Pakistan Veterinary Journal

ISSN: 0253-8318 (PRINT), 2074-7764 (ONLINE)
DOI: 10.29261/pakvetj/2024.136

RESEARCH ARTICLE

A Nomogram Prediction Model for *Mycobacterium avium subspecies paratuberculosis* based on Individual Dairy Herd Improvement Information for Dairy Cows

Mingcheng Wang, Daoqi Liu, Ye Wang, Huili Xia, Chaoying Liu and Gailing Wang*

College of Biological and Food Engineering, Huanghuai University, Zhumadian, Henan 463000, China
*Corresponding author: wanggailing@huanghuai.edu.cn

ARTICLE HISTORY (23-514)

Received: November 25, 2023
Revised: February 4, 2024
Accepted: February 12, 2024
Published online: March 03, 2024

Key words:
Mycobacterium avium subspecies paratuberculosis
Dairy herd improvement
Nomogram
Predictive model

A B S T R A C T

This study developed a nomogram model utilizing dairy cow-level risk factors to predict the risk of *Mycobacterium avium subspecies paratuberculosis* (MAP) infection. MAP antibody status was detected by ELISA in 1,589 dairy cows on commercial farms in Henan Province, China. Dairy Herd Improvement (DHI) data was also collected for each cow. Univariate analysis was used to identify MAP risk factors and multivariate logistic regression with backward bootstrap screening was used to determine the independent predictor for inclusion in the nomogram model. Model performance was evaluated by area under the receiver operating characteristic curve (AUC), calibration plots, and decision curve analysis. Finally, 1,481 cows with complete data were included, with a 24.9% MAP positive rate (n=369). The nomogram model demonstrated good discrimination (AUC 0.71) and accuracy (70.2%). Calibration was excellent (Hosmer-Lemeshow $\chi 2$=3.26, *P*=0.92), and decision curve analysis indicated this predictive model has clinical utility for diagnostic testing. The nomogram predicted individual MAP risk based on routinely available DHI data including age, milk production, mammary health status, milk losses, and milk fat. Our study provides a method for screening high-risk dairy cows and developing intervention strategies based on DHI reports.

To Cite This Article: Wang M, Liu D, Wang Y, Xia H, Liu C and Wang G, 2024. A nomogram prediction model for *Mycobacterium avium subspecies paratuberculosis* based on individual dairy herd improvement information for dairy cows. Pak Vet J, 44(1): 105-110. http://dx.doi.org/10.29261/pakvetj/2024.136

INTRODUCTION

Mycobacterium avium subsp. paratuberculosis (MAP) is a Gram-positive, ubiquitous, obligate intracellular and intestinal pathogen which can cause paratuberculosis or Johne's disease (JD) (Rathnaiah *et al.*, 2017). A chronic intestinal causing by MAP infection mainly occurred in ruminants and many other species, including rabbits, horses, pigs, and deer (Ozsvari *et al.*, 2020). Moreover, MAP infection has been associated with human Crohn's disease and ulcerative colitis (Pierce, 2018). In China, the detection rates of 14.2% and 11.2% were recorded in dairy cows and Tibetan Sheep respectively (Ma *et al.*, 2019; Cheng *et al.*, 2020). Paratuberculosis causing by poor milk production, discarded milk, early culling, veterinary services, and labor costs had resulted in a huge economic losses, Paratuberculosis significantly decreased the economic development by 461-1,940 EUR/case in France (Dufour *et al.*, 2004).

The main clinical symptoms of paratuberculosis in dairy cows include persistent diarrhea, progressive weight loss and milk yield decreasing. While, not all animals exposed to MAP show clinical signs (Nielsen and Toft, 2009). The diagnosis method of paratuberculosis includes clinical diagnosis and subclinical infection diagnosis. Subclinical infection diagnosis is essential to achieve control of paratuberculosis. In adult cows with chronic diarrhea is often indicative of this disease. In the laboratory, antigen and antibodies are often used for diagnosing MAP subclinical infection. Antigen-based tests can be made by isolating the pathogens from feces or diseased tissue, by histological features study of the lesions, and by polymerase chain reaction (PCR) assays (Vilar *et al.*, 2015). However, as MAP shedding can be intermittent, the sensitivity of these methods depends on the development processes of the disease (Nielsen, 2008). Serum- and milk- ELISA may potentially be useful for detecting subclinical paratuberculosis in lactating dairy cows, with good sensitivities and specificities compared to fecal culture, as well as its high efficiency and cost effectiveness (Hendrick *et al.*, 2005).

In dairy cows, milk production and composition are the two most important economic traits. Although variation in milk production traits and functional traits

have a major genetic component (Raven *et al.*, 2014). Many studies have found that MAP was associated with dairy cows' milk production and several well-known milk components closely, such as milk fat, milk protein, Somatic cell count (SCC), and etc. (Machado *et al.*, 2018). Furthermore, milk is an ideal sample material for experiments, as milk is obtained non-invasively from dairy cows regularly (Milovanović *et al.*, 2020). Published articles have developed a mixed model based on milk production and milk composition changing, which has a higher ability for MAP predicting and distinguishing (Machado *et al.*, 2018). However, poor visualization of intuitive may inhibit its potential applications.

Nomogram is a statistics-based tool, that consists of multiple risk factors in a graphical display, and is used to simply predict the overall probability of a specific outcome (Shi *et al.*, 2020). Nomograms based on multiple logistic regression or the Cox Proportional Hazards regression model are widely used to calculate the risk of clinicopathological cancer features and have been developed to assess the survival rate of patients with cancer (Song and Fu, 2019).

With the above background, we used MAP infection data (defined by serum-ELISA results) to analyze the possible risk factors of dairy cow level, and established a nomogram prediction model. Here, we aimed to establish a visualization and efficient model for MAP, identify high-risk dairy cows and provide opportunities for timely intervention.

MATERIALS AND METHODS

Sample size: The sample size was determined by the number of independent variables in the prediction model. The sample size is at least ten-fold more than the number of independent variables (2). Assuming that each factor requires at least 10 cases for verification and the follow-up loss rate is 10%, calculated by the formula: 13×10÷0.9=144, therefore, the suitable sample size should be at least 144 cases.

1598 dairy cows from 12 dairy herds from Henan subjected to MAP testing were provided by the Henan Dairy Herd Improvement Center (HNDHI) between April 2020 and May 2021. Samples submission was voluntary by ranchers and based on animals that shown MAP-like illness. Subjected dairy cows were classified as infected solely on serum MAP ELISA (IDEXX Laboratories Inc., Maine, USA) obtained, positive or negative, according to the manufacturer's instructions.

Explanatory variables: At the same time, DHI reports for lactating dairy cows from the current month and the previous month were collected, excluding those for newly lactating cows. The following information and data were extracted which include birth date, lactation numbers, daily milk production, fat and protein percentage in milk, milk losses, milk urea nitrogen values, days in milk, WHI, and total estimated milk production. Two consecutive somatic cell count (SCC) values were used to assess the mammary gland status of the dairy cows. WHI values were used to evaluate the performance level of the cows in the groups. The stage of lactation is assessed using both current and last milk production values. The variables explanation is shown in Table 1.

Statistical analyses: Only dairy cows with complete data were included in the statistical analysis.

In the present study, statistical analyses, construction, validation and evaluation of the prediction model were performed in R 4.0.2 software (R Foundation, Vienna, Austria, www.r-project.org). Continuous variables are presented as median and range (M (P25, P75)). Univariate logistic analysis was performed to identify MAP-related risk factors. The variables with p <0.1 in the univariate logistic analyses were further assessed by multivariable logistic regression with a backward stepwise method based on the Akaike information criterion (AIC) minimum, in R 4.0.2 software. In univariate logistic regression analysis, factors with a p -value less than 0.1 ($p < 0.1$) are selected for inclusion in a multivariate logistic regression analysis. Subsequently, within the multivariate logistic regression framework, only factors demonstrating a p -value less than 0.05 ($p < 0.05$) are retained in the final model.

Nomogram constructed: A prediction nomogram to assess the risk of MAP was constructed based on the results from the final multivariable logistic regression using the rms package (Li *et al.*, 2021) in R 4.0.2 software. Before determining the final model, the variance inflation factor of each variable was calculated.

Model performance evaluation: The established model of performance focusing on discrimination, calibration, and clinical usefulness was analyzed. The predictive ability of the model for MAP in cows was performed by calculating the area under the curve (AUC), using the pROC package (Robin *et al.*, 2011) in R 4.0.2 software, where 0.5< AUC <0.7 indicates that the model's distinguishing ability is low; 0.7< AUC <0.9 indicates a moderate ability and AUC>0.9 reflects a higher discrimination ability (He *et al.*, 2021).

In addition, the calibration curve was drawn, and 1000 bootstrap validations were performed to assess predictive accuracy. The Hosmer-Lemeshow test was used to evaluate the calibration of the prediction model. P >0.05 revealed a small difference between the predicted value of the model and the actual observation, indicating the model calibration is good (Crowson *et al.*, 2016).

Finally, the clinical effectiveness of the model was evaluated by clinical decision curve analysis, using the rmda package (Shao *et al.*, 2023) in R version 4.0.2, to determine the clinical utility of the nomogram by quantifying the net benefit at different threshold probabilities.

RESULTS

After exclusion, 1481 dairy cows were eligible and included in this study. Out of the included cases, 369 cases exhibited MAP serology positive, accounting for about 24.9% of the total cases (Table 2). 13 variables were statistically analyzed, and the significant variables derived from the univariate analysis (p<0.1) were further subjected to multivariate logistic regression analysis. Following the single factor analysis, significant differences were observed in Days of lactation, Lactation stage, Number of lactations, Milk losses, Mammary gland status, Daily milk production, within herd index (WHI), and Milk fat. All of these significant factors were included in multivariate analysis, using a backward stepwise method, based on the AIC minimum.

Table 1: Predictor variables definition

Predictor variables	Definition
Mammary gland status -Uninfected	Last time SCC＜500 × 10^3cells/mL and this time SCC＜500 × 10^3cells/mL or New lactating diary cow SCC＜500 × 10^3cells/mL
Mammary gland status - New infection	Last time SCC＜500 × 10^3cells/mL and this time SCC≥500 × 10^3cells/mL or New lactating diary cow SCC≥500 × 10^3cells/mL
Mammary gland status - Chronic mastitis	Last time SCC ≥ 500 × 10^3cells/mL and this time SCC≥ 500 × 10^3cells/mL
Mammary gland status - Cure	Last time SCC ≥ 500 × 10^3cells/mL and this time SCC＜500 × 10^3cells/mL
WHI-High level	WHI values ＞100
WHI-Low level	WHI values ＜100
Lactation stage-UP	Current milk production - last milk production ≥ 0
Lactation stage-Down	Current milk production - last milk production < 0

Table 2: Diary cattle and disease characteristics

Category	All	Negative	Positive
	N=1481	N=1135	N=346
Ages(years)			
< 3.5y	711 (48.0%)	622 (54.8%)	89 (25.7%)
≥ 3.5y	770 (52.0%)	513 (45.2%)	257 (74.3%)
Days of lactation	183 (100; 265)	177 (97; 259)	204 (113; 300)
Birth seasons			
Winter (Dec - Feb)	501 (33.8%)	375 (33.0%)	126 (36.4%)
Spring (March - May)	301 (20.3%)	236 (20.8%)	65 (18.8%)
Summer (Jun - Aug)	264 (17.8%)	198 (17.4%)	66 (19.1%)
Autumn (Sep - Nov)	415 (28.0%)	326 (28.7%)	89 (25.7%)
Daily milk production (Kg)			
< 28.4	509 (34.4%)	352 (31.0%)	157 (45.4%)
≥ 28.4	972 (65.6%)	783 (69.0%)	189 (54.6%)
Milk fat (Kg)			
< 4.3	764 (51.6%)	571 (50.3%)	193 (55.8%)
≥ 4.3	717 (48.4%)	564 (49.7%)	153 (44.2%)
Milk protein (Kg)			
< 3.7	1249 (84.3%)	966 (85.1%)	283 (81.8%)
≥ 3.7	232 (15.7%)	169 (14.9%)	63 (18.2%)
Number of lactations			
1th	646 (43.6%)	565 (49.8%)	81 (23.4%)
2th	384 (25.9%)	278 (24.5%)	106 (30.6%)
3th	249 (16.8%)	167 (14.7%)	82 (23.7%)
4th	125 (8.44%)	76 (6.70%)	49 (14.2%)
5th	45 (3.04%)	27 (2.38%)	18 (5.20%)
6th	32 (2.16%)	22 (1.94%)	10 (2.89%)
Mammary gland status			
Uninfected	1213 (81.9%)	968 (85.3%)	245 (70.8%)
New infection	118 (7.97%)	82 (7.22%)	36 (10.4%)
Category	All	Positive	Negative
	N=1481	N=1135	N=346
Chronic mastitis	81 (5.47%)	35 (3.08%)	46 (13.3%)
Cure	69 (4.66%)	50 (4.41%)	19 (5.49%)
Milk Urea Nitrogen(mg/dl)			
< 15.7	632 (42.7%)	492 (43.3%)	140 (40.5%)
≥ 15.7	849 (57.3%)	643 (56.7%)	206 (59.5%)
Milk losses (%)			
< 0.16	944 (63.7%)	774 (68.2%)	170 (49.1%)
≥ 0.16	537 (36.3%)	361 (31.8%)	176 (50.9%)
Lactation stage			
Up	626 (42.3%)	504 (44.4%)	122 (35.3%)
Down	855 (57.7%)	631 (55.6%)	224 (64.7%)
WHI			
Low level	787 (53.1%)	574 (50.6%)	213 (61.6%)
High level	694 (46.9%)	561 (49.4%)	133 (38.4%)
Estimated total milk production (Kg)			
< 9193	1119 (75.6%)	888 (78.2%)	231 (66.8%)
≥ 9193	362 (24.4%)	247 (21.8%)	115 (33.2%)

Table 3: Univariate and multivariate analyses of predictors for MAP

	Univariate analysis		Multivariate analysis	
Characteristics	OR (95% CI)	*P*	OR (95% CI)	*P*
Age(years)				
Less than 3.5	Ref	Ref	Ref	Ref
More than 3.5	3.50 (2.68-4.58)	<0.001	1.80 (1.16-2.81)	<0.001
Days of lactation	1.00 (0.98-1.02)	<0.001		
Birth seasons				
Winter (Dec - Feb)	Ref	Ref		
Spring (March - May)	0.82 (0.58-1.15)	0.25		
Summer (Jun - Aug)	0.99 (0.7-1.40)	0.96		
Autumn (Sep - Nov)	0.81 (0.6-1.11)	0.19		
Milk production daily (Kg)				
Less than 28.4	Ref	Ref	Ref	Ref
More than 28.4	0.54 (0.42-0.69)	<0.001	0.62 (0.47-0.81)	0.002
Milk fat (Kg)				
Less than 4.27	Ref	Ref		
More than 4.27	0.80 (0.63-1.02)	0.08	0.77 (0.59-0.99)	0.04
Milk protein (Kg)				
Less than 3.65	Ref	Ref		
More than 3.65	1.27 (0.93-1.75)	0.14		
Number of lactations				
1th	Ref	Ref		
2th	2.66 (1.93-3.67)	<0.001		
3th	3.43 (2.41-4.87)	<0.001		
4th	4.50 (2.93-6.90)	<0.001		
5th	4.65 (2.45-8.82)	<0.001		
6th	3.17 (1.45-6.94)	0.004		
Mammary gland status				
Uninfected	Ref	Ref	Ref	Ref
New infection	1.73 (1.14-2.63)	0.009	1.47 (0.76-2.01)	0.37
Chronic mastitis	5.19 (3.27-8.24)	<0.001	3.22 (1.66-4.78)	<0.001
Cure	1.50 (0.87-2.59)	0.15	1.05 (0.59-1.87)	0.82
Milk Urea Nitrogen (mg/dl)				
Less than 15.7	Ref	Ref		
More than 15.7	1.13 (0.88-1.44)	0.34		
Milk losses (%)				
Less than 0.16	Ref	Ref	Ref	Ref
More than 0.16	2.22 (1.74-2.84)	<0.001	1.39 (1.02-1.89)	0.02
Lactation stage				
Up	Ref	Ref		
Down	1.47 (1.14-1.88)	0.003		
WHI				
Low level	Ref	Ref		
High level	0.64 (0.5-0.82)	<0.001		
Estimated total milk production (Kg)				
Less than 9193	Ref	Ref		
More than 9193	1.79 (1.37-2.33)	<0.001		

According to the results of backward stepwise regression, the model containing Age, Milk fat, Milk losses, Mammary gland status, and Daily milk production, the minimal value of AIC is 1474.1. No collinearity was observed between screened variables (max VIF <2.7). In the multivariate analysis, the result revealed that ages (p=0.01), milk fat (p=0.04), Milk losses (p=0.03), chronic mastitis (p<0.001), and milk production (p=0.01), were significant independent risk factors. The univariate and multivariate analyses are listed in Table 3.

A model containing these independent predictors was developed and presented as a nomogram, which included 5 significant predictors for MAP prediction (Fig. 1). The assignment of predictors in the nomogram is shown in Table 4. The score for each predictive factor was obtained from the nomogram, and the total score was calculated as the sum of these individual scores. The corresponding total score values show the predicted probability of MAP. Then we established ROC curves to evaluate the diagnostic value, sensitivity and specificity, and threshold value (Fig. 2). The area under the curve (AUC) of the

Table 4: Assignment table of predictors in nomogram

predictors	Assignment
Age	1= < 3.5y; 2= ≥ 3.5y
Daily milk production	1= < 38.4Kg; 2= ≥ 38.4Kg
Mammary gland status	1= Uninfected; 2=New infection;3= Chronic mastitis; 4= Cure
Milk losses	1= < 0.16%; 2= ≥ 0.16%
Milk fat	1= < 4.3Kg; 2= ≥ 4.3Kg

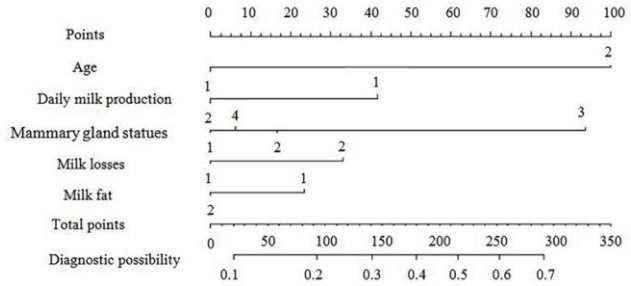

Fig. 1: Risk nomogram model for predicting MAP.

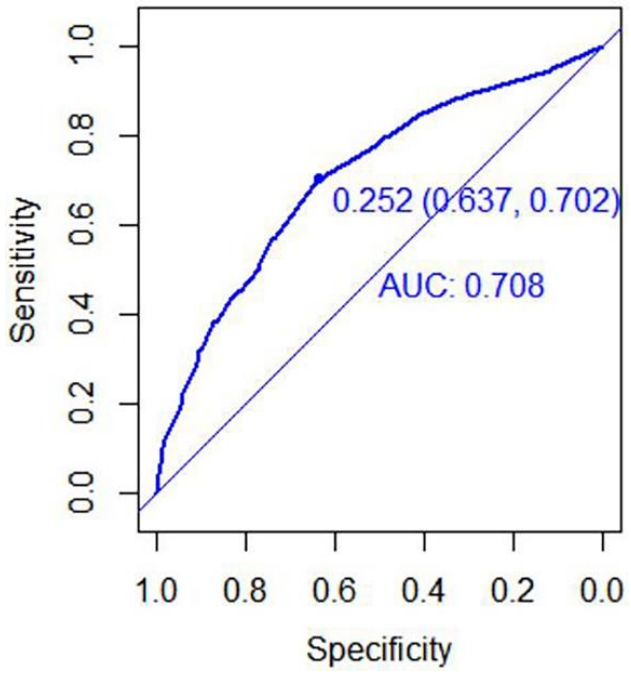

Fig. 2: The ROC of the risk nomogram model.

ROC curve or C-statistics—was 0.71 (95% CI: 0.68-0.74) and its sensitivity and specificity were 0.63 and 0.70, respectively, indicating a certain predictive effect. The threshold value of the ROC curve was 0.25. If the probability of risk was greater than 0.25, the risk of MAP was high.

The performance of this nomogram was graphically evaluated by a calibration curve. The calibration prediction curve fitted with the ideal curve (Fig. 3), which demonstrated the relatively acceptable goodness-of-fit of the nomogram. In addition, the Hosmer–Lemeshow goodness-of-fit test showed $\chi2 = 3.26$, $p = 0.92$, indicating no significant difference between predicted and observed probabilities, which had good consistency.

Decision curve analysis (DCA), a novel method, was used to evaluate the clinical efficiency and benefits of the prediction model. Black indicates that all samples are negative, therefore the net benefit is 0. Grey indicates that all samples are positive. The x-axis represents the threshold probabilities of dairy cows. As shown in Fig. 4,

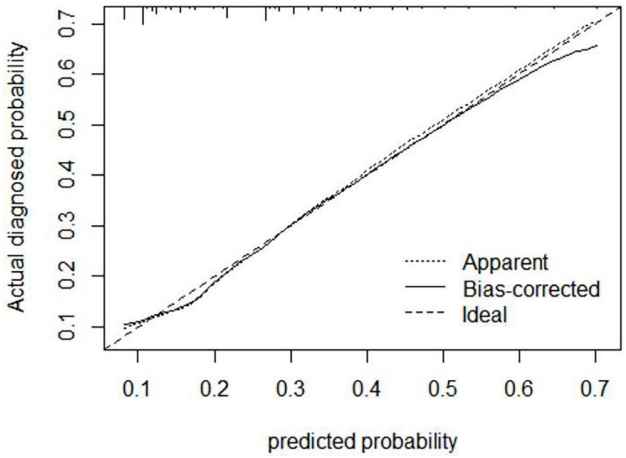

Fig. 3: The calibration degree of the risk nomogram model.

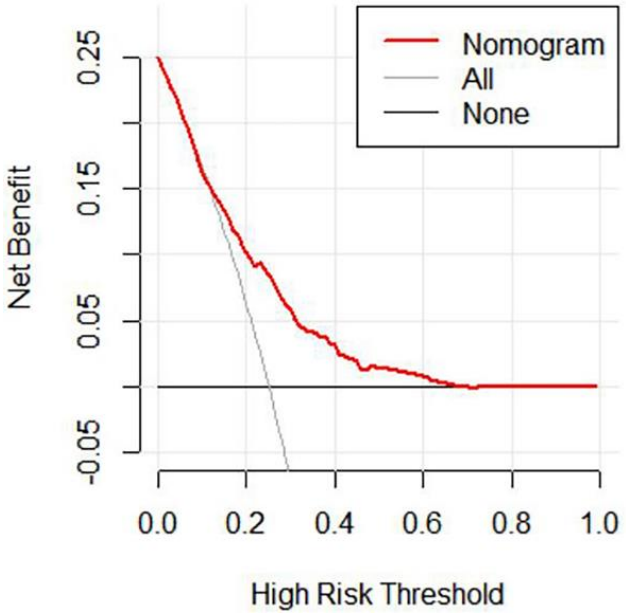

Fig. 4: The decision curve of the risk nomogram model.

most areas of the decision curve of the prediction model were greater than 0 and the model has a net benefit between about 17 and 65% of threshold probabilities, which indicates that the model has clinical value. All these results demonstrated that our model had a certain predictive value.

DISCUSSION

In this study, we developed a simple but certain discriminating, well-calibrated prediction model for MAP. This model was developed based on DHI information of the dairy cows and focusing on the changes of the dairy cow production performance and milk components. To estimate the risk of MAP, the final model showed a good discrimination and calibration performance. The ROC was 0.71 and the Hosmer–Lemeshow test showed this model has a good fitting trend, therefore no significant difference was observed between the predicted and the actual outcomes ($\chi2=3.26$, $p = 0.92$).

SCC per mL of milk was widely used as an indicator of the incidence of mastitis (Qanbari *et al.*, 2014). Usually, SCC = 500 × 10^3cells/mL was used as a

threshold for judging the clinical mastitis status of dairy cows (Alhussien *et al.*, 2021). However, a strong association between SCC and MAP antibodies has been found in UK Holstein-Friesian cows (Pritchard *et al.*, 2017) and Danish cows (Baptista *et al.*, 2008). In US dairy cows, animals with higher SCC were more positively relative with MAP (Machado *et al.*, 2018). In this study, we used the SCC value of the sampling month rather than SCC values directly, and the previous month to evaluate the udder status of the dairy cows according to the threshold. Chronic clinical mastitis group was here associated with test results strongly (3.22 95% CI: 1.66-4.78), which was in agreement with previous reports generally (Rossi *et al.*, 2017), However, there has been no research on the paratuberculosis makes the host susceptible to mastitis, or intramammary infection can help initiate paratuberculosis (Barber *et al.*, 2019). SCC-based assessment of dairy cows' mammary gland status may be a useful factor for predicting MAP.

MAP antibody distribution in dairy cows is age-related, young calves are more susceptible to MAP infection, while older animals are more likely to become seropositive(Faruk *et al.*, 2020). In this study, animals > 3.5 years old were at a significantly higher risk of being positive (1.80, 95% CI: 1.16-2.81). Although there are few studies about the age distribution of MAP serum antibodies, similar results have been reported in milk-ELISA studies. One study found that the prevalence of 0.33% for dairy cows was less than 2 years of age and 0.94% for 5 years old (Nielsen *et al.*, 2013). Results from another study suggest that dairy cows which > 4 years old were more likely to be MAP positive, relative to animals which < 4 years old (Machado *et al.*, 2018). Probably because MAP has different pathogenicity at different age stages in dairy cows (Matthews *et al.*, 2021).

In the present study, higher fat content in milk was associated with decreased odds of seropositivity (0.77, 95% CI: 0.59-0.99) and the results were consistent with that previously reported in different studies (Rasmussen *et al.*, 2021; Vidic *et al.*, 2013). However, whether the reduction in the level of milk fat can increase the odds of seropositivity remains controversial. Several investigations have reported some seropositivity of dairy cows with higher milk fat compared with seronegative animals (Johnson *et al.*, 2001; Wiszniewska-Łaszczych *et al.*, 2020). While in another study found that it was normal that the presence of antibodies can cause both a decrease and an increase of the milk composition (Eisenberg *et al.*, 2015).

Health, physiological, genetic, environmental, and other factors are able to affect milk production in dairy cows (Garcia and Shalloo, 2015). Thus, MAP infection, as a disease factor, can affect animal productivity. In the present study, animals yielding > 28.4Kg of milk per day were at a significantly lower risk of being positive (0.62, 95% CI: 0.47–0.81) compared to the baseline category. Our result was consistent with previous studies (Pritchard *et al.*, 2017). However, previous studies that investigated the association between milk production and seropositivity showed different results. Milk yield in seropositive dairy cows is significantly higher than in seronegative dairy cows (Hendrick *et al*., 2005; Johnson *et al.*, 2001). These studies used different populations, different sample sizes, and different diagnostic methods. This may explain the inconsistency in part between different study results.

Milk loss is an estimated value, which is based on the difference between the expected and the actual production (Adriaens *et al.*, 2021). Milk loss is associated with individual dairy cows milk SCC (Chen *et al.*, 2021). This relationship has been used to estimate the milk loss due to subclinical mastitis at the herd level and also increase farmers' awareness of the mastitis. Here, Milk loss > 0.16% per day were at a significantly higher risk of being positive (1.39, 95% CI:1.02-1.89) compared to the baseline category, also clearly revealing MAP causing economic loss due to milk loss from the side. Little work has been reported in the literature on the relationship between MAP and milk loss. From this study, milk loss may be a useful predictor for MAP.

No statistically significant association between birth season and odds of positivity for MAP was found, which was in line with previous in 691 herds (Machado *et al.*, 2018), not at par with the result in 4 Holstein herds and 24 Jersey herds (Zare *et al.*, 2013). This may be related to different geographical locations, different sanitary conditions, and management strategies in the tested populations.

To the best of our knowledge, this is the first report of the nomogram predictive model in MAP. Although the current study has strengthened and contributed to the current study, certain limitations should be noted. Firstly, we constructed the prediction nomogram based on the DHI information of the dairy cows, and the accuracy of the data analyzed relied on the quality and completeness of input DHI information. Secondly, the database did not release the data about gastroenteritis incidence records and treatment, stool score, which are the main clinical signs of MAP. Thirdly, this is a single-center study and our model may not be applied to other centers. Therefore, in future studies, the sample size and experimental time should be extended and the practicability of the nomogram model should be verified by a multi-center study.

Conclusions: In summary, our study provides a powerful predictive tool in the form of a nomogram model, which leverages dairy cow-level risk factors to accurately forecast the risk of Mycobacterium avium subspecies paratuberculosis (MAP) infection. By integrating MAP antibody status data obtained through ELISA testing and comprehensive Dairy Herd Improvement (DHI) data for 1,481 dairy cows in Henan Province, China, we have successfully identified key risk factors and crafted an effective predictive model. Through meticulous evaluation, this model has exhibited not only good discrimination and accuracy but also impeccable calibration, affirming its reliability for clinical application in diagnostic testing. With its capacity to individualize MAP risk assessment using readily available DHI information, encompassing factors such as age, milk production, mammary health status, milk loss and milk fat, our nomogram not only aids in identifying high-risk dairy cows but also provides a framework for the strategic development of intervention plans based on DHI reports.

Acknowledgements: This study was supported by the Foundation of major science and technology projects of Henan Province (191110110600, 202102110108), Key scientific research projects of colleges and universities of Henan Province (22A180022), Key R&D project of Ningxia Hui Autonomous Region, China (2022BBF03024).

Conflicts of interest: The authors declare no conflicts of interest.

Author's contribution: GLW and MCW designed the study. HLX and DQL performed the blood samples and data collection and cleaning. GLW, CYL, YW processed the blood samples and ELISA measurements. GLW and MCW performed the data analysis. GLW wrote the article, MCW provided critical revision. All authors approved the final version.

REFERENCES

Adriaens I, Van Den Brulle I, Geerinckx, *et al*., 2021. Milk losses linked to mastitis treatments at dairy farms with automatic milking systems. Preventive Vet Med 194:105420.

Alhussien MN, Panda BSK and Dang AK, 2021. A comparative study on changes in total and differential milk cell counts, activity, and expression of milk phagocytes of healthy and mastitic indigenous sahiwal cows. Front Vet Sci 8:670811.

Baptista FM, Nielsen SS and Toft N, 2008. Association between the presence of antibodies to Mycobacterium avium subspecies paratuberculosis and somatic cell count. J of Dairy Sci 91(1):109-18.

Barber B, Bell N and Van Winden S, 2019. Mycobacterium avium subspecies seroconversion in dairy cattle and its association with raised somatic cell count. Vet Rec 185(12):374.

Chen H, Weersink A, Kelton D, *et al*., 2021. Estimating milk loss based on somatic cell count at the cow and herd level. J Dairy Sci 104(7):7919-31.

Cheng Z, Liu M, Wang P, *et al*., 2020. Characteristics and epidemiological investigation of Paratuberculosis in dairy cattle in Tai'an, China. BioMed Res Int 2020:3896754.

Crowson CS, Atkinson EJ and Therneau TM, 2016. Assessing calibration of prognostic risk scores. Stat Methods Med Res 25(4):1692-706.

Dufour B, Pouillot R and Durand B, 2004. A cost/benefit study of paratuberculosis certification in French cattle herds. Vet Res 35(1):69-81.

Eisenberg S, Veldman E, Rutten V, *et al*., 2015. A longitudinal study of factors influencing the result of a Mycobacterium avium ssp. Paratuberculosis Antibody ELISA in Milk of Dairy Cows 98(4):2345-55.

Faruk M, Jung Y, Hur T, *et al*., 2020. Mycobacterium avium longitudinal study of subsp. antibody kinetics in dairy cattle using sera and milk throughout the lactation period. Vet Sci 7(3):81.

Garcia A and Shalloo L, 2015. Invited review:The economic impact and control of paratuberculosis in cattle. J Dairy Sci 98(8):5019-39.

He J, Sun J, Zhang T, *et al*., 2021. Construction and validation of risk prediction model for deep vein thrombosis in acute exacerbations of chronic obstructive pulmonary disease based on serum angiopoietin 2 levels. Annals of Palliative Med 10(7):7905-18.

Hendrick S, Duffield T, Kelton D, *et al*., 2005. Evaluation of enzyme-linked immunosorbent assays performed on milk and serum samples for detection of paratuberculosis in lactating dairy cows. J of the Amer Vet Med Assoc 226(3):424-8.

Hendrick S, Kelton D, Leslie K, *et al*., 2005. Effect of paratuberculosis on culling, milk production, and milk quality in dairy herds. J of the Amer Vet Med Assoc 227(8):1302-8.

Johnson Y, Kaneene J, Gardiner J, *et al*., 2001. The effect of subclinical Mycobacterium paratuberculosis infection on milk production in Michigan dairy cows. J Dairy Sci 84(10):2188-94.

Li G, Bing Y, Tian M, *et al*., 2021. Using a nomogram to preoperatively predict distant metastasis of pancreatic neuroendocrine tumor in elderly patients. Chin Med Sci J 36(3):218-24.

Ma J, Tian A, Zheng W, *et al*., 2019. First report of bovine viral diarrhea virus and Mycobacterium avium subspecies paratuberculosis infection in Tibetan sheep (Ovis aries) in Tibetan Plateau, China. Trop Anim Health Prod 51(3):719-22.

Machado G, Kanankege K, Schumann V, *et al*., 2018. Identifying individual animal factors associated with Mycobacterium avium subsp. paratuberculosis (MAP) milk ELISA positivity in dairy cattle in the Midwest region of the United States. BMC Vet Res 14(1):28.

Matthews C, Cotter P, O' Mahony, *et al*., 2021. MAP, Johne's disease and the microbiome; current knowledge and future considerations. Anim Microbiome 3(1):34.

Milovanović M, Milićević V, Radojičić S, *et al*., 2020. Suitability of individual and bulk milk samples to investigate the humoral immune response to lumpy skin disease vaccination by ELISA. Virol J 17(1):28.

Nielsen S, 2008. Transitions in diagnostic tests used for detection of Mycobacterium avium subsp. paratuberculosis infections in cattle. Vet Microbiol 132(3-4):274-82.

Nielsen S and Toft N, 2009. A review of prevalences of paratuber-culosis in farmed animals in Europe. Preventive Vet Med 88:1-14.

Nielsen S, Toft N and Okura H, 2013. Dynamics of specific anti-Mycobacterium avium subsp. paratuberculosis antibody response through age. PloS One 8(4):e63009.

Ozsvari L, Harnos A, Lang Z, *et al*., 2020. The impact of paratuberculosis on milk production, fertility, and culling in large commercial hungarian dairy herds. Frontiers in Vet Sci 7:565324.

Pierce E, 2018. Mycobacterium avium could subspecies cause Crohn's disease, ulcerative colitis and colorectal cancer? Infectious agents and cancer. Infect Agent Cancer 13:1.

Pritchard T, Coffey M, Bond K, *et al*., 2017. Phenotypic effects of subclinical paratuberculosis (Johne's disease) in dairy cattle. J Dairy Sci 100(1):679-90.

Qanbari S, Pausch H, Jansen S, *et al*., 2014. Classic selective sweeps revealed by massive sequencing in cattle. PLoS Genet 10(2):e1004148.

Rasmussen P, Barkema H, Mason S, *et al*., 2021. Economic losses due to Johne's disease (paratuberculosis) in dairy cattle. J Dairy Sci 104(3):3123-43.

Rathnaiah G, Zinniel D, Bannantine J, *et al*., 2017. Mycobacterium avium pathogenesis, molecular genetics, and genomics of subsp. , the etiologic agent of Johne's disease. Front Vet Sci 4:187.

Raven L, Cocks B and Hayes B, 2014. Multibreed genome wide association can improve precision of mapping causative variants underlying milk production in dairy cattle. BMC Genom 15:62.

Robin X, Turck N, Hainard A, *et al*., 2011. pROC:an open-source package for R and S^+ to analyze and compare ROC curves. BMC Bioinformatics 12:77.

Rossi G, Grohn YT, Schukken YH, *et al*., 2017. The effect of Mycobacterium avium ssp. paratuberculosis infection on clinical mastitis occurrence in dairy cows. J Dairy Sci 100(9):7446-54.

Shao J, Wang C, Shu K, *et al*., 2023. A contrast-enhanced CT-based radiomic nomogram for the differential diagnosis of intravenous leiomyomatosis and uterine leiomyoma. Front Oncol 13:1239124.

Shi X, Xu L, Ma B, *et al*., 2020. Development and validation of a nomogram to predict the prognosis of patients with gastric cardia cancer. Sci Rep 10(1):14143.

Song W and Fu T, 2019. Circular RNA-Associated Competing Endogenous RNA Network and Prognostic Nomogram for Patients With Colorectal Cancer. Front Oncol 9:1181.

Vidic B, Savic S, Vidic V, *et al*., 2013. Economic impact of paratuberculosis on milk production. Biotechnol Anim Husbandry 29(2):183-91.

Vilar A, Santos C, Pimenta C, *et al*., 2015. Herd-level prevalence and associated risk factors for Mycobacterium avium subsp. paratuberculosis in cattle in the State of Paraíba, Northeastern Brazil. Preventive Vet Med 121(1-2):49-55.

Wiszniewska-Łaszczych A, Liedtke K, Szteyn J, *et al*., 2020. The effect of *Mycobacterium avium subsp. Paratuberculosis* infection on the productivity of cows in two dairy herds with a low seroprevalence of Paratuberculosis. Animals(Basel) 10(3):490.

Zare Y, Shook G, Collins M, *et al*., 2013. Evidence of birth seasonality and clustering of Mycobacterium avium subspecies paratuberculosis infection in US dairy herds. Preventive Vet Med 112(3-4):276-84.