was statistically significant ($P < 0.05$); however, no biologically significant differences were observed between shared water sources across all average arrival weight categories.

Our results display that cattle demographics related to average arrival weight and cohort size at arrival may impact the effect of the number of water sources placed in a pen on BRD incidence risk in the first 45 DOF. Further investigation is needed to identify management strategies and methods related to water sources fence lines between that can be utilized to potentially influence BRD risk within the first 45 DOF in feedlot operations.

## Introduction

Bovine respiratory disease remains the primary cause of morbidity and mortality in feedlot cattle despite advancement in management and treatment protocols over the years. (Edwards, 2010). Management and control of BRD is difficult as it is a multifactorial disease with several risk factors that may contribute to onset of disease (Taylor et al., 2010). Numerous risk factors have been identified towards BRD; however, knowledge gaps remain between the relationship of management interventions and health outcomes in feedlot cattle. The spread and transmission of pathogenic organisms in cattle populations has been documented as a risk factor towards BRD and multiple stressors may predispose cattle to viral infection of the respiratory tract. (Griffin et al., 2010; Kilma et al., 2014). Understanding which interventions and management strategies are risk factors influencing BRD risk can be valuable towards mitigating BRD in feedlot cattle populations. A previous study investigated housing factors related to shared pen water and number of adjourning pens and their associations towards BRD risk (Hay et al., 2016). These studies reported associations between these risk factors and BRD incidence, but did not look at the potential interactions between these management risk factors and other known risk factors related to cattle demographics on BRD risk. Quantifying effects of management factors relative to BRD incidence for cattle after arrival may be useful for cattle managers to modify their current management techniques to reduce BRD incidence.

The objective of this study was to evaluate the potential associations between management characteristics, cattle demographics, and BRD incidence in the first 45 days on feed (DOF) in commercial feedlot cattle. The management factors we evaluated are related to pen water sources and fence lines separating pens. Little work has been done evaluating the relevance of number of shared waters or fence lines relative to BRD and this information could be useful for future management decisions. For example, if the number of shared water sources is found to be associated with BRD incidence, then managers could adjust the number of water sources shared for certain groups of cattle to potentially reduce the risk of BRD for that group of cattle. Our goal was to find information regarding potential management interventions that would fill important knowledge gaps and enhance understanding of management strategies that can be utilized by commercial feedlot operations to reduce BRD incidence.

## Materials and Methods

Animal Care and Use Committee approval was not obtained for this study as data were collected retrospectively from commercial feedlots.

Data were collected from 10 central U.S. high plains feedlots between January 2018 to April 2020 and included daily cohort- and individual-level information. Our outcome of interest was BRD incidence within the first 45 DOF as the majority of BRD onset occurs during this timeframe (Edwards 1996); therefore, all data were limited to events that occurred during the first 45 DOF. A cohort was defined as a group of cattle purchased and managed in a similar manner and housed together during the initial 45 DOF. Cohort-level data included demographic characteristics of the cattle which were sex, arrival date, average arrival weight, and cohort size at arrival (Table 1). A variable representing a cohort-pen ID combination was created to track cohorts and their physical housing location (pen) during the 45 DOF period. Cohorts included in the study were restricted to those housed in two or fewer pens within the first 45 DOF and were allowed to stay in their first pen for < 7 DOF as some animals may have been placed in receiving pens during the first few days of arrival. The pen characteristics utilized for the analysis were from the pen where cattle spent from at least day 7 to day 45. Any cohorts that were moved between 3 or more pens during the first 45 DOF were excluded from analysis. Individual animal data was filtered to include only animals with first treatment events for BRD within the first 45 DOF and were joined to their matching cohort.

Collected data were aligned with inclusion criteria, validated, categorized, and limited only to BRD-specific individual health events. Inclusion criteria included: remove cohorts with less than 25 animals, restrict average arrival weight between 227 kg to 453 kg (500 lb. to 1,000 lb.), and include heifer, steer, and mixed-sex cohorts. Health related events not recorded as BRD were excluded from analysis (AIP, bloat, musculoskeletal, etc.). Our case definition for BRD was the first time BRD was diagnosed in an individual animal by feedlot personnel and antimicrobial treatments administered during the first 45 DOF. Variables were categorized to avoid violation of the linearity assumption. Total cohort size at arrival (categorized as follows: 25 to 99, 100 to 175, >175) and average arrival weight (categorized as follows: 500 lb. to 599 lb. (227 kg to 272 kg), 600 lb. to 699 lb. (273 kg to 318 kg), 700 lb. to 799 lb. (319 kg to 363 kg), 800 lb. to 899 lb. (364 kg to 408 kg), 900 lb. to 1,000 lb. (409 kg to 453 kg)) were categorized similarly to previous literature (Babcock et al. 2009). Arrival dates were split into quartiles based on the arrival month to determine which quarter of the year the cohort entered the feedlot: January through March (Q1), April through June (Q2), July through September (Q3), and October through December (Q4). Cohorts with missing or incomplete data for any of these variables were excluded from the study population. The data came from a retrospective analysis from commercial feedlot data that consisted of information collected from multiple feedlot operations. As a result, vaccination programs from each operation, distance cattle traveled to the feedlot, and preconditioning status were unavailable. Additionally, BRD case definitions could have varied between feedlots and over time.

## Water Sources

Water sources placed in each pen were identified utilizing Google Earth Pro (version 7.3.3.7786) and the number of distinct water sources allocated per pen were counted. Cohort-level variables for water sources were created for each cohort which represented the number of water sources (NW) and shared water sources (SW). The definition for NW was defined as the number of usable water sources available to cattle housed in each pen. The variable for NW was categorized into a binary variable representing having access to either one or two water sources per pen (no pens had >2 water sources). The definition for SW was defined as the number of water sources located in a pen that can be accessed by animals from by one or more neighboring pens. The variable for SW was categorized into a binary variable representing no shared water sources or at least one shared water source. Any cohort that had missing water data was removed from the dataset for analysis

## Shared Fence Lines

A cohort-level variable for shared fence lines (SF) was assigned for each cohort. SF was defined as a fence line that is used to divide two separate pens from one another. The number of SF in each pen were documented utilizing Google Earth Pro (version 7.3.3.7786) and tallying the number of SF per pen. The variable for SF was categorized into a binary variable representing one shared fence line or two shared fence lines (no pens had >2 shared fence lines). If a cohort had missing data for shared fence lines it was removed from the dataset prior to analysis.

## Statistical Methods

Three distinct generalized linear mixed-models (NWmod, SWmod, and SFmod), one for each explanatory variable of interest (NW, SW, and SF), were fitted with the" glmer" function in the ‘lme4’ package in R (R Core Team 2021) to assess potential associations between the explanatory variables of interest and cattle demographics with BRD incidence in the first 45 DOF. A logit link function was utilized in each model. The outcome variable of interest in each model was BRD incidence in the first 45 DOF and was calculated as the total number of first BRD treatments in the first 45 DOF (events) / total animals in the pen (trials). Covariates included average arrival weight, cohort size at arrival, arrival date quarter, sex, and one of the three housing factors. Several interaction terms were incorporated in each model based on previously investigated factors determined to impact BRD incidence in feedlot cattle including: sex with average arrival weight; sex with cohort size at arrival; sex with arrival date quarter; average arrival weight with cohort size at arrival; average arrival weight with arrival date quarter; and arrival date quarter with cohort size at arrival. (Cernicchiaro et al., 2012; Babcock et al., 2013; Avra et al., 2017). Interactions between housing factors and cattle demographics were included in each model. For example: the NWmod evaluating potential associations between NW and BRD incidence in the first 45 DOF tested the 2-way interactions between each cattle demographic and NW. A random intercept for feedlot was included in each model to account for the hierarchical structure of the data. Variables that were determined *a priori* to be associated with BRD based on previous literature (quarter of arrival, arrival weight, sex, animal received) were retained in the model as fixed effects regardless of statistical significance. Remaining variables (including interactions) were retained only if they were significantly associated ($P < 0.05$) with the outcome or were part of a significant interaction term. All main effects were included regardless of significance if they were part of a significant ($P < 0.05$) interaction.

## Results

Our study population included 1,563 cohorts representing 168,482 individual animals from 10 feedlots from 2018 to 2020 (Table 2). There were 10,065 recorded cases of first treatment BRD within the initial 45 DOF representing 6.44% of the study population. Figure 1 displays the distribution of BRD incidence in the first 45 DOF in the study population.

Table 3 charts all variables and interactions that were significantly associated ($P < 0.05$) with BRD incidence in the first 45 DOF in all models. All final models (NWMod, SWMod, SFMod) included main effects of sex, cohort size at arrival, average arrival weight, and arrival quarter as-well-as significant ($P < 0.05$) 2-way interactions between sex and average arrival weight, and average arrival weight and arrival quarter. In NWMod and SFMod, 2-way interactions between cattle demographics and the respective covariate of interest (NW or SF) were all significantly ($P < 0.05$) associated with BRD incidence. In the SWmod, the only significant ($P < 0.05$) interaction between the cattle demographics and the covariate of interest (SW) was cohort size at arrival and SW.

The cohort size at arrival modified the effect of NW on BRD incidence in the first 45 DOF (Fig. 2). Cohorts with a cohort size at arrival of 100 to 175 animal as cohorts in this category had a higher BRD incidence in the first 45 DOF when provided one water source (5.50% ± .10) compared to cohorts that had two water sources (3.11% ± 0.56). Cohorts in other categories of cohort size at arrival (25 to 99, >175) had statistically similar BRD incidence regardless of the total number of water sources available.

The average arrival weight modified the effect of NW on BRD incidence in the first 45 DOF (Fig. 3). There was a significant contrast seen in cohorts with an arrival weight between 500 lb. to 599 lb. (227 kg to 272 kg) in our results. In this average arrival weight category, the

BRD incidence was higher when cattle had one water source (8.80% ± 1.50) compared to cattle that two water sources available (5.55% ± 0.98). There was also a significant contrast seen in cohorts with an arrival weight between 600 lb. to 699 lb. (273 kg to 218 kg) as the BRD incidence was lower when cattle had one total water source (8.17% ± 1.40) compared to cattle that had two water sources (11.60% ± 1.92). Cohorts in other average arrival weight categories did not display significant differences in BRD incidence regardless of NW available.

## Discussion

This study analyzed the associations between the number of water sources, shared water sources, and shared fence lines on BRD incidence in the first 45 DOF while accounting for cattle demographics that previously have been associated with BRD incidence. Quantifying the effects of this relationship is important to determine the potential associations cohort-level management factors and whether modifying these conditions could be used to mitigate BRD risk in feedlot cattle.

Prior research has investigated the relationship of these factors with BRD incidence in Australian feedlots (Hay et al, 2016). In this Australian report shared pen water (yes/no) and shared fence lines (one/two) were tested for association with BRD. The study did no incorporate the number of water sources in their analysis. In our study, the categories for SW and SF were similar between this study and previous research; however, our study evaluated three separate models incorporating interactions between housing factors of interest and cattle demographics. Each factor of interest was included in its own separate model (3 models) in order to assess the impact of each variable on BRD incidence on its own. Hay et al. (2016), reported that cattle with shared water sources were 4.3 times as likely (OR = 4.3, 95% CI = (1.4 to 10.3) to be treated for BRD compared to cattle without shared water sources (equivalent to our definition of shared water). In our study, the SWmod displayed a significant association between SW and the cohort size at arrival. Although this interaction was statically significant it was not found to have biological significance. Hay et al. (2016) also reported that the risk of BRD was not different between cattle that shared 1 fence lines and cattle that shared 2 fence lines (equivalent to our definition of shared fence line). In our study, the SFMod displayed that SF, as well as the interactions with all cattle demographics, were significantly ($P < 0.05$) associated with BRD incidence; however, although SF was statistically significantly associated with BRD incidence, there were no interactions that displayed a biological significance between SF, the interactions with other variables, and BRD incidence. As a result, in both studies there was no evidence of an association between shared fence lines and BRD incidence.

Previous research reported an association between the number of animals in a cohort and risk of BRD (Sanderson et al., 2008; Cernicchiaro et al., 2012). In our study, we observed a BRD incidence within the 100 to 175 cohort size at arrival category when there was one total water source available compared to two total water sources available. There were no significant differences in BRD incidence between the one or two total water sources within the 25 to 99 and >175 cohort size at arrival categories. It can be hypothesized that there are other unmeasured management protocols in place related to either variable that are influencing these results. Overall, the level of BRD incidence decreased as the size of the cohort at arrival increased. Cattle in cohorts with >175 animals displayed lower levels of BRD incidence which may be a result of low-risk cattle being placed in larger groups. However, our dataset did not capture risk classification of the cattle or management changes related to risk classification that might have varied across feedlots and thus no firm conclusions can be made.  Cattle in cohorts that are 25 to 99 animals in size displayed higher levels of morbidity which may be the result of high-risk animals being placed in smaller cohort sizes. In cohort sizes of 100 to 175 the level of morbidity varied depending on whether animals were given one or two sources of water. Animals with one water source displayed a higher BRD incidence compared to animals with two water sources in cohort sizes of 100 to 175 animals. Feedlot managers may be managing animals in this cohort size differently depending on the expected BRD risk of cattle. Additionally, we noticed that larger pens were more likely to have two water sources and smaller pens were likely to have one water source. An additional consideration beyond simp cohort size is that commingling has been documented as a risk factor for BRD and it can be inferred that the risk of BRD in different cohort sizes could be due to increased commingling in different sized groups (Step et al., 2008; Taylor et al., 2010). However, we can only measure the impact of the variables in our model and we were limited to cohort size without any additional metadata regarding commingling. In the NWmod, BRD incidence was similar between the 500 lb. to 599 lb. (227 kg to 272 kg) and 600 lb. to 699 lb. (273 kg to 218 kg) arrival weight categories which were our lowest arrival weight categories. These findings are consistent with previous research that reported light weight cattle are at higher risk for BRD compared to heavier cattle (Reinhardt et al., 2009; Taylor et al., 2010). In addition, in the NWmod BRD incidence was lowest in the heavier weight arrival weight categories (800 lb. to 899 lb. (364 kg to 408 kg) and 900 lb. to 1,000 lb. (409 kg to 453 kg)). Within the 500 lb. to 599 lb. (227 kg to 272 kg) categories BRD incidence was higher for cohorts that had one total water source available compared to cohorts that had two water. In contrast, within the 600 lb. to 699 lb. (273 kg to 218 kg) category cohorts with one total water source had a lower BRD incidence compared to cohorts with two water sources. It is possible that these results could be confounded by another factor such as a management decision on where to place high risk cattle arriving to the feedlot, but it is impossible to determine this from out data and results. In the SWmod the interaction between average arrival weight and SW was not statistically significant towards BRD incidence in the first 45 DOF. In the SFmod the interaction between average arrival weight and SF was statistically significant, but did not display a biological significance with BRD in the first 45 DOF.

One limitation of our study is that it is a retrospective analysis looking at pre-existing observational data and may be subject to confounding or bias. Results are also limited to feedlots that were included and may not be applicable to other feedlots due to differences in management, geography, cattle types, different case definitions for BRD, and many other potential differences. Because our data originated from commercial feedlots, it is inherently "messy" and may contain unknown biases or errors despite our efforts to clean and validate it. Additionally, we only evaluated BRD incidence during the first 45 DOF, which did not encompass all total BRD treatments throughout the entire feeding phase. It is possible that some of the risk factors we explored related to BRD incidence are significantly associated with risk of retreatment, risk of becoming a chronic animal, or risk of dying; our analysis did not include those outcomes due to limitations of the data provided in our dataset. Additional studies will be needed to evaluate those other important outcomes and their relationship with management-related risk factors. Several cohorts were also removed from the dataset if they were housed in more than 2 pens throughout the first 45 DOF. There are many possible reasons related to cattle flow and management decisions that may have caused cohorts to move several times throughout the first 45 DOF. As the objective of this study was to evaluate potential associations between pen waters and shared fence lines, it would become extremely difficult (if not impossible) to evaluate these characteristics from more than one pen and their potential associations with a single value of BRD incidence within the first 45 DOF. Finally, the risk status of cohorts entering the feedlot was not known in our dataset, so we could not determine which cohorts of cattle were classified as high risk for BRD at arrival. A future, well-controlled prospective study examining the risk of BRD during the first 45 DOF in association with pen housing conditions that includes some of this missing metadata should be conducted to help determine the differences in BRD incidence.

## Conclusion

Our study determined that BRD incidence in the first 45 DOF had significant interactions between management variables related to total and shared water sources and shared fence lines. However, these interactions did not display meaningful differences that can be used to manage BRD incidence in feedlot cattle. These management variables and their impact on BRD incidence have been briefly evaluated in previous literature, and our study further elaborates on the associations of those variables with BRD incidence early in the feeding period and how they are affected by other well-known cattle demographic risk factors. Results from this study provide estimates on how these management factors may be influencing the risk of initial BRD treatment for feedlot cattle. Additional research in this area will lead to a better understanding of the impacts of management conditions for feedlot cattle and how these conditions can potentially be modified to reduce the risk of BRD in the feedlot industry.

# Figures

**Figure 4.1. Histogram of the distribution of the level of BRD incidence in the first 45 days on feed (DOF) in the study population. The study population contained 1,563 cohorts which included 168,482 individual animals from 2018-2020. The x-axis displays the percentage of BRD incidence in a cohort during the first 45 DOF. The y-axis displays the count of cohorts for each value.**

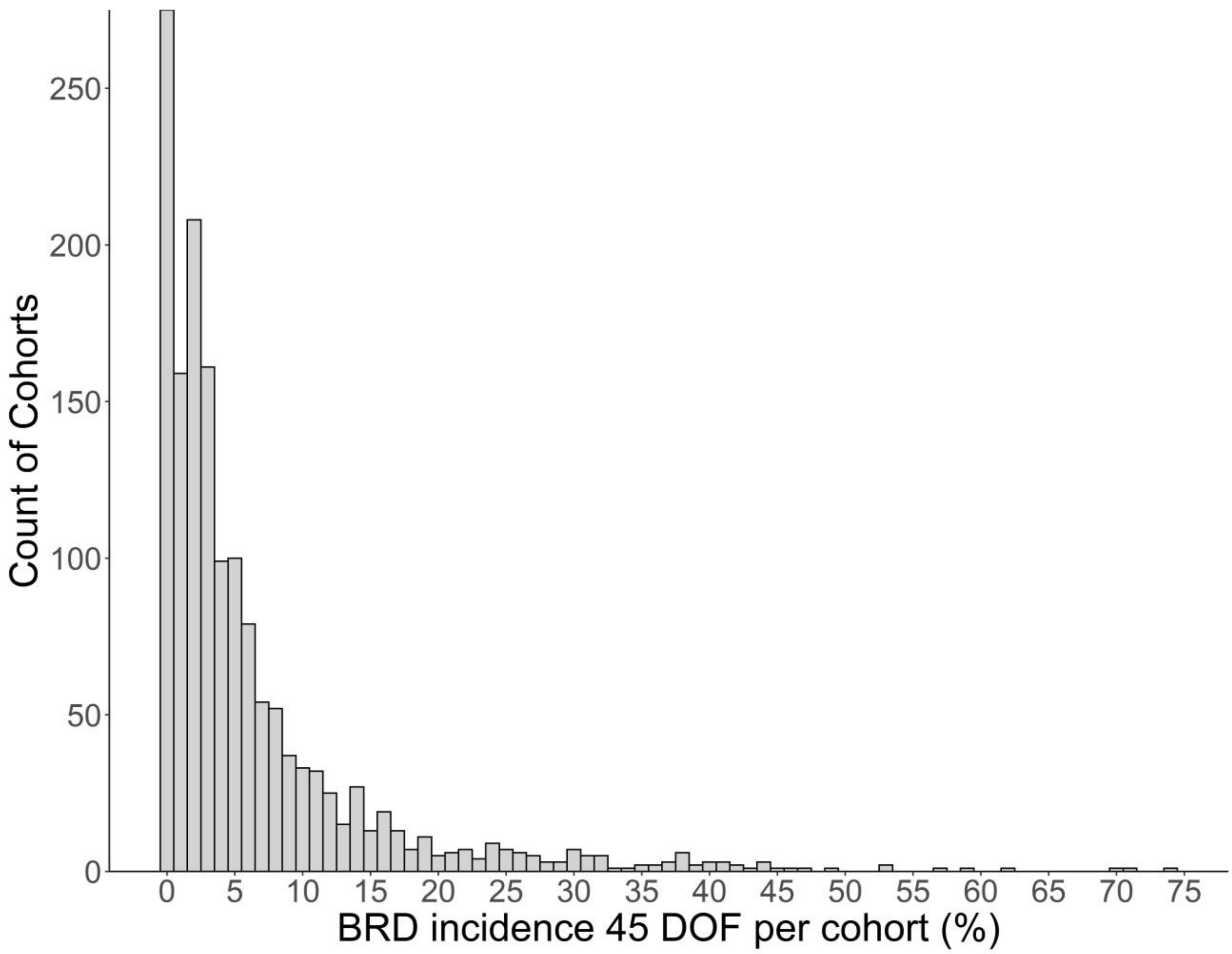


**Figure 4.2. Model estimated probabilities for BRD incidence in the first 45 DOF by total water sources (one source/two sources) and cohort size at arrival category (25-99, 100-175, >175 animals in each cohort) in commercial feedlot cattle during the first 45 DOF using the NWmod results. Error bars represent SE of the probability.**

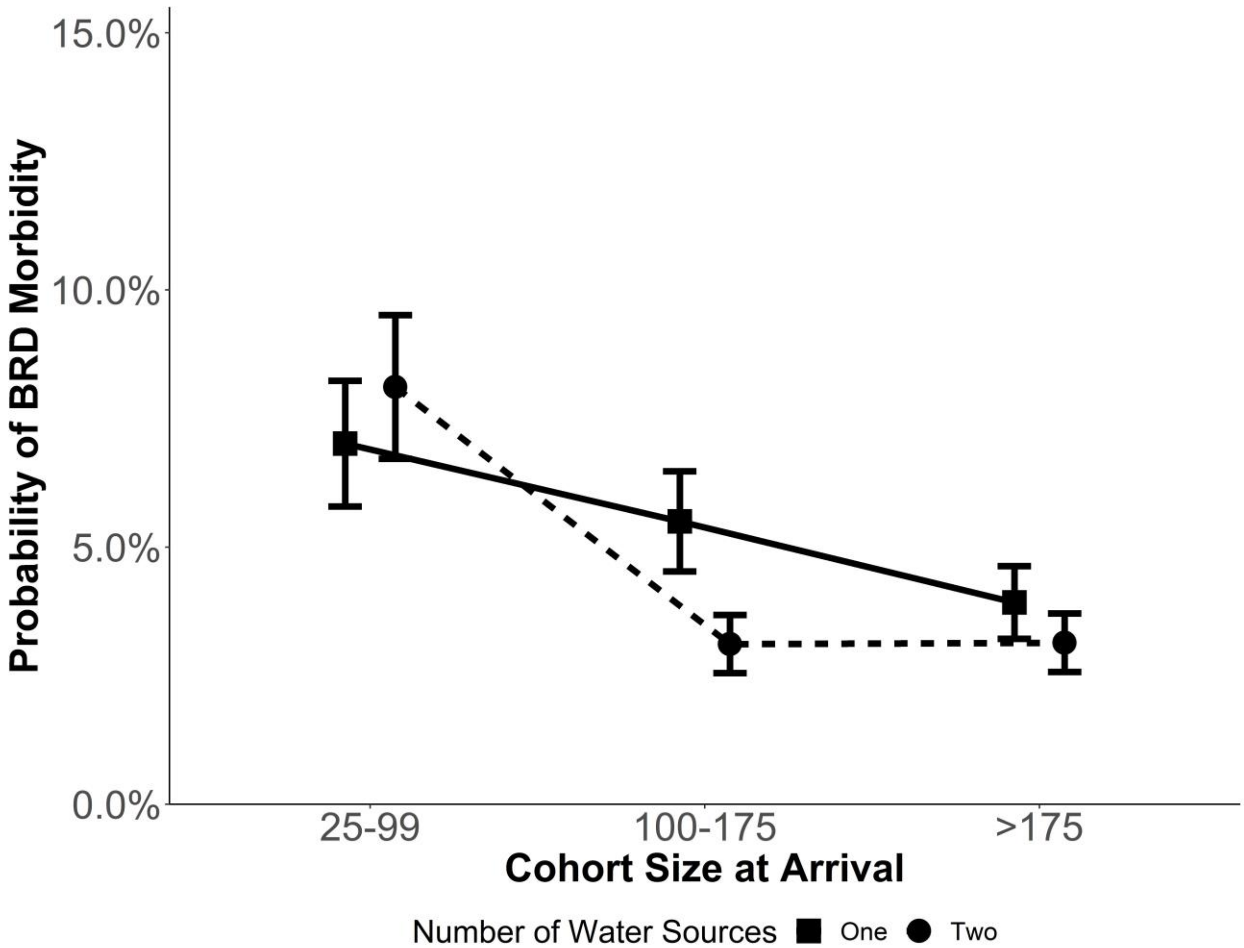
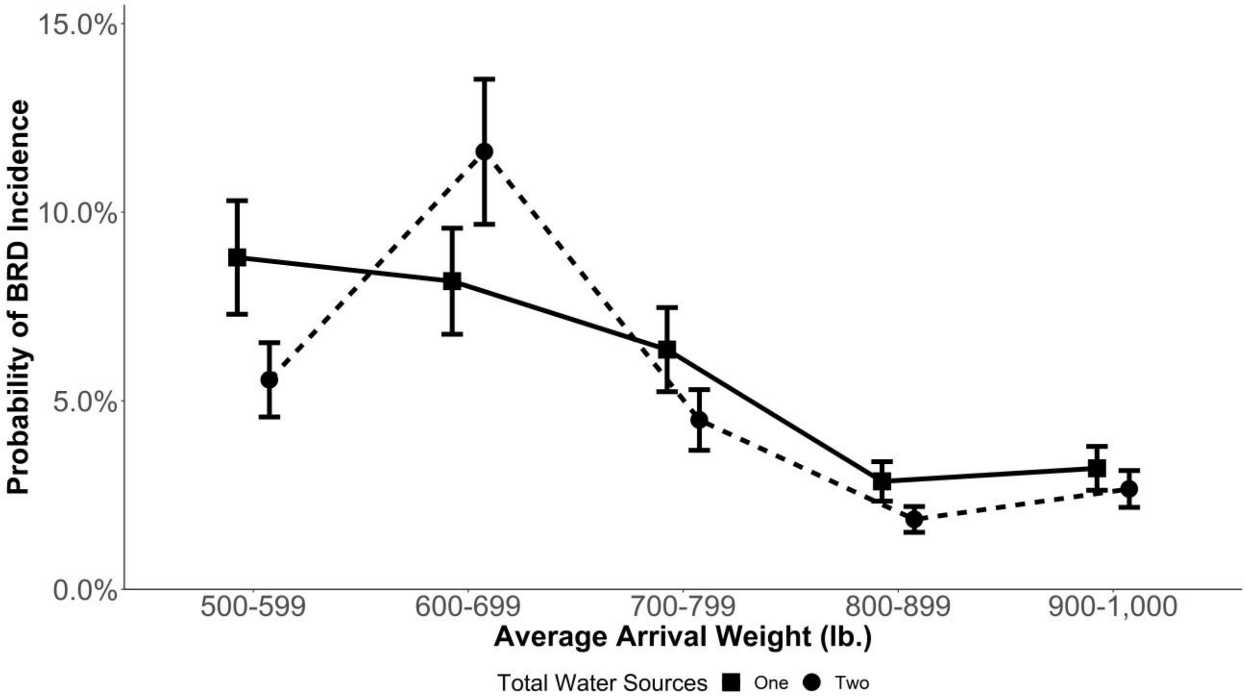


**Figure 4.3. Model estimated probabilities for BRD incidence in the first 45 DOF by total water sources (one source/two sources) and average arrival weight category (500-599 lb., 600-699 lb., 700-799 lb., 800-899 lb., and 900-1,000 lb.) in commercial feedlot cattle during the first 45 DOF using the NWmod results. Error bars represent SE of the probability.**



# Tables

**Table 4.1. Distribution of variables from the study population used for analysis from 10 feedlots from 2018-2020. The study population consisted of 1,563 cohorts which represented 168,482 animals.**

| Variable & Category | Number (%) of cohorts |
|---|---|
| Cohort size at arrival | |
| 25-99 | 854 (55.74) |
| 100-175 | 447 (29.18) |
| >175 | 231 (15.10) |
| Average Arrival Weight, lb. | |
| 500 – 599 | 116 (7.57) |
| 600 – 699 | 317 (20.69) |
| 700 – 799 | 558 (36.23) |
| 800 – 899 | 407 (26.57) |
| 900 – 1,000 | 134 (8.74) |
| Sex | |
| Heifers | 819 (53.46) |
| Steers | 599 (39.09) |
| Mixed | 114 (7.44) |
| Arrival Date Quarter | |
| Jan-March (1) | 398 (25.97) |
| April-June (2) | 433 (28.26) |
| July-September (3) | 419 (27.34) |
| October-December (4) | 282 (18.40) |
| Total Water Sources | |
| One source (0) | 1342 (87.60) |
| Multiple sources (1) | 190 (12.40) |
| Shared Pen Waters | |
| No (0) | 531 (31.10) |
| Yes (1) | 1,001 (68.9) |
| Shared Fence Lines | |
| One (0) | 428 (27.94) |
| Two (1) | 1,104 (72.06) |

**Table 4.2. Descriptive statistics for continuous variables in the study population cohorts (n = 1,563) from 10 feedlots. Data in study population consisted of 1,563 cohorts which represented 168,482 animals.**

| Variable | Mean | SD[2] | Median | Range | IQR[3] |
|---|---|---|---|---|---|
| Cohort Size at arrival | 108.69 | 55.88 | 87 | 25 to 324 | 64 to 144 |
| Average arrival weight, lb | 762.57 | 226.75 | 763 | 503 to 1,000 | 693 to 835 |
| BRD incidence[1], % | 6.44 | 9.00 | 3.26 | 0 to 74.07 | 1.28 to 7.84 |

[1] First treatment bovine respiratory disease (BRD) incidence was our outcome variable and was calculated only for the initial 45 days on feed
[2] SD = standard deviation
[3] IQR= Interquartile Range

**Table 4.3. Final generalized linear mixed-models demonstrating housing characteristics and cattle demographic factors and their association with bovine respiratory disease (BRD) incidence during the first 45 days on feed (DOF). The level of significance was set a P < 0.05 for each model. Three models (NWmod, SWmod, SFmod) for each hosing factor of interest (number of waters, shared waters, and shared fence lines) were evaluated.**

| Variable | NWmod *P*-Values | SWmod *P*-Values | SFmod *P*-Values |
|---|---|---|---|
| Sex | <0.01 | <0.01 | <0.01 |
| Cohort size at arrival | <0.01 | <0.01 | <0.01 |
| Average arrival weight | <0.01 | <0.01 | <0.01 |
| Arrival date quarter | <0.01 | <0.01 | <0.01 |
| NW[1] | <0.01 | ---[4] | --- |
| SW[2] | --- | 0.27 | --- |
| SF[3] | --- | --- | <0.01 |
| Sex x Average arrival weight | <0.01 | <0.01 | <0.01 |
| Average arrival weight x Arrival date quarter | <0.01 | <0.01 | <0.01 |
| Sex x NW | <0.01 | --- | --- |
| Average arrival weight x NW | <0.01 | --- | --- |
| Cohort size at arrival NW | <0.01 | --- | --- |
| Arrival date quarter x NW | <0.01 | --- | --- |
| Sex x SW | --- | 0.77 | --- |
| Average arrival weight x SW | --- | 0.07 | --- |
| Cohort size at arrival SW | --- | <0.01 | --- |
| Arrival Date Quarter x SW | --- | 0.63 | --- |
| Sex x SF | --- | --- | <0.01 |
| Average arrival weight x SF | --- | --- | <0.01 |
| Cohort size at arrival x SF | --- | --- | <0.01 |
| Arrival Date Quarter x SF | --- | --- | <0.01 |

[1] NW = number of water sources

[2] SW = shared water sources

[3] SF = shared fence lines

[4] Dashed lines (---) signify that the variable was not incorporated into the model

# Chapter 5 - Conclusions

BRD in feedlot cattle continues to negatively impact animal health and performance which ultimately leads to incurred treatment costs and lost value and is the costliest disease affecting the beef cattle industry. BRD is multifactorial disease and thus several studies focusing on the epidemiology and potential risk factors towards BRD have been performed. Although numerous factors associated with BRD risk in the feedlot phase of production have been identified gaps in knowledge of BRD epidemiology remain. Research focusing on pen-level management factors related to pen housing conditions have been studied, but are limited. The purpose of this thesis was to continue to evaluate the associations and predictability of pen housing conditions towards BRD risk in the first 45 DOF in order to determine if these factors may be used by feedlot operations to decrease risk of BRD and increase animal health/welfare and performance.

Management of morbidity within a feedlot environment includes focusing on proper health protocols. Feedlot managers may implement certain management strategies to manage cattle health and reduce the risk of BRD. Management practices may be modified to manage the BRD risk of animals. One area where managers may implement changes is the pen housing conditions that cattle are placed in. This includes the amount of pen area and bunk space that are allocated to an incoming group of cattle. We conducted a study to evaluate the associations between pen area per head and bunk space per head, along with cattle demographics, on BRD incidence in the first 45 days on feed. Results displayed that pen area per head and bunk space per head were significantly associated with BRD incidence, but the effect of these factors were modified by cattle demographics. For example, heavy weight cattle experienced an increased risk of BRD incidence in the first 45 DOF when provided pen area per head and bunk space per head below industry recommendations. Managers may be able to use these management factors to manage BRD risk, but further investigation of the associations with BRD incidence is needed to gain a better understanding of their impact on cattle health.

In addition to pen area and bunk space, water sources and amount of shared fence lines may also play a role in a management decision related to animal health protocols. The number of water sources, as well as if they are shared with a neighboring pen, may be a factor that managers consider when placing cattle in pens. In addition, the number of shared fence lines with neighboring pens may also be considered. We conducted a study to evaluate the effect of these factors on BRD incidence in the first 45 DOF. Results from this study demonstrated that water sources and shared fence lines do not have a large impact on the risk of BRD incidence.

The ability to predict the expected BRD risk of incoming cattle to a feedlot operation is valuable and can provide tremendous health and economic benefits to a feedlot production system. Previous research has evaluated the use of predictive modeling to predict health outcomes of cattle in a feedlot setting. We evaluated the ability of several classification algorithms to identify incoming cohorts that are expected to be at high risk for BRD. Results demonstrated the models have a low (AUC = 0.6 to 0.69) to average (AUC = 0.7 to 0.79) ability to predict the expected BRD based on the area under the curve of each classification algorithm using the 15% cutoff. In addition, an economic analysis was performed on the results from each models' final diagnostics to determine if there is a potential economic health cost advantage to using a predictive algorithm compared to a human to classify expected risk. Our results from the economic analysis revealed that most models provided a comparable or lower economic advantage in relation to a human classifying expected risk. The decision tree model displayed the highest potential economic advantage in relation to the human, but this was dependent on the prevalence of high-risk cohorts that were in the population. This study lays the ground work towards attaching economic values to predictive algorithm outputs in order to determine if using a predictive algorithm would be beneficial to a feedlot operation in place of management protocols already set in place.

# References


Amrine, D. E., McLellan, J. G., White, B. J., Larson, R. L., Renter, D. G., & Sanderson, M. (2019). Evaluation of three classification models to predict risk class of cattle cohorts developing bovine respiratory disease within the first 14 days on feed using on-arrival and/or pre-arrival information. *Computers and Electronics in Agriculture*, *156*, 439–446. https://doi.org/10.1016/j.compag.2018.11.035

Amrine, D. E., White, B. J., & Larson, R. L. (2014). Comparison of classification algorithms to predict outcomes of feedlot cattle identified and treated for bovine respiratory disease. *Computers and Electronics in Agriculture*, *105*, 9–19. https://doi.org/10.1016/j.compag.2014.04.009

Avra, T. D., Abell, K. M., Shane, D. D., Theurer, M. E., Larson, R. L., & White, B. J. (2017). A retrospective analysis of risk factors associated with bovine respiratory disease treatment failure in feedlot cattle1. *Journal of Animal Science*, *95*(4), 1521–1527. https://doi.org/10.2527/jas.2016.1254

Babcock, A. H., Cernicchiaro, N., White, B. J., Dubnicka, S. R., Thomson, D. U., Ives, S. E., Scott, H. M., Milliken, G. A., & Renter, D. G. (2013a). A multivariable assessment quantifying effects of cohort-level factors associated with combined mortality and culling risk in cohorts of U.S. commercial feedlot cattle. *Preventive Veterinary Medicine*, *108*(1), 38–46. https://doi.org/10.1016/j.prevetmed.2012.07.008

Babcock, A. H., White, B. J., Dritz, S. S., Thomson, D. U., & Renter, D. G. (2009). Feedlot health and performance effects associated with the timing of respiratory disease treatment1. *Journal of Animal Science*, *87*(1), 314–327. https://doi.org/10.2527/jas.2008-1201

Babcock, A. H., White, B. J., Renter, D. G., Dubnicka, S. R., & Scott, H. M. (2013b). Predicting cumulative risk of bovine respiratory disease complex (BRDC) using feedlot arrival data and daily morbidity and mortality counts. *Canadian Journal of Veterinary Research = Revue Canadienne De Recherche Veterinaire*, *77*(1), 33–44.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Modelings Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Boyer, W., Davidson, J., George, H., Graber, R., Murphy, J. P., & Harner, J. (2017). Guidelines for Planning Cattle Feedlots. *Kansas State University Research & Exetension*. https://bookstore.ksre.ksu.edu/pubs/MF3392.pdf

Cawley, G. C., & Talbot, N. L. C. (2010). On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation. *Journal of Machine Learning Research*, *11*, 2079–2107.

Cernicchiaro, N., Renter, D. G., White, B. J., Babcock, A. H., & Fox, J. T. (2012). Associations between weather conditions during the first 45 days after feedlot arrival and daily respiratory disease risks in autumn-placed feeder cattle in the United States. *Journal of Animal Science*, *90*(4), 1328–1337. https://doi.org/10.2527/jas.2011-4657

Chirase, N. K., & Greene, L. W. (2001). Dietary zinc and manganese sources administered from the fetal stage onwards affect immune response of transit stressed and virus infected offspring steer calves. *Animal Feed Science and Technology*, *93*(3–4), 217–228. https://doi.org/10.1016/S0377-8401(01)00277-2

Edwards, A. (1996). Respiratory diseases of feedlot cattle in central USA. *The Bovine Practitioner*, 5-7 Pages. https://doi.org/10.21423/BOVINE-VOL1996NO30P5-7

Edwards, T. A. (2010). Control Methods for Bovine Respiratory Disease for Feedlot Cattle. *Veterinary Clinics of North America: Food Animal Practice*, *26*(2), 273–284. https://doi.org/10.1016/j.cvfa.2010.03.005

Friedman, F., Hastie, T., & Tibshirani, R. (2010). Regularization Paths for Generalized Models via Coordinate Descent. *Journal of Statistical Software*, *33*(1), 1–22.

Gardner, I. A., & Greiner, M. (2006). Receiver-operating characteristic curves and likelihood ratios: Improvements over traditional methods for the evaluation and application of veterinary clinical pathology tests. *Veterinary Clinical Pathology*, *35*(1), 8–17. https://doi.org/10.1111/j.1939-165X.2006.tb00082.x

Google Earth Pro. (2020). [Computer software] (7.3.3.7786). https://www.google.com/earth/index.html

Griffin, D., Chengappa, M. M., Kuszak, J., & McVey, D. S. (2010). Bacterial Pathogens of the Bovine Respiratory Disease Complex. *Veterinary Clinics of North America: Food Animal Practice*, *26*(2), 381–394. https://doi.org/10.1016/j.cvfa.2010.04.004

Hastie, T., & Tibshirani, R. (202 C.E.). *mda: Mixture and Flexible Discriminant Analysis*. https://CRAN.R-project.org/package=mda

Hay, K. E., Barnes, T. S., Morton, J. M., Clements, A. C. A., & Mahony, T. J. (2014). Risk factors for bovine respiratory disease in Australian feedlot cattle: Use of a causal diagram-informed approach to estimate effects of animal mixing and movements before feedlot entry. *Preventive Veterinary Medicine*, *117*(1), 160–169. https://doi.org/10.1016/j.prevetmed.2014.07.001

Hay, K. E., Morton, J. M., Clements, A. C. A., Mahony, T. J., & Barnes, T. S. (2016). Associations between feedlot management practices and bovine respiratory disease in Australian feedlot cattle. *Preventive Veterinary Medicine*, *128*, 23–32. https://doi.org/10.1016/j.prevetmed.2016.03.017

Ives, S. E., & Richeson, J. T. (2015). Use of Antimicrobial Metaphylaxis for the Control of Bovine Respiratory Disease in High-Risk Cattle. *The Veterinary Clinics of North*

*America. Food Animal Practice*, *31*(3), 341–350, v. https://doi.org/10.1016/j.cvfa.2015.05.008

Kayser, W. C., Carstens, G. E., Jackson, K. S., Pinchak, W. E., Banerjee, A., & Fu, Y. (2019). Evaluation of statistical process control procedures to monitor feeding behavior patterns and detect onset of bovine respiratory disease in growing bulls. *Journal of Animal Science*, *97*(3), 1158–1170. https://doi.org/10.1093/jas/sky486

Kingsford, C., & Salzberg, S. L. (2008). What are decision trees? *Nature Biotechnology*, *26*(9), 1011–1013. https://doi.org/10.1038/nbt0908-1011

Klima, C. L., Zaheer, R., Cook, S. R., Booker, C. W., Hendrick, S., Alexander, T. W., & McAllister, T. A. (2014). Pathogens of bovine respiratory disease in North American feedlots conferring multidrug resistance via integrative conjugative elements. *Journal of Clinical Microbiology*, *52*(2), 438–448. https://doi.org/10.1128/JCM.02485-13

Koehler, G. J., & Erenguc, S. S. (1990). Minimizing Misclassifications in Linear Discriminant Analysis*. *Decision Sciences*, *21*(1), 63–85. https://doi.org/10.1111/j.1540-5915.1990.tb00317.x

Kuhn, M., & Johnson, K. (2020). *Feature engineering and selection: A practical approach for predictive models*. CRC Press, Taylor & Francis Group.

Kuhn, M., & Vaughan, D. (2021). *yardstick: Tidy Characterizations of Model Performance*. https://CRAN.R-project.org/package=yardstick

Kuhn, M., & Wickham, H. (2020). *Tidymodels: A collection of packages for modeling and machine learning using tidyverse principles.* https://www.tidymodels.org

Liaw, A., & Wiener, M. (2001). Classificatoin and Regression by RandomForest. *Forest*, *23*.

Majka, M. (2019). *naivebayes: High Performance Implementation of the Naive Bayes Algorithm in R*. https://CRAN.R-project.org/package=naivebayes

Mandrekar, J. N. (2010). Receiver operating characteristic curve in diagnostic test assessment. *Journal of Thoracic Oncology: Official Publication of the International Association for the Study of Lung Cancer*, *5*(9), 1315–1316. https://doi.org/10.1097/JTO.0b013e3181ec173d

McNeill, J. W. (2001). *2000-2001 Texas A&M Ranch to Rail—North/South Summary Report*. Texas Agricultural Extension Service, Texas A&M University. College Station, TX.

Menardi, G., & Torelli, N. (2014). Training and assessing classification rules with imbalanced data. *Data Mining and Knowledge Discovery*, *28*(1), 92–122. https://doi.org/10.1007/s10618-012-0295-5

Miles, D. G. (2009). Overview of the North American beef cattle industry and the incidence of bovine respiratory disease (BRD). *Animal Health Research Reviews*, *10*(2), 101–103. https://doi.org/10.1017/S1466252309990090

Myles, A. J., Feudale, R. N., Liu, Y., Woody, N. A., & Brown, S. D. (2004). An introduction to decision tree modeling. *Journal of Chemometrics*, *18*(6), 275–285. https://doi.org/10.1002/cem.873

Petrie, A., & Watson, P. F. (2013). *Statistics for veterinary and animal science* (3rd ed). Blackwell Pub. Professional.

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation For Statistical Computing. https://www.R-project.org/

Reinhardt, C. D., Busby, W. D., & Corah, L. R. (2009). Relationship of various incoming cattle traits with feedlot performance and carcass traits1. *Journal of Animal Science*, *87*(9), 3030–3042. https://doi.org/10.2527/jas.2008-1293

Safari, S., Baratloo, A., Elfil, M., & Negida, A. (2016). Evidence Based Emergency Medicine; Part 5 Receiver Operating Curve and Area under the Curve. *Emergency (Tehran, Iran)*, *4*(2), 111–113.

Sanderson, M. W., Dargatz, D. A., & Wagner, B. A. (2008). Risk factors for initial respiratory disease in United States' feedlots based on producer-collected daily morbidity counts. *The Canadian Veterinary Journal = La Revue Veterinaire Canadienne*, *49*(4), 373–378.

Snowder, G. D., Van Vleck, L. D., Cundiff, L. V., & Bennett, G. L. (2006). Bovine respiratory disease in feedlot cattle: Environmental, genetic, and economic factors. *Journal of Animal Science*, *84*(8), 1999–2008. https://doi.org/10.2527/jas.2006-046

Step, D. L., Krehbiel, C. R., DePra, H. A., Cranston, J. J., Fulton, R. W., Kirkpatrick, J. G., Gill, D. R., Payton, M. E., Montelongo, M. A., & Confer, A. W. (2008). Effects of commingling beef calves from different sources and weaning protocols during a forty-two-day receiving period on performance and bovine respiratory disease1,2. *Journal of Animal Science*, *86*(11), 3146–3158. https://doi.org/10.2527/jas.2008-0883

Taylor, J. D., Fulton, R. W., Lehenbauer, T. W., Step, D. L., & Confer, A. W. (2010). The epidemiology of bovine respiratory disease: What is the evidence for predisposing factors? *The Canadian Veterinary Journal = La Revue Veterinaire Canadienne*, *51*(10), 1095–1102.

Tenny, S., & Hoffman, M. R. (2021). Prevalence. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK430867/

Therneau, T., & Atkinson, B. (2019). *rpart: Recursive Partitioning and Regression Trees*. https://CRAN.R-project.org/package=rpart

Theurer, M. E., White, B. J., Larson, R. L., & Schroeder, T. C. (2015). A stochastic model to determine the economic value of changing diagnostic test characteristics for identification of cattle for treatment of bovine respiratory disease. *Journal of Animal Science*, *93*(3), 1398–1410. https://doi.org/10.2527/jas.2014-8487

United States Department of Agriculture. (2013a). *Types and Cost of Respiratory Disease Treatments in U.S. Feedlots*. https://www.aphis.usda.gov/animal_health/nahms/feedlot/downloads/feedlot2011/Feed11_is_RespDis_1.pdf

United States Department of Agriculture. (2013b). *Feedlot 2011. Part IV: Health and Health Management on U.S. Feedlots with a Capacity of 1,000 or More Head*. https://www.aphis.usda.gov/animal_health/nahms/feedlot/downloads/feedlot2011/Feed11_dr_PartIV_1.pdf

Webb, G. I. (2016). Naïve Bayes. In C. Sammut & G. I. Webb (Eds.), *Encyclopedia of Machine Learning and Data Mining* (pp. 1–2). Springer US. https://doi.org/10.1007/978-1-4899-7502-7_581-1

Wiegand, J. B., Cooke, R. F., Brandão, A. P., Schubach, K. M., Colombo, E. A., Sowers, C., Duff, G. C., & Gouvêa, V. N. (2020). Impacts of commingling on health and productive responses of beef heifers during feedlot receiving. *Translational Animal Science*, *4*(Supplement_1), S79–S83. https://doi.org/10.1093/tas/txaa102

Wright, M. N., & Ziegler, A. (2017). A Fast Implementation of Random Forests for High Dimensional Data in C++and R. *Journal of Statistical Software*, *77*(1). https://doi.org/10.18637/jss.v077.i01

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, *3*(1), 32–35. https://doi.org/10.1002/1097-0142(1950)3:1<32::aid-cncr2820030106>3.0.co;2-3

# Appendix A - Supplementary material

**Chapter 2 supplementary table 1**

| Interaction | Estimated Probability | Standard Error |
|---|---|---|
| Arrival Weight, kg x Area per Animal, $m^2$ ($P$ < 0.5) | | |
| 227 – 272 x ≤23.22 | 6.46% | 1.18 |
| 273 – 318 x ≤23.22 | 9.21% | 1.63 |
| 319 – 363 x ≤23.22 | 5.10% | 0.94 |
| 364 – 408 x ≤23.22 | 3.04% | 0.57 |
| 409 – 454 x ≤23.22 | 3.60% | 0.64 |
| 227 – 272 x 23.23-32.52 | 8.73% | 1.55 |
| 273 – 318 x 23.23-32.52 | 8.91% | 1.58 |
| 319 – 363 x 23.23-32.52 | 5.49% | 1.01 |
| 364 – 408 x 23.23-32.52 | 2.21% | 0.42 |
| 409 – 454 x 23.23-32.52 | 2.05% | 0.39 |
| 227 – 272 x >32.52 | 7.24% | 1.31 |
| 273 – 318 x >32.52 | 6.94% | 1.26 |
| 319 – 363 x >32.52 | 5.54% | 1.02 |
| 364 – 408 x >32.52 | 2.02% | 0.39 |
| 409 – 454 x >32.52 | 2.47% | 0.47 |
| Arrival Weight, kg x Bunk Space per Animal, m ($P$ < 0.05) | | |
| 227 – 272 x ≤0.3 | 6.86% | 1.24 |
| 273 – 318 x ≤0.3 | 7.92% | 1.42 |
| 319 – 363 x ≤0.3 | 5.15% | 0.95 |
| 364 – 408 x ≤0.3 | 1.98% | 0.38 |
| 409 – 454 x ≤0.3 | 1.91% | 0.37 |
| 227 – 272 x 0.31-0.46 | 8.66% | 1.54 |
| 273 – 318 x 0.31-0.46 | 7.76% | 1.39 |
| 319 – 363 x 0.31-0.46 | 5.88% | 1.08 |
| 364 – 408 x 0.31-0.46 | 2.56% | 0.49 |
| 409 – 454 x 0.31-0.46 | 3.34% | 0.63 |
| 227 – 272 x >0.46 | 6.88% | 1.25 |
| 273 – 318 x >0.46 | 9.29% | 1.64 |
| 319 – 363 x >0.46 | 5.12% | 0.95 |
| 364 – 408 x >0.46 | 2.68% | 0.51 |
| 409 – 454 x >0.46 | 2.86% | 0.54 |
| Cohort Size at Arrival x Area per Animal, $m^2$ ($P$ < 0.05) | | |
| 25-99 x ≤23.22 | 6.36% | 1.16 |

| | | |
|---|---|---|
| 100-175 x ≤23.22 | 4.62% | 0.86 |
| >175 x ≤23.22 | 4.46% | 0.83 |
| 25-99 x 23.23-32.52 | 6.50% | 1.18 |
| 100-175 x 23.23-32.52 | 5.20% | 0.96 |
| >175 x 23.23-32.52 | 2.83% | 0.54 |
| 25-99 x >32.52 | 6.86% | 1.24 |
| 100-175 x >32.52 | 4.70% | 0.87 |
| >175 x >32.52 | 2.40% | 0.46 |
| Cohort Size at Arrival x Bunk Space per Animal, m ($P < 0.05$) | | |
| 25-99 x ≤0.3 | 5.88% | 1.08 |
| 100-175 x ≤0.3 | 4.24% | 0.79 |
| >175 x ≤0.3 | 2.66% | 0.50 |
| 25-99 x ≤0.3 | 7.30% | 1.32 |
| 100-175 x ≤0.3 | 5.65% | 1.04 |
| >175 x ≤0.3 | 3.20% | 0.60 |
| 25-99 x ≤0.3 | 6.60% | 1.20 |
| 100-175 x ≤0.3 | 4.71% | 0.87 |
| >175 x ≤0.3 | 3.57% | 0.67 |
| Arrival Date Quarter x Area per Animal, $m^2$ ($P < 0.05$) | | |
| Jan-March (1) x ≤23.22 | 5.15% | 0.95 |
| April-June (2) x ≤23.22 | 4.07% | 0.76 |
| July-September (3) x ≤23.22 | 5.48% | 1.01 |
| October-December (4) x ≤23.22 | 5.81% | 1.06 |
| Jan-March (1) x 23.23-32.52 | 3.48% | 0.65 |
| April-June (2) x 23.23-32.52 | 4.22% | 0.79 |
| July-September (3) x 23.23-32.52 | 4.10% | 0.77 |
| October-December (4) x 23.23-32.52 | 7.29% | 1.32 |
| Jan-March (1) x >32.52 | 4.39% | 0.82 |
| April-June (2) x >32.52 | 4.32% | 0.80 |
| July-September (3) x >32.52 | 3.07% | 0.58 |
| October-December (4) x >32.52 | 5.75% | 1.05 |
| Arrival Date Quarter x Bunk Space per Animal, m ($P < 0.05$) | | |
| Jan-March (1) x ≤0.3 | 3.48% | 0.65 |
| April-June (2) x ≤0.3 | 3.88% | 0.73 |
| July-September (3) x ≤0.3 | 3.43% | 0.64 |
| October-December (4) x ≤0.3 | 5.84% | 1.07 |
| Jan-March (1) x 0.31-0.46 | 5.35% | 0.98 |
| April-June (2) x 0.31-0.46 | 4.29% | 0.80 |
| July-September (3) x 0.31-0.46 | 4.60% | 0.85 |
| October-December (4) x 0.31-0.46 | 6.44% | 1.17 |
| Jan-March (1) x >0.46 | 4.22% | 0.79 |

| | | |
|---|---|---|
| April-June (2) x >0.46 | 4.46% | 0.83 |
| July-September (3) x >0.46 | 4.40% | 0.82 |
| October-December (4) x >0.46 | 6.48% | 1.18 |
| Sex x Area per Animal, $m^2$ (*P* <0.05) | | |
| Heifers x ≤23.22 | 4.42% | 0.82 |
| Steers x ≤23.22 | 4.87% | 0.90 |
| Mixed x ≤23.22 | 6.09% | 1.11 |
| Heifers x 23.23-32.52 | 3.70% | 0.69 |
| Steers x 23.23-32.52 | 4.48% | 0.83 |
| Mixed x 23.23-32.52 | 5.81% | 1.06 |
| Heifers x >32.52 | 3.83% | 0.72 |
| Steers x >32.52 | 4.01% | 0.75 |
| Mixed x >32.52 | 5.10% | 0.94 |
| Sex x Bunk Space per Animal, m (*P* <0.05) | | |
| Heifers x ≤0.3 | 3.78% | 0.71 |
| Steers x ≤0.3 | 3.42% | 0.64 |
| Mixed x ≤0.3 | 5.17% | 0.95 |
| Heifers x 0.31-0.46 | 4.01% | 0.75 |
| Steers x 0.31-0.46 | 6.27% | 1.14 |
| Mixed x 0.31-0.46 | 5.28% | 0.97 |
| Heifers x >0.46 | 4.13% | 0.77 |
| Steers x >0.46 | 4.07% | 0.76 |
| Mixed x >0.46 | 6.61% | 1.20 |

**Chapter 3 supplementary figure 1**