

## OPEN ACCESS

## EDITED BY

Markus Rupp,  
University Medical Center Regensburg,  
Germany

## REVIEWED BY

Vinayak Narayan,  
Northwell Health, United States  
Piotr Yablonskii,  
St-Petersburg Research Institute of  
Phthisiopulmonology, Russia  
Ning Lang,  
Peking University Third Hospital, China

## \*CORRESPONDENCE

Mardan Mamat  
mardanmmtmx@163.com

## SPECIALTY SECTION

This article was submitted to Orthopedic  
Surgery, a section of the journal *Frontiers in  
Surgery*

RECEIVED 29 May 2022

ACCEPTED 02 November 2022

PUBLISHED 06 January 2023

## CITATION

Yasin P, Mardan M, Xu T, Cai X, Abulizi Y,  
Wang T, Sheng W and Mamat M (2023)  
Development and validation of a diagnostic  
model for differentiating tuberculous  
spondylitis from brucellar spondylitis using  
machine learning: A retrospective cohort study.  
*Front. Surg.* 9:955761.  
doi: 10.3389/fsurg.2022.955761

## COPYRIGHT

© 2023 Yasin, Mardan, Xu, Cai, Abulizi, Wang,  
Sheng and Mamat. This is an open-access  
article distributed under the terms of the  
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).  
The use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in this  
journal is cited, in accordance with accepted  
academic practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# Development and validation of a diagnostic model for differentiating tuberculous spondylitis from brucellar spondylitis using machine learning: A retrospective cohort study

Parhat Yasin<sup>1</sup>, Muradil Mardan<sup>2</sup>, Tao Xu<sup>1</sup>, Xiaoyu Cai<sup>1</sup>,  
Yakefu Abulizi<sup>1</sup>, Ting Wang<sup>1</sup>, Weibin Sheng<sup>1</sup>  
and Mardan Mamat<sup>1\*</sup>

<sup>1</sup>Department of Spine Surgery, The First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang, China, <sup>2</sup>School of Medicine, Tongji University, Shanghai, China

**Background:** Tuberculous spondylitis (TS) and brucellar spondylitis (BS) are commonly observed in spinal infectious diseases, which are initially caused by bacteremia. BS is easily misdiagnosed as TS, especially in underdeveloped regions of northwestern China with less sensitive medical equipment. Nevertheless, a rapid and reliable diagnostic tool remains to be developed and a clinical diagnostic model to differentiate TS and BS using machine learning algorithms is of great significance.

**Methods:** A total of 410 patients were included in this study. Independent factors to predict TS were selected by using the least absolute shrinkage and selection operator (LASSO) regression model, permutation feature importance, and multivariate logistic regression analysis. A TS risk prediction model was developed with six different machine learning algorithms. We used several metrics to evaluate the accuracy, calibration capability, and predictability of these models. The performance of the model with the best predictability was further verified with the area under the curve (AUC) of the receiver operating characteristic (ROC) curve and the calibration curve. The clinical performance of the final model was evaluated by decision curve analysis.

**Results:** Six variables were incorporated in the final model, namely, pain severity, CRP, x-ray intervertebral disc height loss, x-ray endplate sclerosis, CT vertebral destruction, and MRI paravertebral abscess. The analysis of appraising six models revealed that the logistic regression model developed in the current study outperformed other methods in terms of sensitivity ( $0.88 \pm 0.07$ ) and accuracy ( $0.79 \pm 0.07$ ). The AUC of the logistic regression model predicting TS was 0.86 (95% CI, 0.81–0.90) in the training set and 0.86 (95% CI, 0.78–0.92) in the validation set. The decision curve analysis indicated that the logistic regression model displayed a higher clinical efficiency in the differential diagnosis.

**Conclusions:** The logistic regression model developed in this study outperformed other methods. The logistic regression model demonstrated by a calculator exerts good discrimination and calibration capability and could be applicable in differentiating TS from BS in primary health care diagnosis.

#### KEYWORDS

tuberculous spondylitis (TS), brucellar spondylitis (BS), magnetic resonance imaging (MRI), computed tomography (CT), x-ray, machine learning

## Introduction

Tuberculosis (TB) and brucellosis are severe infectious diseases that are threatening human beings. According to the global tuberculosis report (2014), TB remains one of the world's deadliest communicable diseases, and in 2013, approximately 9.0 million people developed TB, among which 1.5 million died from the disease (1), and another recent report showed that 1.6 million people died from TB in 2017 (2). Brucellosis, which is caused by *Brucella melitensis*, is a serious zoonotic disease that causes more than 500,000 human infections worldwide annually (3). Spinal tuberculosis (STB) is not a rare presentation of extrapulmonary tuberculosis. About 1%–2% of all cases of TB are diagnosed as STB, and these patients represent 10%–15% of extrapulmonary TB, of which nearly half involve the musculoskeletal system (4). About 6%–12% of brucellosis cases may suffer a spinal illness, which is the latent reason for the deformities and permanent neurologic deficiencies (5–8). TS and BS are commonly observed in spinal infectious diseases, which are initially caused by bacteremia. They mostly occur in the thoracolumbar segment of the spine. Both TS and BS present several similar clinical performances, such as low-grade fever, including dull pain or discomfort of the dorsum, and elevated inflammatory mediators; hence, distinguishing TS from BS is challenging and BS is commonly misdiagnosed as TS. Currently, the most effective and accurate method for distinguishing TS from BS is based on biopsy and the isolation, culture, and identification of mycobacteria from patient specimens, but it is laborious and time-consuming (9). Hence, developing rapid, cost-effective, and accurate diagnostic methods is urgently desired and of great clinical significance. In this study, we report the development and validation of a machine learning algorithm-based diagnostic model to differentiate between the acute and subacute stages: TS and BS. The predictive model presented in this article follows the TRIPOD Checklist (10).

## Materials and methods

The research was conducted under the approval of the ethics committee of Xinjiang Medical University Affiliated

First Hospital, Urumqi, and individual agreements for this retrospective analysis were waived.

## Patients

Patients admitted to the Department of Spine Surgery between January 2018 and December 2021 and considered as spinal TS ( $n = 275$ , primary cohort: 612) or BS ( $n = 135$ , primary cohort: 209) (Table 1) were included in this population-based retrospective cohort study with ethical approval of the ethical review committee board of Xinjiang Medical University Affiliated First Hospital. Patients included in this study met the following criteria: (1) diagnosed with spinal tuberculosis or brucellar spondylitis in the acute and subacute stages; (2) accepted surgery therapy; (3) the collected information, especially imaging materials, was complete and available; and (4) age  $\geq 18$  years. Patients who met the following exclusion criteria were excluded from analysis: (1) diagnosed with malignant cancer, hematological diseases, and hepatology disease; (2) spine out of alignment; (3) revision spinal surgery; (4) scoliosis deformity; (5) pyogenic spondylitis; (6) spinal hydatid; (7) age  $< 18$  years; and (8) patients with missing data were  $\geq 10\%$ .

The diagnosis, referred to as a response variable in our research, was obtained from symptoms, signs, laboratory tests, and imaging features. TS and BS share similar clinical presentation along with the systemic constitutional manifestation, characterized by sweating, fever, local pain, fatigue, etc. Imaging revealed mild or severe vertebral destruction, intervertebral disc height loss, cold abscess, etc. Laboratory tests included erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), and routine blood tests, which are considered nonspecific. Specific tests comprised positive results of enzyme-linked immunospot assay (T-SPOT.TB), the presence of *Mycobacterium tuberculosis* based on acid-fast bacilli in Ziehl-Neelsen-stained smears, growth in cultures, and/or biopsy examination for TS and *Brucella* agglutination titer test (1:160 or higher) and isolation of *Brucella* species from blood, bone marrow, or other tissues for BS.

TABLE 1 Baseline characteristics of patients.

Variables	Total (N = 410)	BS (N = 135)	TS (N = 275)	p- Value
Age (years)	51.6 ± 16.1	51.8 ± 12.2	51.4 ± 17.7	0.780
Gender				<b>&lt;0.001</b>
Female	166 (40.5%)	36 (26.7%)	130 (47.3%)	
Male	244 (59.5%)	99 (73.3%)	145 (52.7%)	
Ethnicity				0.181
Han	126 (30.7%)	52 (38.5%)	74 (26.9%)	
Kazak	33 (8.05%)	10 (7.41%)	23 (8.36%)	
Mongolian	7 (1.71%)	1 (0.74%)	6 (2.18%)	
Others	23 (5.61%)	6 (4.44%)	17 (6.18%)	
Uygur	221 (53.9%)	66 (48.9%)	155 (56.4%)	
BMI (kg/m <sup>2</sup> )	23.1 ± 3.03	23.61 ± 2.86	22.8 ± 3.09	0.011
Fever				<b>&lt;0.001</b>
High	73 (17.8%)	43 (31.9%)	30 (10.9%)	
Low	337 (82.2%)	92 (68.1%)	245 (89.1%)	
Pain severity				<b>&lt;0.001</b>
Moderate	192 (46.8%)	41 (30.4%)	151 (54.9%)	
Severe	218 (53.2%)	94 (69.6%)	124 (45.1%)	
History of weight loss				0.483
No	259 (63.2%)	89 (65.9%)	170 (61.8%)	
Yes	151 (36.8%)	46 (34.1%)	105 (38.2%)	
Past history of tuberculosis in other solid organs				0.181
No	324 (79.0%)	101 (74.8%)	223 (81.1%)	
Yes	86 (21.0%)	34 (25.2%)	52 (18.9%)	
WBC (×10 <sup>9</sup> /L)	6.56 ± 2.15	6.62 ± 2.10	6.53 ± 2.17	0.667
ESR (mm/h)	45.60 ± 17.61	44.4 ± 15.4	46.1 ± 18.6	0.327
CRP (mg/L)	44.30 ± 37.00	30.9 ± 23.31	50.8 ± 40.5	<b>&lt;0.001</b>
Hb (g/L)	126 ± 17.7	130 ± 16.6	124 ± 18.0	<b>0.004</b>
TG (mmol/L)	1.25 ± 0.54	1.37 ± 0.62	1.19 ± 0.49	<b>0.003</b>
TC (mmol/L)	3.95 ± 0.92	4.15 ± 0.86	3.85 ± 0.93	<b>0.002</b>
HDL-C (mmol/L)	1.00 ± 0.31	0.97 ± 0.30	1.02 ± 0.31	0.155
LDL-C (mmol/L)	2.69 ± 0.77	2.79 ± 0.69	2.65 ± 0.81	0.058
ALB (g/L)	37.8 ± 5.71	37.1 ± 6.08	38.1 ± 5.49	0.093
AST (U/L)	24.0 ± 17.5	27.0 ± 17.0	22.5 ± 17.5	0.012
ALT (U/L)	25.9 ± 27.6	34.3 ± 29.0	21.8 ± 26.0	<b>&lt;0.001</b>
GGT (U/L)	51.9 ± 45.4	55.8 ± 43.1	50.0 ± 46.5	0.215
ALP (U/L)	103 ± 42.8	108 ± 39.7	101 ± 44.0	0.072
Location				
C	10 (2.44%)	7 (5.19%)	3 (1.09%)	
C + T	2 (0.49%)	1 (0.74%)	1 (0.36%)	
L	233 (56.8%)	90 (66.7%)	143 (52.0%)	
L + S	52 (12.7%)	27 (20.0%)	25 (9.09%)	
S	1 (0.24%)	0 (0.00%)	1 (0.36%)	
T	90 (22.0%)	6 (4.44%)	84 (30.5%)	
T + L	22 (5.37%)	4 (2.96%)	18 (6.55%)	
Segment	2.48 ± 0.96	2.29 ± 0.66	2.57 ± 1.07	<b>0.001</b>

(continued)

TABLE 1 Continued

Variables	Total (N = 410)	BS (N = 135)	TS (N = 275)	p- Value
MRI spinal stenosis				<b>0.001</b>
No	273 (66.6%)	106 (78.5%)	167 (60.7%)	
Yes	137 (33.4%)	29 (21.5%)	108 (39.3%)	
MRI paravertebral abscess				<b>&lt;0.001</b>
No	169 (41.2%)	85 (63.0%)	84 (30.5%)	
Yes	241 (58.8%)	50 (37.0%)	191 (69.5%)	
MRI psoas abscess				0.116
No	259 (63.2%)	93 (68.9%)	166 (60.4%)	
Yes	151 (36.8%)	42 (31.1%)	109 (39.6%)	
MRI epidural abscess				0.973
No	320 (78.0%)	106 (78.5%)	214 (77.8%)	
Yes	90 (22.0%)	29 (21.5%)	61 (22.2%)	
CT vertebral destruction				<b>&lt;0.001</b>
Mild (≤1/3)	129 (31.5%)	68 (50.4%)	61 (22.2%)	
Severe (>1/3)	281 (68.5%)	67 (49.6%)	214 (77.8%)	
CT marginal osteophytes				0.429
No	186 (45.4%)	57 (42.2%)	129 (46.9%)	
Yes	224 (54.6%)	78 (57.8%)	146 (53.1%)	
CT endplate sclerosis				0.022
No	267 (65.1%)	77 (57.0%)	190 (69.1%)	
Yes	143 (34.9%)	58 (43.0%)	85 (30.9%)	
CT spinal stenosis				<b>0.005</b>
No	317 (77.3%)	116 (85.9%)	201 (73.1%)	
Yes	93 (22.7%)	19 (14.1%)	74 (26.9%)	
CT paravertebral abscess				<b>0.001</b>
No	198 (48.3%)	81 (60.0%)	117 (42.5%)	
Yes	212 (51.7%)	54 (40.0%)	158 (57.5%)	
CT epidural abscess				0.737
No	366 (89.3%)	122 (90.4%)	244 (88.7%)	
Yes	44 (10.7%)	13 (9.63%)	31 (11.3%)	
X-ray intervertebral disc height loss				<b>&lt;0.001</b>
No	166 (40.5%)	84 (62.2%)	82 (29.8%)	
Yes	244 (59.5%)	51 (37.8%)	193 (70.2%)	
X-ray endplate sclerosis				<b>&lt;0.001</b>
No	248 (60.5%)	53 (39.3%)	195 (70.9%)	
Yes	162 (39.5%)	82 (60.7%)	80 (29.1%)	
X-ray osteophytes				<b>&lt;0.001</b>
No	251 (61.2%)	63 (46.7%)	188 (68.4%)	
Yes	159 (38.8%)	72 (53.3%)	87 (31.6%)	

BMI, body mass index (kg/m<sup>2</sup>); WBC, preoperative white blood cell (×10<sup>9</sup>/L); ESR, preoperative erythrocyte sedimentation rate (mm/h); CRP, preoperative C-reactive protein (mg/L); Hb, preoperative hemoglobin (g/L); TG, preoperative total triglyceride (mmol/L); TC, preoperative total cholesterol (mmol/L); HD-C, preoperative high-density lipoprotein cholesterol (mmol/L); LDL-C, preoperative low-density lipoprotein cholesterol (mmol/L); ALB, preoperative operative albumin (g/L); AST, preoperative aspartate aminotransferase (U/L); ALT, preoperative alanine aminotransferase (U/L); GGT, preoperative gamma-glutamyl transferase (U/L); ALP, alkaline phosphatase U/L; C, cervical spine; T, thoracic spine; T + L, thoracolumbar spine; L, lumbar spine; L + S, lumbosacral spine.

## Collection of data

Demographic, clinical, and imaging data were collected for each case, including age, gender, location that can be used to estimate the disease epidemiology characteristic (map source: [http://datav.aliyun.com/portal/school/atlas/area\\_selector](http://datav.aliyun.com/portal/school/atlas/area_selector)) (as is shown in [Figure 1](#)), the body mass index (BMI), the level of pain degree divided into two categories based on the visual analog scale (moderate,  $VAS \leq 5$ ; severe,  $VAS > 5$ ), the fever grade measured at the patient's first visit also divided into two categories (low,  $<38.5^\circ\text{C}$ ; high,  $\geq 38.5^\circ\text{C}$ ), preoperative ESR, preoperative CRP, preoperative white blood cell (WBC) count, preoperative hemoglobin, history of weight loss, history of tuberculosis in other solid organs, preoperative low-density lipoprotein cholesterol (LDL-C), preoperative high-density lipoprotein cholesterol (HDL-C), preoperative total cholesterol (TC), preoperative total triglyceride (TG), preoperative albumin (Alb), preoperative gamma-glutamyl transferase (GGT), preoperative alanine aminotransferase (ALT), preoperative aspartate aminotransferase (AST), preoperative alkaline phosphatase (ALP), the level of involvement, the number of affected vertebra, magnetic resonance imaging (MRI) findings including abscess (paravertebral abscess, epidural abscess, psoas abscess) and spinal stenosis, and computed tomography (CT) findings including vertebral destruction, marginal osteophytes, endplate sclerosis, spinal stenosis, paravertebral abscess, and epidural abscess. We defined severe vertebral destruction as one-third or higher vertebral damage. X-ray findings included intervertebral disc height, osteophytes, endplate sclerosis, and bone bridge. All images used in this study were reviewed

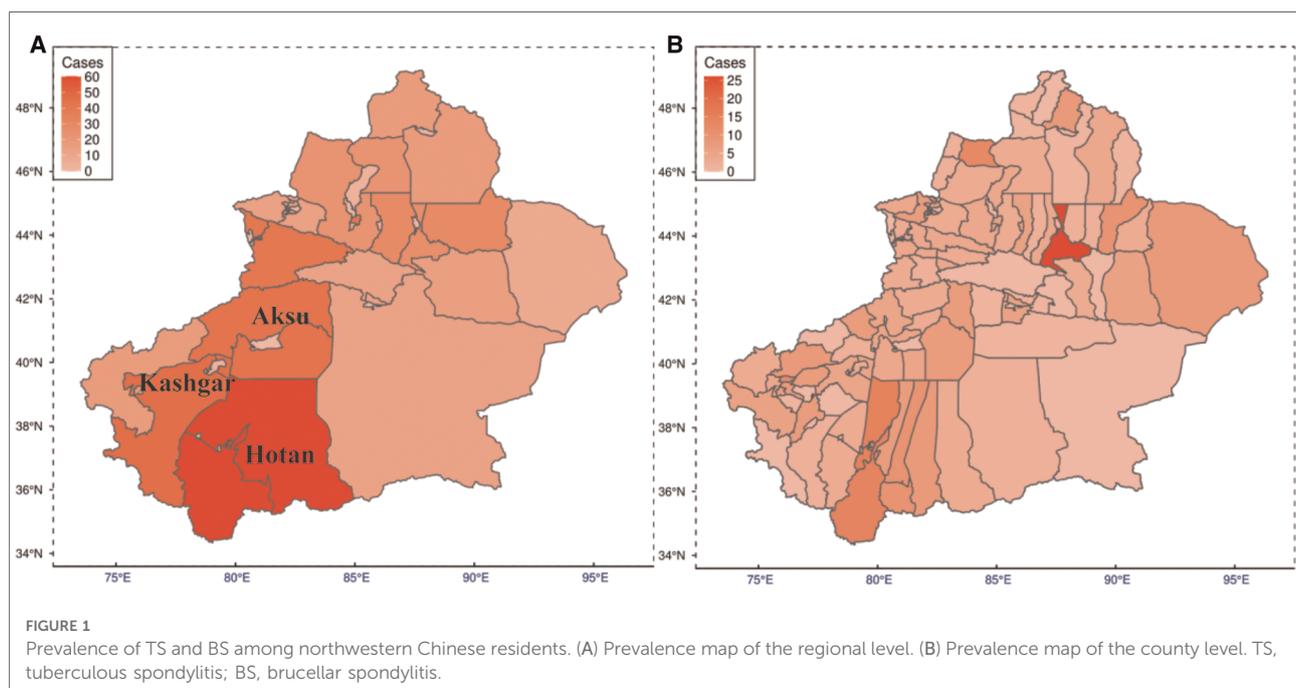
and analyzed by a chief physician blinded to clinical and laboratory results. We imputed the missing data ( $<10\%$ ) using the MICE package (version 3.14.0) (11).

## Feature selection

We identified candidate predictors through the least absolute shrinkage and selection operator (LASSO) model owing to its attribution of compression estimation algorithms in high-dimensional regression and the importance score of each predictor via the permutation importance approach using the random forest classification model. After applying the LASSO regression model and permutation feature importance method to the training set, respectively, we initially screened variables (12). We chose the top 10 variables according to their importance arranged by the model, which simultaneously were selected in the LASSO method. Then, a multivariable logistic regression analysis was conducted. Variables with a two-sided  $p$ -value  $\leq 0.05$  and frequently used in routine clinical practice were included in the model along with their odds ratios (ORs), associated 95% confidence intervals (CI),  $\beta$ -coefficients, and corresponding  $p$ -values.

## Machine learning model construction

Regarding machine learning, we used six risk algorithms to develop a predictive model for TS: logistic regression (LR),



neural network (NN) (13), random forest (RF) (14), decision tree (DT) (15), Gaussian naïve Bayes (Gaussian NB) (16), and K-nearest neighbor (KNN) (17). LR is basically a classification algorithm that comes under the supervised category. DT is a nonparametric supervised learning algorithm consisting of upside-down trees that make decisions based on the conditions present in the data. RF is a combination of a multitude of decision trees that can be constructed for prediction when facing regression tasks. NN is one of the supervised machine learning methods that simulates the way the human brain processes information. NB is a method based on Bayes theorem mainly used for classification. KNN is a nonparametric classification approach widely used in real-life issues (18–22).

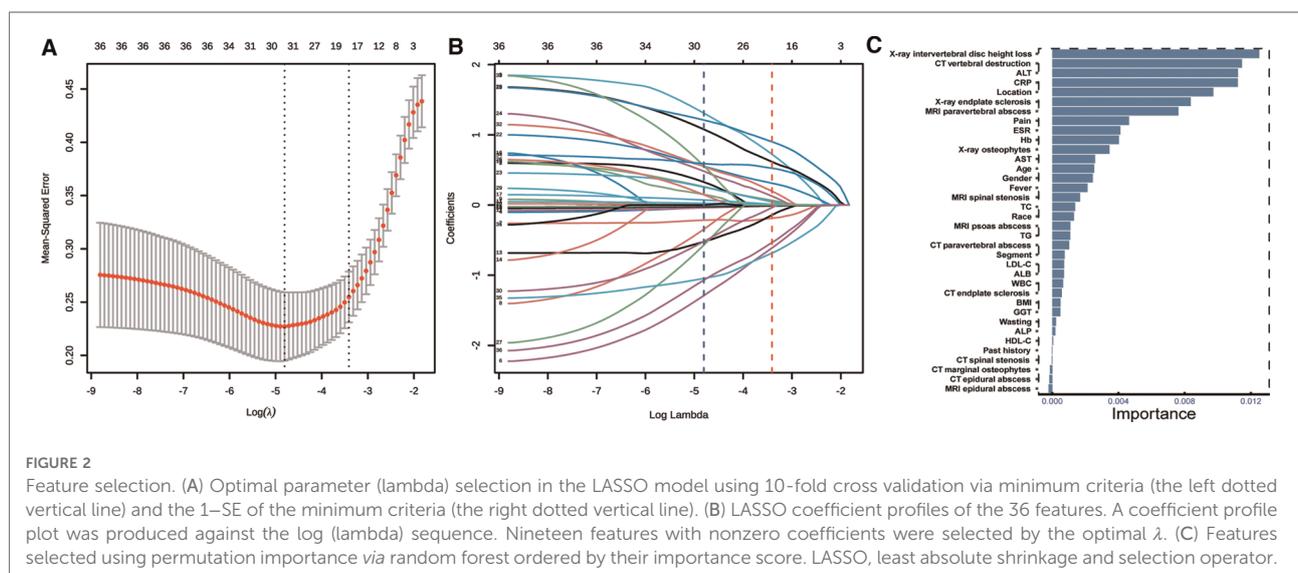
Once the features were inputted, these algorithms enabled predictions regarding important signs for the diagnosis of TS in a sample of patients with TS or BS. R programming software (version 4.1.2) was used to build the predictive models.

### Evaluation and improvement of model performance

The data used in this study were randomly divided into two groups including a training set and a validation set with a ratio of 7:3. Model establishment consists of some unavoidable processes: data preprocessing, training the model with tuned hyperparameters (also called model performance improvement), evaluating the model performance, and testing the model on unknown data. However, previous research studies present an error-prone manipulation, which is reporting the performance

estimated in the tuning procedure as model performance, which is somehow biased and overestimated (23). Evaluating the model performance should not be carried on the same datasets used for tuning since this kind of operation would cause biased performance during evaluation. Thus, we adopted a nested resampling strategy (nested cross-validation) to obtain an unbiased score. It used outer and inner loops to separate resampling optimization from model performance evaluation. The model was fitted on the outer training data set using the tuned hyperparameter configuration obtained by inner resampling. Repeated *k*-fold cross-validation (KCV, *k* = 10, *n* = 10, *n* is the number of repeats) was used as the outer resampling strategy, and *k*-fold cross-validation (KCV, *k* = 5) was the inner resampling method to tune the hyperparameters of each model. In the process of KCV, *k* – 1 folds of the data were used as the training set and the reserved part of data was used as the testing set to evaluate nine metrics, namely, sensitivity, specificity, accuracy, precision, positive predictive value (PPV), negative predictive value (NPV), F1 score, area under the curve of the receiver operating characteristic curve (AUROC), and the precision–recall curve (AUPRC) iteratively until every fold experienced inner validation. The whole process was repeated 100 times. This was believed to reduce the probability of overfitting and underfitting in a tiny data set and would help to reflect its practical performance.

Ultimately, the values of AUROC and AUPRC from the six models were compared to decide the best performing model. The opted model, logistic regression (LR), was constructed as a scoring system using the entire training data, and it was validated using the validation data set. The ROC and PRC



analyses were carried out utilizing the R package: ModelMetrics (version 1.2.2.2) (24).

## Scoring system development and validation

The logistic regression model, selected after the aforementioned individual models were evaluated based on the required criteria, is displayed as a scale system embedded into Excel (Microsoft, USA), which is convenient to use (25). We estimate the discrimination performance of the scale system with AUROC and the calibration curve in the training and validation sets, respectively. At last, decision curve analysis (DCA) was used to examine the clinical efficiency of the model to quantify the benefits and the area under the curve to be appraised (26).

## Statistical analysis

We performed all statistical analyses by using R software 4.1.2. The normality of the data with the Q-Q plots of all data was assessed. Continuous variables were presented as mean ± standard deviation (SD) in the case of normal distribution; otherwise, they were presented as median values (quartiles). Student's *t*-test was used to compare two mean values of continuous data considered normally distributed after normality evaluation. Otherwise, the Mann-Whitney *U*-test was performed. Categorical variables were expressed as frequency (percentage). The chi-square test or Fisher's exact test was used to compare two frequencies.

## Results

### Epidemiology of cases enrolled in this study

Regional distributions of patients diagnosed with TS or BS enrolled in this study are shown in Figure 1. For each region, the darker shade represents a higher incidence of disease. As can be seen, in general, the southern part of Xinjiang China, especially the Hotan region, reveals a higher prevalence.

### Patients

A total of 410 patients (*n* = 275 TS patients and *n* = 135 BS patients) were enrolled; 70% of them were included in the training set (*n* = 292), and the remaining patients were included in the validation set (*n* = 118). The differences in all baseline demographic characteristics and predictors, including clinical personation, laboratory tests, and radiology findings between the TS and BS, are given in Table 1. Patients with TS had higher CRP

levels, ESR, and proportion of lower pain, while patients with BS showed higher WBC count. In addition, most imaging-related data showed significant differences between patients with TS and BS.

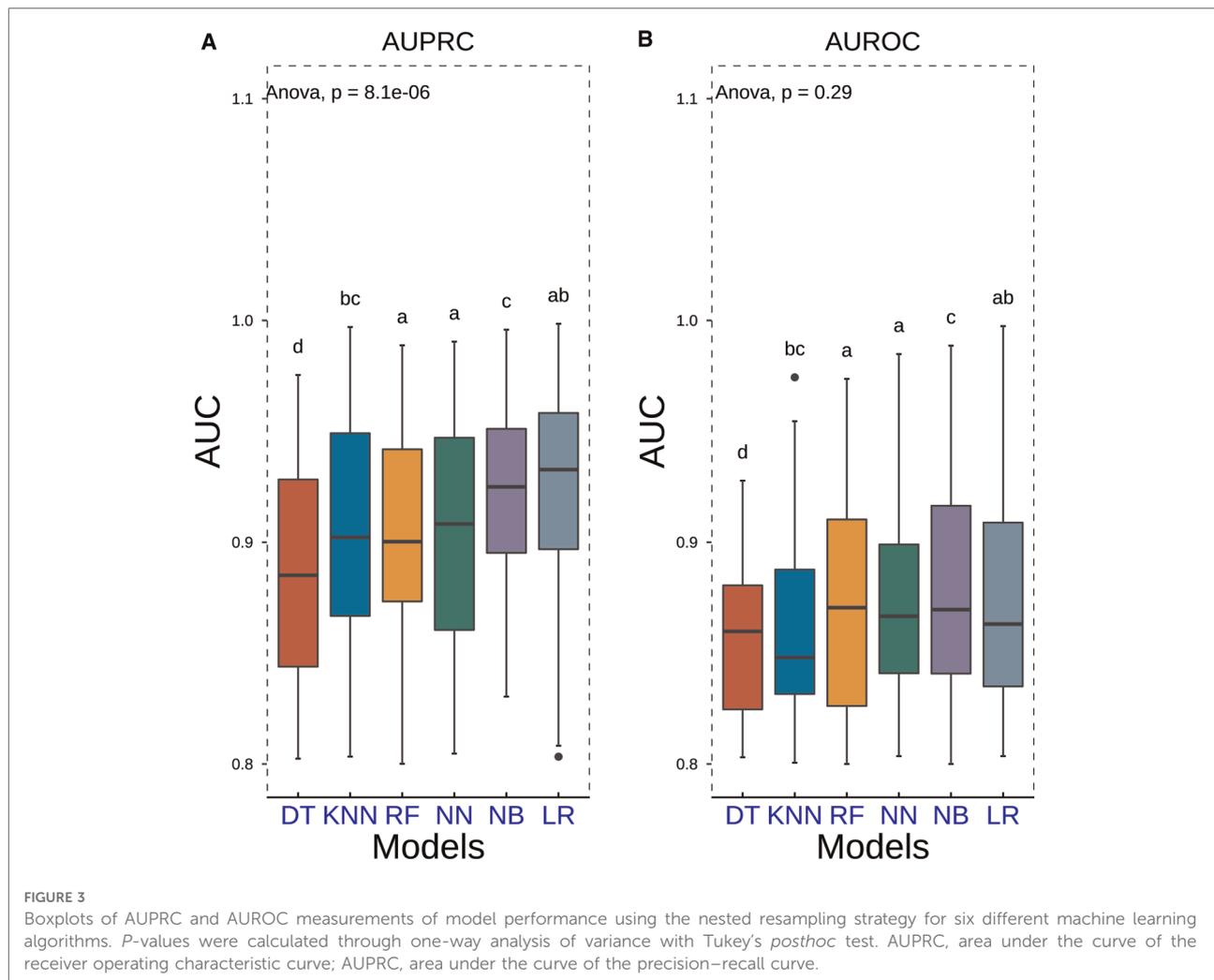
## Feature selection

Thirty-six variables were reduced to 19 predictors with the LASSO method (Figures 2A,B). The top 10 variables with relative importance score selected by the LASSO method were CRP, ESR, Hb, ALT, pain severity, CT vertebral destruction, x-ray intervertebral disc height loss, x-ray endplate sclerosis, MRI paravertebral abscess, and location (Figure 2C). Multivariate analysis was conducted based on the above results. Predictors associated with the TS patients included pain severity, CRP, x-ray intervertebral disc height loss, x-ray endplate sclerosis, CT vertebral destruction, and MRI paravertebral abscess (Table 2).

TABLE 2 Prediction factors for TS from study population by multiple logistic regression model.

Characteristic	OR	95% CI	<i>p</i> -Value
Pain severity			
Moderate	—	—	
Severe	0.37	0.20, 0.66	<0.001
CRP (mg/L)	1.02	1.01, 1.03	<0.001
Hb (g/L)	0.97	0.95, 0.99	0.008
ALT (U/L)	0.99	0.98, 1.00	0.045
ESR (mm/h)	0.99	0.97, 1.01	0.2
X-ray endplate sclerosis			
No	—	—	
Yes	0.20	0.11, 0.36	<0.001
X-ray intervertebral disc height loss			
No	—	—	
Yes	3.31	1.87, 5.98	<0.001
CT vertebral destruction			
Mild (≤ 1/3)	—	—	
Severe (>1/3)	3.21	1.78, 5.87	<0.001
MRI paravertebral abscess			
No	—	—	
Yes	3.05	1.72, 5.51	<0.001
Location			
C	—	—	
C + T	1.92	0.04, 102	0.7
L	4.46	0.89, 26.9	0.079
L + S	3.95	0.69, 26.8	0.13
T	38.6	6.15, 292	<0.001
T + L	16.5	1.93, 171	0.013

OR, odds ratio; CI, confidence interval; CRP, preoperative C-reactive protein (mg/L); Hb, preoperative hemoglobin (g/L); ESR, preoperative erythrocyte sedimentation rate (mm/h); ALT, preoperative alanine aminotransferase (U/L); C, cervical spine; T, thoracic spine; T + L, thoracolumbar spine; L, lumbar spine; L + S, lumbosacral spine.



## Evaluation of model prediction capability

Repeated 10-fold cross-validation was carried out in the outer loop to assess model performance with ROC and PRC analyses. This process was repeated 10 times. We discovered that DT was related to relatively lower AUROC and AUPRC values. However, LR, NN, and NB methods exhibited higher AUROC and AUPRC values (Figure 3). Furthermore, seven popular metrics (sensitivity, specificity, accuracy, precision, F1 score, PPV, and NPV) were also used to assess the performance of these models (Table 3). As LR shows higher specificity than NN and NB and has best accuracy and F1 score, it is the most commonly used algorithm with its convenience displaying high accuracy with lower standard deviance. This indicated that the LR model did possess an outstanding ability to be implemented into clinical decision-making.

## Establishment of the scoring system

Based on the candidate predictors screened on the training set, a scale calculator, which comprised six major features, was developed for predicting the probability of TS. Each factor in the calculator was assigned a unique score in light of the value of the corresponding factor. The sum of all scores computed by rounding up the scores of all predictors can be used to compute the probability of TS (Figure 4). For details, please refer to Table S1.

## Model performance and validation

We validated the differentiation capacity of the model in the training set and validation set, respectively. The C-statistics and AUC of the model to predict the diagnosis of TS were 0.860

TABLE 3 Predictive performance of each model.

Model	Sen.	Spe.	Acc.	Pre.	P.P.V.	N.P.V.	F1
LR	0.88 ± 0.07	0.58 ± 0.18	0.79 ± 0.07	0.82 ± 0.08	0.82 ± 0.08	0.69 ± 0.19	0.85 ± 0.06
NN	0.87 ± 0.08	0.56 ± 0.2	0.77 ± 0.07	0.81 ± 0.09	0.81 ± 0.09	0.68 ± 0.17	0.84 ± 0.06
RF	0.89 ± 0.08	0.56 ± 0.17	0.79 ± 0.08	0.82 ± 0.08	0.82 ± 0.08	0.71 ± 0.18	0.85 ± 0.06
DT	0.86 ± 0.08	0.53 ± 0.19	0.75 ± 0.08	0.8 ± 0.09	0.8 ± 0.09	0.64 ± 0.18	0.82 ± 0.06
NB	0.84 ± 0.09	0.64 ± 0.15	0.77 ± 0.07	0.83 ± 0.08	0.83 ± 0.08	0.66 ± 0.17	0.83 ± 0.06
KNN	0.86 ± 0.06	0.62 ± 0.17	0.78 ± 0.07	0.83 ± 0.08	0.83 ± 0.08	0.67 ± 0.15	0.84 ± 0.05

Sen., sensitivity; Spe., specificity; Acc., accuracy; Pre., precision; P.P.V., positive predictive value; N.P.V., negative predictive value; LR, logistic regression; NN, neural network; RF, random forest; DT, decision tree; NB, naive Bayes; KNN, K-nearest neighbor.

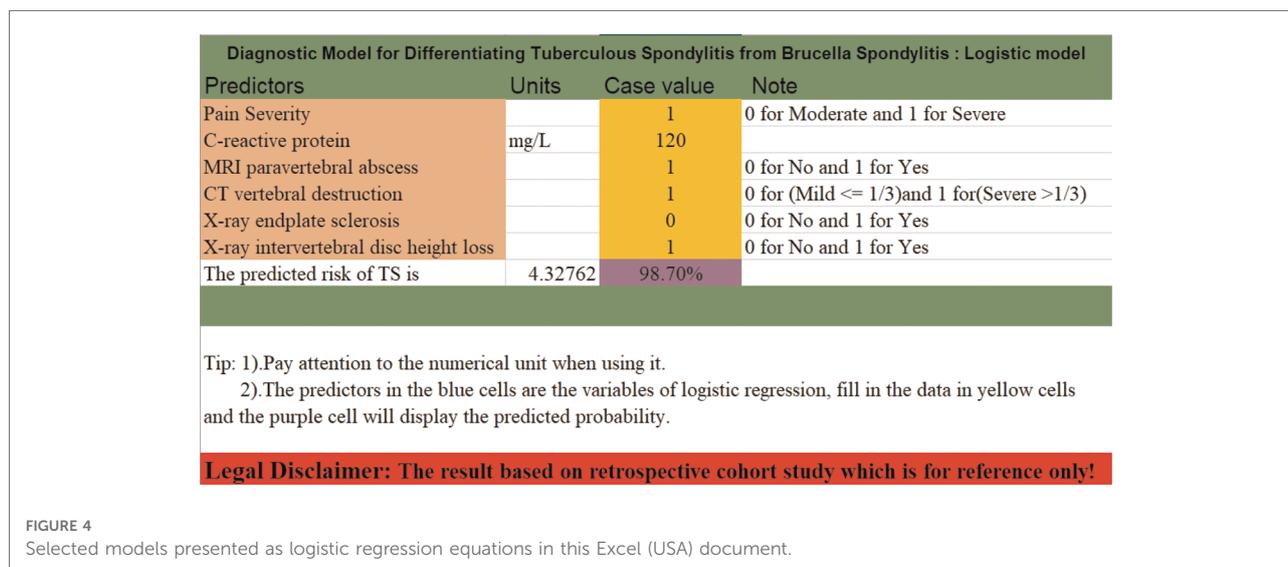


FIGURE 4 Selected models presented as logistic regression equations in this Excel (USA) document.

(95% CI, 0.814–0.900) (Figure 5A) and 0.857 (95% CI, 0.778–0.920) (Figure 5C). The calibration curve showed that the model excellently predicted actual probabilities (Figures 5B,D).

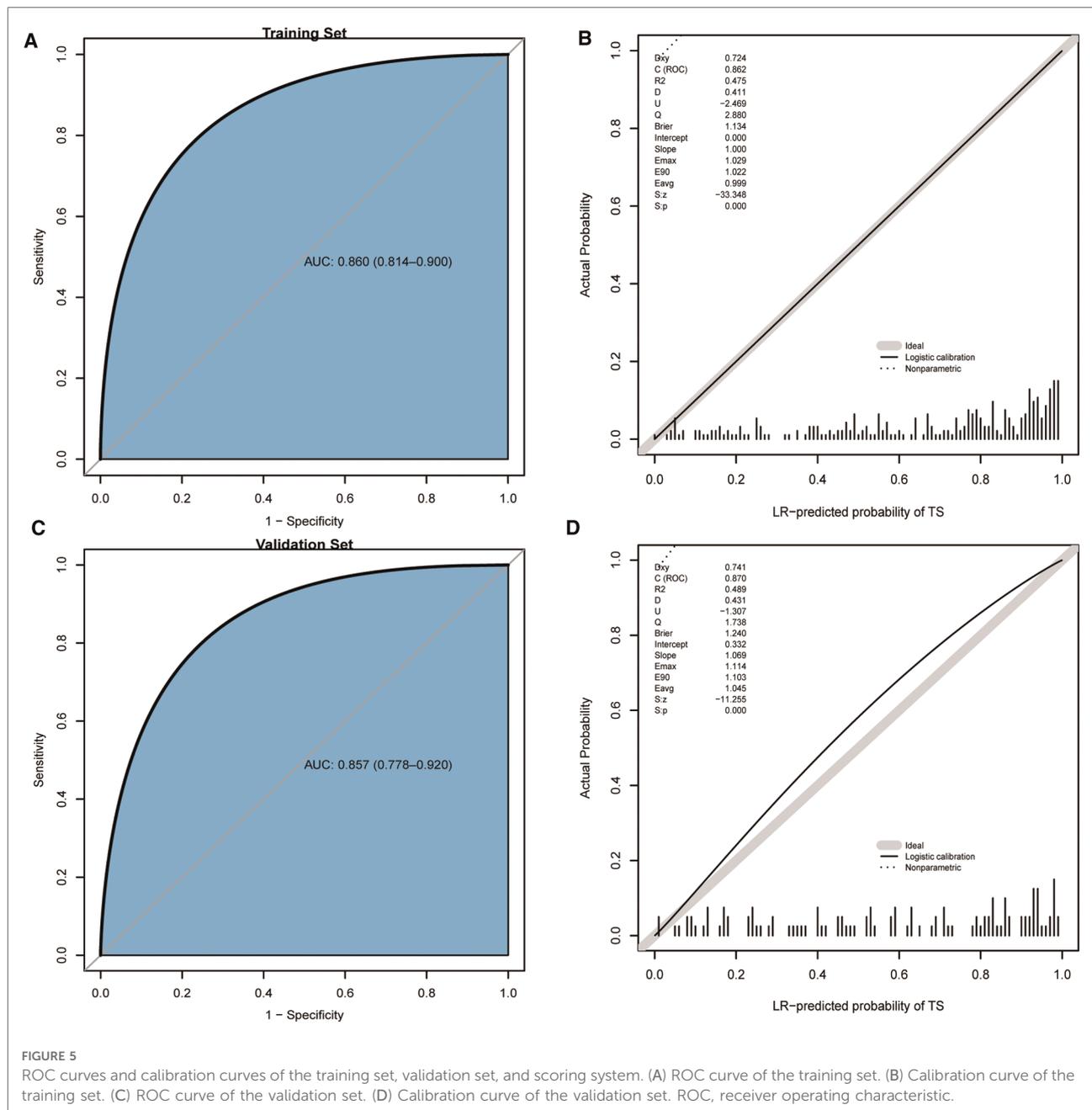
### Clinical efficiency of the model

We implemented DCA to confirm whether it could bring benefit to clinical practice. It can be found that the model had a prominent ability to improve clinical efficiency in predicting TS, as shown in Figure 6.

### Discussion

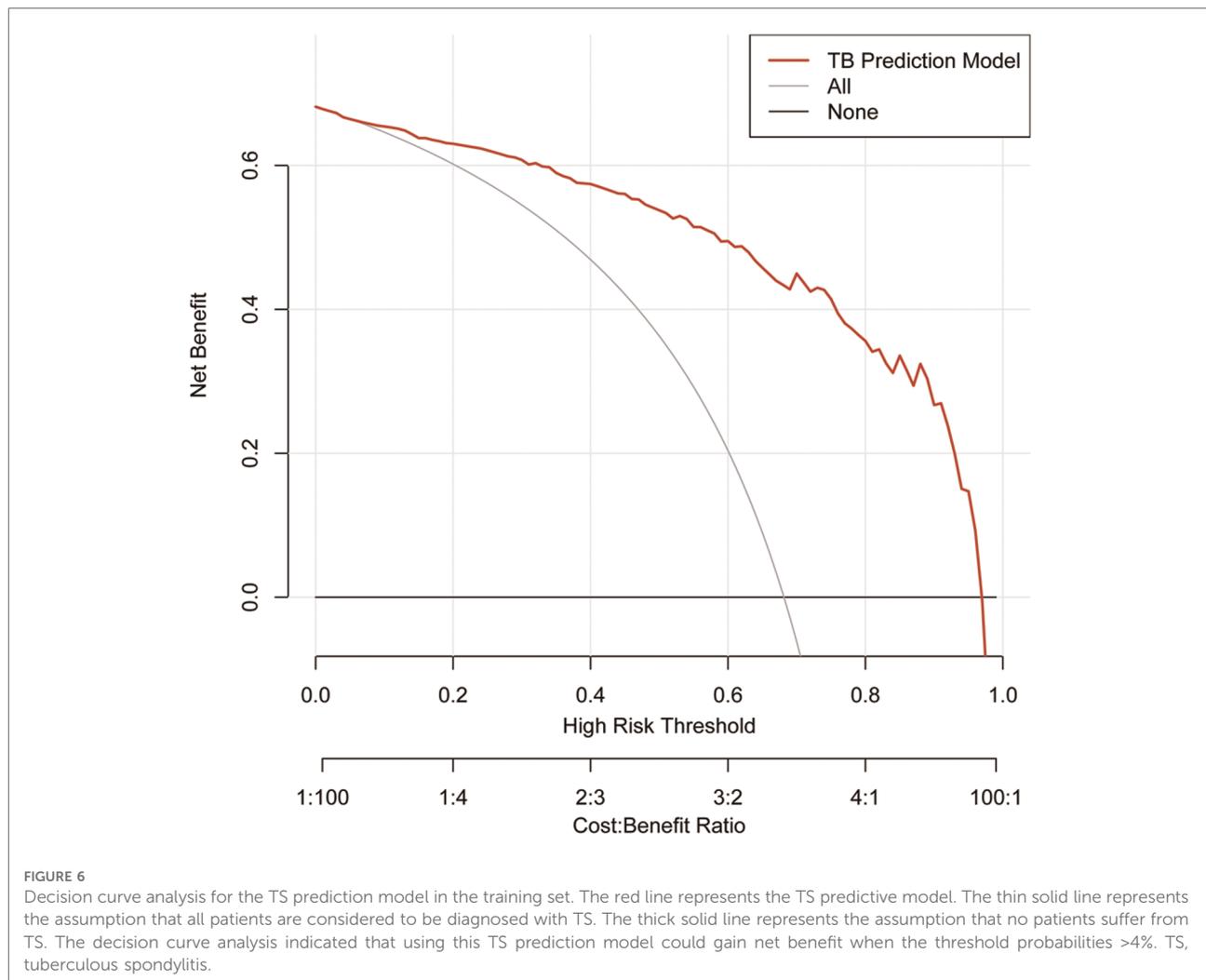
Machine learning has been widely used in many types of research on diseases. As per our best knowledge, this is the first report on exploiting different machine learning algorithms to develop a diagnostic model with noninvasive clinical indices to differentiate between TS and BS. ML approaches vary their performance depending on various

hyperparameters, which play a significant role in decision-making. Finding a set of configurations of hyperparameters is called tuning. It is realized that performance evaluation and tuning are strongly correlated. The nested resampling method we implemented in this research could combined these two procedures to minimize the bias occurring in the whole process. Moreover, the opted model has been visualized as a calculator embedded into an Excel document to encourage further study of its clinical utility. All distinctive predictors selected in the prediction model were basic clinical appearance, laboratory tests, and different imaging data, allowing for routine accessibility in clinical practice. The results displayed that our model possessed excellent discrimination and calibration capacity in two data sets, with AUC values of 0.860 in the training set and 0.857 in the validation set. However, we can find from the above results that the model has the likelihood of misclassification. We assume that this is because of the instability of data. In addition, it somehow depends on the interpretation of the radiologist evaluating the image of patients because the five predictors are related to radiological manifestations.



Both tuberculosis and brucellosis are systemic diseases and remain to be considered public health issues, especially in developing countries, showing higher incidence in the northwest part of China than the other parts of China (27). TS has been mainly discovered in less developed regions because of low income and hygienic status (28). Xinjiang has the second highest incidence of human brucellosis, according to data from the China Public Health Data Center, where patients are mainly pastoralists and veterinarians (29). Previous studies have shown human brucellosis is associated with contact with animals and consumption of uncooked milk

and products from goat and sheep (30–32). In addition, there are other factors also connected to brucellosis like high temperatures, air pollution, wind speed, etc. (33). However, the aforementioned factors can be found in Southern Xinjiang, China. Our statistical results based on the patients enrolled in this study displayed that the southern part of Xinjiang, China shows a higher incidence than the northern part, which agrees well with previous research studies. The clinical diagnosis of spinal tuberculosis usually comprises clinical manifestations, laboratory studies, and imaging data (34). The gold standard for diagnosing spinal TB or BS is



bacterial isolation (culture) from blood, bone marrow, or tissues (35, 36). Nevertheless, confined to the low positive rate of mycobacteria culture or isolation, diagnosis commonly incorporates clinical symptoms, physical examinations, radiographic findings, tissue a microbiological culture, polymerase chain reaction (PCR), and gene detection (37). Due to the resemblance in the clinical manifestation laboratory tests and imaging findings, many patients may be misdiagnosed during the primary phase of the sickness due to delays from insufficient knowledge (38). Early recognition and effective cure are critical in preventing devastating complications (39). Thus, it is urgent to investigate the related features, develop a convenient and sensitive prediction model, and help primary health care clinicians in less developed areas.

In this article, we select six predictors strongly associated with TS, including pain severity, CRP, x-ray intervertebral disc height loss, x-ray endplate, CT vertebral destruction, and MRI paravertebral abscess. To minimize the heterogeneity of

the model to differentiate TS from BS, we chose to acquire features based on the first blood test. We believe that this measure can reduce heterogeneity and boost the model performance.

Patient complaints in TS or BS may initially be effortful to discriminate because of the nature of the illness. Patients with BS often report moderate fever, sweating, malaise, back pain (local pain), and anorexia, whereas patients with TS report back pain, evening pyrexia, generalized body ache, fatigue, body weight loss, neurological abnormalities, and night sweats. Unfortunately, one or more of these symptoms are shown in merely 20%–38% of patients with skeletal tuberculosis (40, 41). Back pain is considered the most frequent complaint of TS. It can be axial pain or radicular pain, which is believed to be the result of the damage to the anterior spinal bodies and mass effect by cold abscess or instability of the spine, nerve root compression, and vertebral body collapse (41, 42). In

clinical practice, pain severity showed variance between TS and BS, and the latter can be found with severe pain degrees the former, which is concordant with previous findings. The result of multivariate logistic regression also proved that point (OR: 0.37, 95% CI, 0.20–0.66,  $p < 0.001$ ). Fever types of the two diseases also show differences in that brucellosis appears to be a moderate ( $\geq 38.5^\circ\text{C}$ ) fever, while tuberculosis is low ( $< 38.5^\circ\text{C}$ ) fever with sweats ( $p < 0.001$ ). However, it was not included in our model. Given the wide range among the patients, their age, gender, and ethnicity, to some degree, may affect the result. However, gender shows a great difference between TS and BS, which might be the result of sampling bias. None of these were selected as predictors in ML models because the training set cannot be represented with a small number of samples. Thus, we maintained that there were no significant differences in demographic characteristics, including ethnicity, gender, history of weight loss, history of tuberculosis in other solid organs, and age, between the BS and TS patients after the scientific and precise analysis of our data, which is in line with previous studies (43).

Clinical laboratory tests, such as WBC count, ESR, and CRP level, which are all nonspecific in showing infectious processes and linked to spondylitis in the majority of cases, are a significant part of clinical diagnoses (40, 42, 44, 45). It can be easily found from our result that CRP levels were higher in TS patients than those in BS patients ( $p < 0.001$ ), which was similar to the results reported in previous studies (46–48). At the same time, contrary to the findings, we did not find a significant difference in WBC count and ESR between patients with TS and BS.

Radiological findings are the keystone of the diagnostic process (49). Plain radiography is usually examined first in patients suspected to have TS or BS, and plain radiography images may exhibit no positive result at the early stage of the disease (50). CT has high sensitivity for early diagnosis. In addition, the identification of the extent of the inflammatory process can also be evaluated in time. Moreover, CT has unreplacable merits of better visualization of the bony details of irregular lytic lesions, sclerosis, disc collapse, and damage to vertebral circumference (51, 52). Previous findings suggest that the diagnosis and differential diagnosis based on MRI of spondylitis patients was qualitative (53, 54). TS and BS are the results of *M. tuberculosis* and *B. melitensis* infections, respectively, which can cause vertebral edema and abscesses, which is reflected by increased T2 values. The lesion level and segments of spinal disease are known to vary according to its etiology. It has been observed that thoracic involvement and multifocal involvement were generally associated with TS (55, 56), a finding consistent with our result. Previous studies have demonstrated that paravertebral abscess, severe bone destruction, and

intervertebral disc height loss were suggestive of TS, while local bone damage and confined paravertebral involvements were suggestive of BS, which can be proved by our results (57). In addition to that, endplate sclerosis and osteophytes are more common in BS than in TS, while disc height loss is more frequent in TS, which is in agreement with previous studies (37, 58, 59).

A previous study indicates no sign of predicting the benefit of ML over LR for clinical prediction models (60). The LR model showed good performance with AUROC, AUPRC, and specificity and no significant difference when compared to SVM and NB. Thus, we selected the logistic regression model to differentiate TS from BS. Previous research studies have largely used nomograms exhibiting predictive models. It is not precise enough and somewhat rough to use, and some factors in this model cannot be computed directly, so a scaling system is chosen to visualize the model (25).

## Limitations

There are several limitations of this research. First, this analysis was based on data acquired from electronic medical records in a single center, and it would be more convincing to use multicenter clinical data. Second, it was hard to determine the phase of disease in this series. In addition, as a retrospective design, the research has a few innate demerits compared to a prospective study. What is more, further prospective studies to validate its efficacy with a larger sample size are still needed.

## Conclusions

The model established in this research revealed better discrimination and calibration capability, and internal cross-validation disclosed that this model can still maintain stability when facing diverse tasks. Then, this model was visualized by a calculator that can quickly identify individuals at risk of TS and help physicians in primary health care in less developed areas with a higher incidence of TS or BS in time.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

## Ethics statement

The studies involving human participants were reviewed and approved by the ethics committee of Xinjiang Medical University Affiliated First Hospital. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

MM designed the study. PY collected and analyzed the data and wrote the manuscript. MM, TX, XC, YFA, TW, WS and MM reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

This study was supported by the Science and Technology Planning Project of Xinjiang Uygur Autonomous Region (No. 2016B03047-3).

## References

- World Health Organization. *Global Tuberculosis report 2014*. Geneva, Switzerland: World Health Organization (2014). Available at: <https://www.who.int/publications/i/item/9789241564809>
- Reid MJA, Arinaminpathy N, Bloom BR, Boehme C, Chaisson R, et al. Building a tuberculosis-free world: the Lancet Commission on Tuberculosis. *Lancet*. (2019) 393(10178):1331–84. doi: 10.1016/S0140-6736(19)30024-8
- Seleem MN, Boyle SM, Sriranganathan NJ. Brucellosis: a re-emerging zoonosis. *Vet Microbiol*. (2010) 140(3–4):392–8. doi: 10.1016/j.vetmic.2009.06.021
- Gautam MP, Karki P, Rijal S, Singh R. Pott's spine and paraplegia. *JNMA J Nepal Med Assoc*. (2005) 44(159):106–15. doi: 10.1002/bjs.9736
- Bundle DR, McGiven J. Brucellosis: improved diagnostics and vaccine insights from synthetic glycans. *Acc Chem Res*. (2017) 50(12):2958–67. doi: 10.1021/acs.accounts.7b00445
- Colmenero JD, Ruiz-Mesa JD, Plata A, Bermudez P, Martin-Rico P, Queipo-Ortuno MI, et al. Clinical findings, therapeutic approach, and outcome of brucellar vertebral osteomyelitis. *Clin Infect Dis*. (2008) 46(3):426–33. doi: 10.1086/525266
- Ulu-Kilic A, Sayar MS, Tutuncu E, Sezen F, Sencan I. Complicated brucellar spondylodiscitis: experience from an endemic area. *Rheumatol Int*. (2013) 33(11):2909–12. doi: 10.1007/s00296-012-2555-5
- Buzgan T, Karahocagil MK, Irmak H, Baran AI, Karsen H, Evirgen O, et al. Clinical manifestations and complications in 1028 cases of brucellosis: a retrospective evaluation and review of the literature. *Int J Infect Dis*. (2010) 14(6):e469–78. doi: 10.1016/j.ijid.2009.06.031
- Wang Y-K, Kuo F-C, Liu C-J, Wu M-C, Shih H-Y, Wang SS, et al. Diagnosis of *Helicobacter pylori* infection: current options and developments. *World J Gastroenterol*. (2015) 21(40):11221. doi: 10.3748/wjg.v21.i40.11221
- Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (tripod): the tripod statement. *Ann Intern Med*. (2015) 162(1):W1–73. doi: 10.7326/M14-0698
- Zhang Z. Multiple imputation with multivariate imputation by chained equation (MICE) package. *Ann Transl Med*. (2016) 4(2):30. doi: 10.3978/j.issn.2305-5839.2015.12.63

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fsurg.2022.955761/full#supplementary-material>.

- Sauerbrei W, Royston P, Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. *Stat Med*. (2007) 26(30):5512–28. doi: 10.1002/sim.3148
- Wan EA. Neural network classification: a Bayesian interpretation. *IEEE Trans Neural Netw*. (1990) 1(4):303–5. doi: 10.1109/72.80269
- Nguyen C, Wang Y, Nguyen HN. Random forest classifier combined with feature selection for breast cancer diagnosis and prognostic. *J Biomed Sci Eng*. (2013) 06(05):551–60. doi: 10.4236/jbise.2013.65070
- Song Y-Y, Ying L. Decision tree methods: applications for classification and prediction. *Shanghai Arch Psychiatry*. (2015) 27(2):130–5. doi: 10.11919/j.issn.1002-0829.215044
- Xu S. Bayesian Naïve Bayes classifiers to text classification. *J Inf Sci*. (2018) 44(1):48–59. doi: 10.1177/0165551516677946
- Peterson LE. K-nearest neighbor. *Scholarpedi*. (2009) 4(2):1883. doi: 10.4249/scholarpedia.1883
- Gimenez O, Lebreton J-D, Choquet R, Pradel R. R2ucare: an R package to perform goodness-of-fit tests for capture-recapture models. *Methods Ecol Evol*. (2018):192468. doi: 10.1111/2041-210x.13014
- Crookston NL, Finley AO. Yaimpute: an R package for kNN imputation. *J Stat Softw*. (2008) 23:1–16. doi: 10.18637/jss.v023.i10
- Benoit K, Watanabe K, Wang H, Nulty P, Obeng A, Müller S, et al. Quanteda: an R package for the quantitative analysis of textual data. *J Open Source Softw*. (2018) 3(30):774. doi: 10.21105/joss.00774
- Genuer R, Poggi J-M, Tuleau-Malot C. Vsurf: an R package for variable selection using random forests. *R J*. (2015) 7(2):19–33. doi: 10.32614/RJ-2015-018
- Ghandi M, Mohammad-Noori M, Ghareghani N, Lee D, Garraway L, Beer MA. Gkmsvm: an R package for gapped-KMER SVM. *Bioinformatics*. (2016) 32(14):2205–7. doi: 10.1093/bioinformatics/btw203
- Zhao H, You J, Peng Y, Feng Y. Machine learning algorithm using electronic chart-derived data to predict delirium after elderly hip fracture surgeries: a retrospective case-control study. *Front Surg*. (2021) 8:634629. doi: 10.3389/fsurg.2021.634629

24. Hunt T. Modelmetrics: rapid calculation of model metrics. *R Package Version 1.2.2.2 1* (2). (2020). Available at: <https://CRAN.R-project.org/package=ModelMetrics>
25. Xu Q, Wang L, Ming J, Cao H, Liu T, Yu X, et al. Using noninvasive anthropometric indices to develop and validate a predictive model for metabolic syndrome in Chinese adults: a nationwide study. *BMC Endocr Disord.* (2022) 22(1):53. doi: 10.1186/s12902-022-00948-1
26. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making.* (2006) 26(6):565–74. doi: 10.1177/0272989X06295361
27. Ozaksoy D, Yucesoy K, Yucesoy M, Kovanlikaya I, Yuce A, Naderi S. Brucellar spondylitis: MRI findings. *Eur Spine J.* (2001) 10(6):529–33. doi: 10.1007/s005860100285
28. Millet JP, Moreno A, Fina L, del Bano L, Orcau A, de Olalla PG, et al. Factors that influence current tuberculosis epidemiology. *Eur Spine J.* (2013) 22(Suppl 4):539–48. doi: 10.1007/s00586-012-2334-8
29. Zheng Y, Zhang L, Wang C, Wang K, Guo G, Zhang X, et al. Predictive analysis of the number of human brucellosis cases in Xinjiang, China. *Sci Rep.* (2021) 11(1):11513. doi: 10.1038/s41598-021-91176-5
30. Lytras T, Danis K, Dounias G. Incidence patterns and occupational risk factors of human brucellosis in Greece, 2004–2015. *Int J Occup Environ Med.* (2016) 7(4):221–6. doi: 10.15171/ijoem.2016.806
31. Jia B, Zhang F, Lu Y, Zhang W, Li J, Zhang Y, et al. The clinical features of 590 patients with brucellosis in Xinjiang, China with the emphasis on the treatment of complications. *PLoS Negl Trop Dis.* (2017) 11(5):e0005577. doi: 10.1371/journal.pntd.0005577
32. Lou P, Wang L, Zhang X, Xu J, Wang K. Modelling seasonal brucellosis epidemics in Bayingolin Mongol autonomous prefecture of Xinjiang, China, 2010–2014. *Biomed Res Int.* (2016) 2016:5103718. doi: 10.1155/2016/5103718
33. Zhu HS, Wang LL, Lin DH, Hong RT, Ou JM, Chen W, et al. [Analysis on epidemiology and spatial-temporal clustering of human brucellosis in Fujian Province, 2011–2016]. *Zhonghua Liu Xing Bing Xue Za Zhi.* (2017) 38(9):1212–7. doi: 10.3760/cma.j.issn.0254-6450.2017.09.014
34. Khanna K, Sabharwal S. Spinal tuberculosis: a comprehensive review for the modern spine surgeon. *Spine J.* (2019) 19(11):1858–70. doi: 10.1016/j.spinee.2019.05.002
35. Jain AK, Kumar J. Tuberculosis of spine: neurological deficit. *Eur Spine J.* (2013) 22(Suppl 4):624–33. doi: 10.1007/s00586-012-2335-7
36. Dong S, Li W, Tang ZR, Wang H, Pei H, Yuan B. Development and validation of a novel predictive model and web calculator for evaluating transfusion risk after spinal fusion for spinal tuberculosis: a retrospective cohort study. *BMC Musculoskelet Disord.* (2021) 22(1):825. doi: 10.1186/s12891-021-04715-6
37. Chen CH, Chen YM, Lee CW, Chang YJ, Cheng CY, Hung JK. Early diagnosis of spinal tuberculosis. *J Formos Med Assoc.* (2016) 115(10):825–36. doi: 10.1016/j.jfma.2016.07.001
38. Jutte PC, van Loenhout-Rooyackers JH, Borgdorff MW, van Horn JR. Increase of bone and joint tuberculosis in the Netherlands. *J Bone Joint Surg Br.* (2004) 86(6):901–4. doi: 10.1302/0301-620X.86B6.14844
39. Ulu-Kilic A, Karakas A, Erdem H, Turker T, Inal AS, Ak O, et al. Update on treatment options for spinal brucellosis. *Clin Microbiol Infect.* (2014) 20(2):O75–O82. doi: 10.1111/1469-0691.12351
40. Garg RK, Somvanshi DS. Spinal tuberculosis: a review. *J Spinal Cord Med.* (2011) 34(5):440–54. doi: 10.1179/2045772311Y.0000000023
41. Cormican L, Hammal R, Messenger J, Milburn HJ. Current difficulties in the diagnosis and management of spinal tuberculosis. *Postgrad Med J.* (2006) 82(963):46–51. doi: 10.1136/pgmj.2005.032862
42. Esteves S, Catarino I, Lopes D, Sousa C. Spinal tuberculosis: rethinking an old disease. *J Spine.* (2017) 06(01). doi: 10.4172/2165-7939.1000358
43. Lan S, He Y, Tiheiran M, Liu W, Guo H. The angiopoietin-like protein 4: a promising biomarker to distinguish brucella spondylitis from tuberculous spondylitis. *Clin Rheumatol.* (2021) 40(10):4289–94. doi: 10.1007/s10067-021-05752-1
44. Jain AK. Tuberculosis of the spine: a fresh look at an old disease. *J Bone Joint Surg Br.* (2010) 92(7):905–13. doi: 10.1302/0301-620X.92B7.24668
45. Shetty A, Kanna RM, Rajasekaran S. Tb spine—current aspects on clinical presentation, diagnosis and management options. *Semin Spine Surg.* (2015) 28(3):150–62. doi: 10.1053/j.semss.2015.07.006
46. Turunc T, Demiroglu YZ, Uncu H, Colakoglu S, Arslan H. A comparative analysis of tuberculous, brucellar and pyogenic spontaneous spondylodiscitis patients. *J Infect.* (2007) 55(2):158–63. doi: 10.1016/j.jinf.2007.04.002
47. Buha I, Skodric-Trifunovic V, Adzic-Vukicevic T, Ilic A, Blanka-Protic A, Stjepanovic M, et al. Relevance of Tnf-alpha, Il-6 and Irak1 gene expression for assessing disease severity and therapy effects in tuberculosis patients. *J Infect Dev Ctries.* (2019) 13(5):419–25. doi: 10.3855/jidc.10949
48. Hammami F, Koubaa M, Feki W, Chakroun A, Rezik K, Smaoui F, et al. Tuberculous and brucellar spondylodiscitis: comparative analysis of clinical, laboratory, and radiological features. *Asian Spine J.* (2021) 15(6):739–46. doi: 10.31616/asj.2020.0262
49. Garcia-Estrada J, Garzon-de la Mora P, Ballesteros-Guadarrama A, Macias-Compan JD, Murillo-Leano M, Navarro-Ruiz A, et al. Electrochemical fixation techniques. II. Electrochemical dog body fixation. Histological study. *Arch Med Res.* (1996) 27(2):127–32. PMID: 8867365; WOS: A1996UR31800005
50. Alvi AA, Raees A, Khan Rehmani MA, Aslam HM, Saleem S, Ashraf J. Magnetic resonance image findings of spinal tuberculosis at first presentation. *Int Arch Med.* (2014) 7(1):12. doi: 10.1186/1755-7682-7-12
51. Ansari S, Amanullah MF, Ahmad K, Rauniyar RK. Pott's spine: diagnostic imaging modalities and technology advancements. *N Am J Med Sci.* (2013) 5(7):404–11. doi: 10.4103/1947-2714.115775
52. Alp E, Doganay M. Current therapeutic strategy in spinal brucellosis. *Int J Infect Dis.* (2008) 12(6):573–7. doi: 10.1016/j.ijid.2008.03.014
53. Gao M, Sun J, Jiang Z, Cui X, Liu X, Wang G, et al. Comparison of tuberculous and brucellar spondylitis on magnetic resonance images. *Spine (Phila Pa 1976).* (2017) 42(2):113–21. doi: 10.1097/BRS.0000000000001697
54. Galhotra RD, Jain T, Sandhu P, Galhotra V. Utility of magnetic resonance imaging in the differential diagnosis of tubercular and pyogenic spondylodiscitis. *J Nat Sci Biol Med.* (2015) 6(2):388–93. doi: 10.4103/0976-9668.160016
55. Erdem H, Elaldi N, Batirel A, Aliyu S, Sengoz G, Pehlivanoglu F, et al. Comparison of brucellar and tuberculous spondylodiscitis patients: results of the multicenter “backbone-1 study”. *Spine J.* (2015) 15(12):2509–17. doi: 10.1016/j.spinee.2015.09.024
56. Sharif HS, Aideyan OA, Clark DC, Madkour MM, Aabed MY, Mattsson TA, et al. Brucellar and tuberculous spondylitis: comparative imaging features. *Radiology.* (1989) 171(2):419–25. doi: 10.1148/radiology.171.2.2704806
57. Liu X, Li H, Jin C, Niu G, Guo B, Chen Y, et al. Differentiation between brucellar and tuberculous spondylodiscitis in the acute and subacute stages by mri: a retrospective observational study. *Acad Radiol.* (2018) 25(9):1183–9. doi: 10.1016/j.acra.2018.01.028
58. Guo H, Lan S, He Y, Tiheiran M, Liu W. Differentiating *Brucella spondylitis* from tuberculous spondylitis by the conventional MRI and MR T2 mapping: a prospective study. *Eur J Med Res.* (2021) 26(1):125. doi: 10.1186/s40001-021-00598-4
59. Celik AK, Aypak A, Aypak C. Comparative analysis of tuberculous and brucellar spondylodiscitis. *Trop Doct.* (2011) 41(3):172–4. doi: 10.1258/td.2011.110013
60. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol.* (2019) 110:12–22. doi: 10.1016/j.jclinepi.2019.02.004