Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Construction of a diagnostic model for rotator cuff injury based on serum inflammatory markers and routine biochemical parameters

xin li

lixin9806250163.com

the third people's hospital of bengbu

xiaojie xue

the third people's hospital of bengbu

fanxing zhuo

the third people's hospital of bengbu

Research Article

Keywords: Rotator cuff injury, Nomogram, Serological markers, Diagnostic model, Retrospective study

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11192377/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Construction of a diagnostic model for rotator cuff injury based on serum inflammatory markers and routine biochemical parameters

First author and corresponding author: xin li
Email Address: lixin980625@163.com
Co-first author: Xiaojie xue
Email Address: xue_xiaojie@163.com
Second author: Fanxing zhuo
Email Address: 1172066575@qq.com
Primary affiliation: the Third People's Hospital of Bengbu

Abstract
Objective: Rotator cuff injury is a common cause of shoulder pain and its diagnosis relies on MRI, which is costly. We aimed to develop a diagnostic prediction model using routine serum markers and clinical features.
Methods: This retrospective case-control study includes 150 patients with MRI-confirmed rotator cuff injury and 150 healthy controls. Clinical data and routine serological markers are collected. Participants are randomly divided into training (n = 212) and validation (n = 88) sets. Variable selection is performed using univariate logistic regression, LASSO regression with 10-fold cross-validation and the Boruta algorithm. Independent predictors identified by multivariable logistic regression are used to construct a nomogram. Model performance is evaluated by ROC curves, calibration plots, Hosmer-Lemeshow test and decision curve analysis, with internal bootstrap validation.
Results: Six variables—gender, age, total cholesterol, globulin, gamma-glutamyl transferase and lymphocyte-to-monocyte ratio—are independent predictors. The nomogram achieves an AUC of 0.880 (bootstrap-corrected 0.879) in the training cohort and 0.921 in the validation cohort. Calibration is satisfactory (Hosmer-Lemeshow P=0.941 and 0.147, respectively). Decision curve analysis confirms clinical utility.
Conclusion: The nomogram shows robust discrimination, calibration, and clinical net benefit, providing a promising non-invasive tool for early screening of rotator cuff injury.

Keywords
Rotator cuff injury, Nomogram, Serological markers, Diagnostic model, Retrospective study

Introduction

Rotator cuff injury is one of the most common causes of shoulder pain and dysfunction. Its pathological process involves multiple mechanisms such as tendon degeneration, oxidative stress and chronic inflammatory response[1]. Epidemiological data show that the prevalence of this disease in the general population increases significantly with age, particularly in middle-aged and elderly populations. Clinical

manifestations mainly include shoulder pain, limited mobility and muscle weakness, severely impairing patients' quality of life. With the intensification of population aging, the burden of rotator cuff injury is also increasing. A study based on New Zealand's national insurance database recorded over 350,000 claims for traumatic rotator cuff injury over a 9-year period, with related medical and socioeconomic costs reaching up to 960 million New Zealand dollars. These figures indicate that this disease has become an important public health issue[2]. A rotator cuff injury is not a static condition but a progressive pathological process involving tendons, muscles and potentially even bony structures. The initial phase after injury is dominated by acute inflammation and catabolism, gradually transitioning to irreversible degenerative changes such as fatty infiltration, fibrosis and muscle atrophy[3]. Without timely intervention, some patients with partial-thickness tears may experience progressive enlargement of the tear over time, potentially progressing to full-thickness or massive tears, ultimately leading to irreparable tendon damage[4]. This structural destruction not only directly weakens the dynamic stability of the shoulder joint but also alters the kinematics of the glenohumeral joint, leading to significant limitations in patients' daily activities, particularly in movements requiring overhead or rotational functions[5]. Furthermore, the negative impact of rotator cuff injury on functional prognosis can persist. Even in professional athletes, partial-thickness tears can lead to a significant decline in throwing-related indicators two years after injury. It is noteworthy that the same injury exhibits different progression trajectories at different ages. Young patients often present with localized injuries, while elderly patients show more extensive tendon degeneration and complex tears, suggesting that the natural course of the injury is significantly age-dependent[6]. Although there has been considerable discussion on the diagnosis and treatment strategies for rotator cuff injury, most studies focus on optimizing treatment plans, such as the application of platelet-rich plasma and rehabilitation interventions. Clinical research on early identification and risk prediction remains relatively limited. Therefore, a comprehensive understanding of the progression patterns and risk factors of rotator cuff injury, and the exploration of objective, easily accessible biological markers associated with this condition, are of significant clinical relevance to prevent disease progression to an irreversible stage.

Currently, the clinical diagnosis of rotator cuff injury primarily relies on physical examination and imaging modalities such as magnetic resonance imaging (MRI). Although physical examination is widely used as a preliminary screening tool, its accuracy is highly susceptible to factors such as the examiner's experience and overlapping patient symptoms, leading to considerable variability in diagnostic efficacy[7]. MRI, with its excellent soft tissue resolution, is recognized as the imaging gold standard for diagnosing rotator cuff injury, demonstrating high sensitivity and specificity for detecting full-thickness tears and providing anatomical basis for surgical planning[8]. However, MRI is relatively expensive, exhibits limited availability in some primary care settings and is subject to interpretative variability regarding partial-thickness tears and early degenerative injuries, thereby compromising inter-reader consistency[9]. Additionally, the non-dynamic nature of MRI examination limits a comprehensive assessment of rotator cuff functional status

to some extent[10]. It is noteworthy that the degree of structural repair is often not entirely consistent with the recovery of clinical symptoms, suggesting that relying solely on imaging assessment may not fully reflect the overall pathophysiological state of rotator cuff injury[11].

In recent years, increasing evidence suggests that rotator cuff injury is not merely a localized tendon lesion. Its etiology and progression may be accompanied by systemic metabolic disorders and inflammatory responses. For example, coordinated changes in the circulating metabolite profile have been observed in patients with full-thickness rotator cuff tears, suggesting that systemic biological signals may elucidate the pathological mechanisms of rotator cuff injury[11]. Specifically, the chronic degenerative process of the rotator cuff tendon is not an isolated case of localized mechanical attrition but is closely related to the body's systemic metabolic state and inflammatory-immune network. Research indicates that the expression of inflammatory, adipogenic and fibrotic markers including high mobility group box 1 (HMGB1), peroxisome proliferator-activated receptor gamma and alpha-smooth muscle actin is significantly upregulated in the rotator cuff muscle tissue of smokers, indicating that systemic exposures can exacerbate tendon degeneration by driving local inflammation and lipid metabolism reprogramming[12]. Furthermore, the extent of intramuscular fatty infiltration following rotator cuff injury shows a significant correlation with MRI-based fatty infiltration grading at the cellular level. Furthermore, the quantity of lipid accumulation within muscle fibers is significantly greater in patients with full-thickness tears compared to those with partial-thickness tears and healthy controls, indicating that disruptions in lipid metabolism play a pivotal role in the degeneration of the tendon-muscle unit[13]. Consistent with these findings, metabolomics studies have also reported significantly elevated concentrations of lipid metabolites such as sphingomyelin and glycerophospholipids in the serum and joint capsule tissue of patients with rotator cuff injury complicated by joint stiffness, wherein these levels are found to be positively correlated with serum total cholesterol, further supporting the hypothesis that lipid metabolism abnormalities contribute to the pathogenesis of rotator cuff disease[14]. Meanwhile, the role of inflammatory response in rotator cuff injury has also received widespread attention. Animal experiments have confirmed that TNF-α expression is elevated locally after rotator cuff injury. This elevation, in turn, exacerbates the inflammatory cascade, oxidative stress and apoptosis via the upregulation of nerve growth factor, indicating that the inflammatory-immune pathway may be a key bridge linking injury and degeneration[15]. Moreover, previous understanding of the pathogenesis of rotator cuff injury has mostly focused on local factors such as tendon degeneration, mechanical stress injury in the hypovascular zone and trauma. However, recent studies have gradually revealed a more complex systemic pathophysiological background[16]. In the context of inflammatory markers, studies have indicated that the preoperative peripheral blood lymphocyte-to-monocyte ratio (LMR) is a potential predictor of functional recovery after arthroscopic rotator cuff repair, suggesting that systemic inflammatory status may play a significant role in the progression of rotator cuff pathology[17]. Genetic evidence based on Mendelian randomization also suggests a

potential association between elevated levels of certain circulating inflammatory proteins and an increased risk of rotator cuff tears[18]. In the field of lipid metabolism, some cross-sectional studies have found that total cholesterol and low-density lipoprotein cholesterol levels are significantly higher in patients with rotator cuff tears compared to those without tears[19]. However, other studies have failed to replicate this association, showing no significant differences in serum triglyceride and total cholesterol concentrations between rotator cuff tear patients and controls[20]. Furthermore, lipidomics-based analysis reveals that changes in specific lipid molecular species are associated with rotator cuff injury and the extent of fatty infiltration within the rotator cuff, suggesting that traditional lipid markers may not fully capture the local metabolic disorders in the rotator cuff[21]. In summary, existing evidence indicates that the etiology and progression of rotator cuff injury involve multiple pathophysiological processes, including local and systemic inflammatory responses, metabolic abnormalities and lipid metabolism disorders. Inflammatory-immune-related serum markers and metabolic indicators may be involved.

Systemic inflammatory response ratios, including the neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR) and lymphocyte-to-monocyte ratio (LMR), integrate information from multiple immune cell subsets, are easy to measure and exhibit high stability. In recent years, they have shown potential auxiliary diagnostic value in the clinical evaluation of various joint diseases. These indicators combine changes in the number of effector cells of innate and adaptive immunity, allowing for a more sensitive reflection of subtle fluctuations in the body's systemic inflammatory status. In diseases such as osteoarthritis and rheumatoid arthritis, studies have confirmed correlations between these ratios and disease activity, inflammatory burden and prognosis, serving as effective supplements to traditional inflammatory markers. However, in the field of rotator cuff injury, where local tendon degeneration and repair impairment are the core pathological features, research on the diagnostic value of systemic inflammatory indicators such as NLR, PLR and LMR remains relatively scarce. Whether they can reflect the systemic inflammatory response associated with rotator cuff injury and assist in clinical diagnosis currently lacks sufficient evidence.

In recent years, constructing disease prediction models based on serological markers has become an important direction in clinical research. In various disease fields, researchers have attempted to integrate multiple serological markers and clinical features, using machine learning algorithms to construct diagnostic or prognostic prediction models, achieving good predictive performance. Notably, studies on the early diagnosis of bone metastasis in non-small cell lung cancer, the prediction of recurrence in idiopathic nephrotic syndrome and the prognostic assessment of nasopharyngeal carcinoma have demonstrated that combined multi-marker models outperform single-marker models[22-24]. These studies suggest that routine laboratory markers contain diagnostic information that has not yet been fully explored. Integrating these markers through appropriate statistical methods is expected to enhance capabilities for early disease identification. In the development process of

diagnostic prediction models, model validation and performance evaluation are key steps connecting model construction and clinical application. A well-performing model not only needs to demonstrate excellent discrimination ability but also requires good calibration and clinical utility. Relying solely on the goodness-of-fit within the training set to judge model quality often leads to overestimation of the model's generalization ability due to overfitting, thereby weakening its application value in actual clinical settings. Therefore, adopting a scientific internal validation strategy to quantify and correct the model's overfitting bias is crucial. Currently, the Bootstrap resampling method, owing to its ability to fully utilize original sample information and effectively estimate model optimism, has emerged as a standard approach for internal validation of diagnostic models. Many studies have employed this method in conjunction with calibration curves and decision curve analysis (DCA) to conduct a comprehensive evaluation of models across multiple dimensions including discrimination, calibration and clinical net benefit[25,26]. For example, in studies on diagnostic models for frailty status in chronic heart failure and MASLD, researchers have demonstrated robust performance of the models through bootstrap internal validation and evaluated their calibration accuracy and clinical utility using calibration curves and DCA[27,28]. However, in the context of rotator cuff injury, a common musculoskeletal disease, most existing studies primarily focus on analyzing associations between individual or limited numbers of serological markers and disease risk, while largely overlooking the interactions between multiple serological markers and their combined diagnostic efficacy. Furthermore, although some studies have attempted to construct multifactorial models, there are few clinically predictive tools based on routine laboratory markers presented in a visual format. This hinders clinicians from applying research findings to individualized risk assessment, thereby impeding the translation of research evidence into clinical practice. Given that diagnosing rotator cuff injuries remains challenging and the performance of diagnostic models utilizing routine serological markers necessitates rigorous validation. This study systematically evaluates the discrimination, calibration and net clinical benefit of the diagnostic model through internal cross-validation and bootstrap resampling to ensure model stability and generalizability, thereby establishing a methodological framework for subsequent external validation and clinical implementation.

Materials and Methods

Study Design and Subjects

This study employs a retrospective case-control design. A total of 150 patients diagnosed with rotator cuff injury via MRI at the Third People's Hospital of Bengbu between July 1, 2024 and December 31, 2025, are included as the case group. Concurrently, healthy individuals without shoulder joint diseases or related systemic disorders who undergo health check-ups at the same hospital's health examination center during the same period are selected as the control group (n = 150).

Data Collection

Clinical data of the patients are collected, including age, gender, height, weight and body mass index (BMI). Venous blood samples are collected from all subjects in an early-morning fasting state and routine serological indicators are analyzed, specifically including: liver function indicators comprising alanine aminotransferase (ALT), aspartate aminotransferase (AST), gamma-glutamyl transferase (GGT), albumin (ALB) and globulin (GLB), lipid indicators: total cholesterol (TC), triglycerides (TG), high-density lipoprotein cholesterol (HDL) and low-density lipoprotein cholesterol (LDL), blood glucose indicator: fasting blood glucose (FBG), renal function indicators: urea, creatinine (Cr) and uric acid (UA), complete blood count indicators: white blood cells (WBC), neutrophil count (Neut), lymphocyte count (Lymph), monocyte count (Mono), hemoglobin (Hb) and platelets (PLT), systemic inflammatory indicators: neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR) and lymphocyte-to-monocyte ratio (LMR).

Rotator Cuff Injury Patients

Inclusion Criteria

1. A confirmed diagnosis of rotator cuff injury established via MRI, encompassing both full-thickness and partial-thickness tears.
2. Patient age of 18 years or older.
3. Completion of routine serological testing and clinical data collection within 24 hours of admission.

Exclusion Criteria

1. Concurrent shoulder pathologies of other etiologies, including subacromial impingement syndrome, adhesive capsulitis (frozen shoulder), shoulder joint infection or tumors.
2. Concomitant systemic infections, autoimmune diseases or active inflammatory conditions.
3. A history of chronic use of glucocorticoids, immunosuppressants or biologic agents.
4. Severe hepatic or renal insufficiency, hematologic disorders or a history of malignant neoplasms.
5. Excessive missing data, defined as > 20% absence of key serological indicators.

Healthy Check-Up Population

Inclusion Criteria

1. Routine health check-ups at the same health examination center concurrently.
2. No clinical symptoms such as shoulder pain or restricted range of motion, and no abnormalities upon physical examination.
3. No history of rotator cuff injury or shoulder joint surgery.

Exclusion Criteria

1. Imaging findings indicating any rotator cuff abnormality, even if asymptomatic.
2. Presence of systemic diseases matching those in the case group, including infections, autoimmune diseases and advanced kidney disease.
3. History of acute infection or surgical intervention within the preceding 3 months.
4. Use of medications that may affect serological indicators, such as glucocorticoids, lipid-lowering drugs or anticoagulants.
5. Incomplete medical records with significant missing data.

Diagnostic Criteria for Rotator Cuff Injury

MRI findings are characterized by high signal intensity within the rotator cuff tendon, discontinuity of the tendon, tearing or retraction.

Statistical Analysis

Statistical analysis is performed using R (version 4.2.1). The dataset is randomly divided into training and validation sets at a 7:3 ratio, with a fixed random seed (seed = 52) to ensure reproducibility. All variable screening and model construction analyses are conducted exclusively on the training set, while the validation set remains blinded throughout the process to evaluate the model's generalizability.

In the training set, univariate logistic regression analysis is first used to screen for statistically significant variables ($P < 0.05$). Subsequently, Lasso regression analysis with 10-fold cross-validation is employed to identify variables with non-zero coefficients at lambda.1se through coefficient screening and variable trajectory analysis. Next, the Boruta algorithm is used for variable selection. Finally, the intersection of variables with non-zero coefficients at lambda.1se and those selected by the Boruta algorithm is computed to define the final included variables. Variables meeting the criteria are incorporated into a multivariate logistic regression model to identify independent predictors with diagnostic value ($P < 0.05$). Based on these independent predictors, a visual diagnostic nomogram is constructed and the included variables are evaluated using ROC curves.

To assess internal stability and correct for overfitting bias, bootstrap resampling (1,000 iterations) is employed to evaluate discrimination via ROC curves, calibration via calibration plots and the Hosmer–Lemeshow test, and clinical utility via decision curve analysis (DCA).

To validate the model's generalizability, the fixed prediction model constructed from the training set is directly applied to the validation set (without variable screening or model reconstruction), and bootstrap resampling (B = 1000 times) is similarly used to evaluate discrimination, calibration and clinical application value via ROC curves, calibration curves and DCA.

Results

Univariate Logistic Regression Analysis

As previously described, the dataset is divided in a 7:3 ratio. Univariate logistic

regression is performed on the training set to preliminarily screen for statistically significant indicators. As shown in Figure 1, the results of the univariate logistic regression analysis indicate that Gender, Age, UA, TC, ALB, GLB, GGT, Hb and LMR are all statistically significant ($P < 0.05$). Other indicators, such as alanine aminotransferase, aspartate aminotransferase, triglycerides, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, Fasting blood glucose, urea, Creatinine, white blood cell count, neutrophil count, lymphocyte count, monocyte count, platelet count, neutrophil-to-lymphocyte ratio and platelet-to-lymphocyte ratio, do not reach the level of statistical significance ($P > 0.05$).

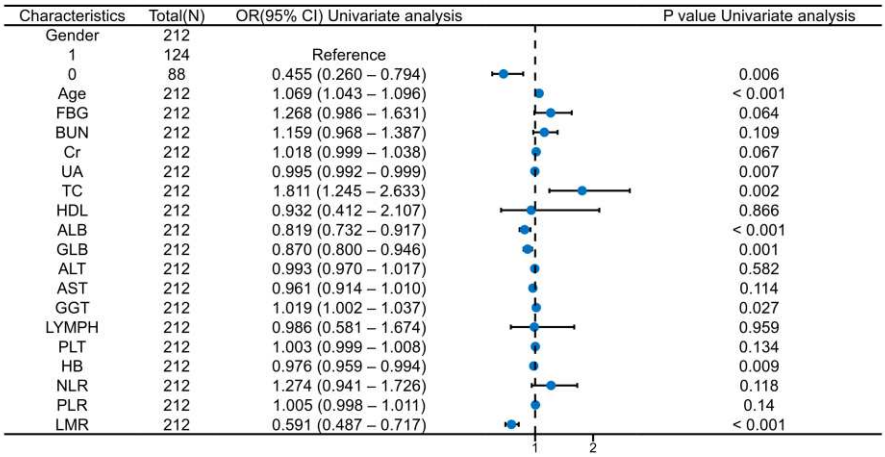


Figure 1: Results of the univariate logistic regression analysis for patients with rotator cuff tears. For the gender variable, 0 denotes male and 1 denotes female.

LASSO Regression Analysis and Boruta Algorithm Analysis

Based on the candidate indicators screened by univariate logistic regression, a LASSO regression model is developed using the training set via 10-fold cross-validation. By combining coefficient screening with coefficient path analysis, seven variables with non-zero coefficients at lambda.1se are ultimately identified: Gender, Age, UA, TC, GLB, GGT and LMR (Figure 2A).

To evaluate the importance of each candidate variable, the Boruta algorithm, which relies on random forests, is applied to the same set of indicators. This algorithm generates shadow attributes by randomly permuting the predictor variables and compares the importance scores of the genuine variables with those of the shadow variables. Variables whose importance consistently exceed the maximum shadow importance across iterations are classified as "confirmed," those that do not are classified as "rejected," and ambiguous ones are labeled "tentative." A multiple-comparison-corrected statistical test is used to determine variable significance, with a significance threshold of two-sided $P < 0.01$ and a maximum of 200 iterations. Tentative variables remaining after the maximum iterations are resolved using the TentativeRoughFix method, based on their median importance relative to the shadow variables. Ultimately confirmed variables are included in the final screening results. As shown in Figure 2C, the Boruta algorithm identifies Gender, Age, UA, TC, ALB, GLB, GGT, Hb and LMR.

The final variable set is determined by taking the intersection of the results from the

LASSO regression and Boruta algorithm analyses, comprising Gender, Age, UA, TC, GLB, GGT and LMR.

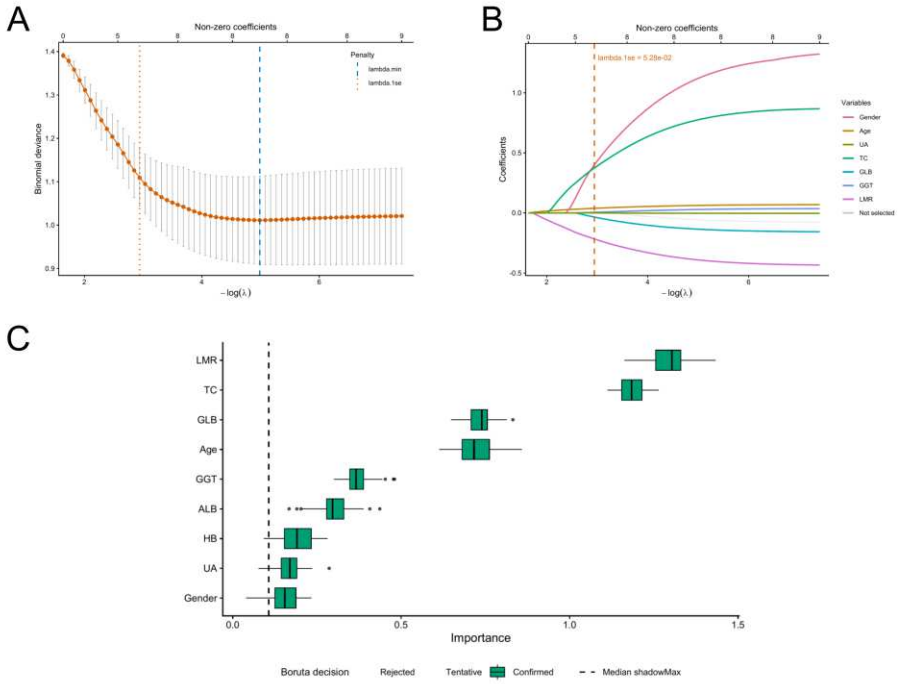


Figure 2: Results of Lasso regression and Boruta algorithm analyses in patients with rotator cuff tears. Figure 2A shows Lasso coefficient screening, Figure 2B illustrates the solution path of variables in the Lasso model and Figure 2C displays the feature selection results obtained via the Boruta algorithm.

Multivariate Logistic Analysis

In the training set, we conduct multivariate logistic regression analysis to test the significance of the above intersection variables, using $P < 0.05$ as the screening criterion. As shown in Figure 3, multivariate logistic regression analysis identifies Gender, Age, TC, GLB, GGT and LMR as statistically significant ($P < 0.05$).

A Variance Inflation Factor (VIF) between 0 and 10 is generally considered indicative of no severe multicollinearity. To further evaluate multicollinearity among the predictor variables, we assess it using the VIF. The results in Table 1 show that the VIF values for all variables are well below the common threshold, specifically: Gender (1.7201), Age (1.1374), TC (1.0853), GLB (1.1515), GGT (1.2731) and LMR (1.0348). None of the above variables exhibit severe multicollinearity, indicating sufficient independence among them.

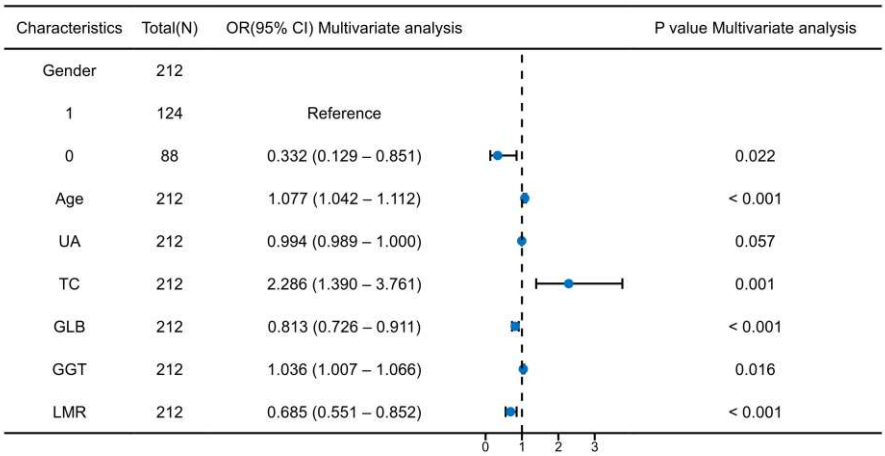


Figure 3: Multivariate Logistic regression analysis of patients with rotator cuff injury. For the gender variable, 0 denotes male and 1 denotes female.

Analysis of multicollinearity among the predictive indicators

Table 1. Variance inflation factors (VIF) for multicollinearity diagnosis among indicators.

Variables	VIF
Gender	1.7201
Age	1.1374
TC	1.0853
GLB	1.1515
GGT	1.2731
LMR	1.0348

ROC Curve Analysis of Independent Predictors

Following a sequential screening process involving univariate logistic regression, LASSO regression, the Boruta algorithm and multivariate logistic regression, LMR, Age, GLB, Gender, TC and GGT are identified as independent predictors. Subsequently, we use ROC curves to evaluate the diagnostic value of these factors. As shown in Figure 4, the AUC values of each independent predictor, in descending order, are as follows: LMR (0.817), Age (0.728), GLB (0.656), Gender (0.594), TC (0.588) and GGT (0.565).

Table 2 presents the results of the ROC curve analysis for each independent predictor of rotator cuff injury. Among all variables, LMR exhibits the highest diagnostic efficacy, with an AUC of 0.818 (95% CI: 0.759–0.876). At an optimal cut-off value of 5.016, it achieves a sensitivity of 0.755 and a specificity of 0.821. Age also shows moderate predictive value, with an AUC of 0.729 (95% CI: 0.660–0.797). At an optimal cut-off value of 49.500 years, sensitivity and specificity are 0.821 and 0.538, respectively. The AUC for GLB is 0.656 (95% CI: 0.580–0.731), and at a cut-off value of 24.650, the sensitivity is 0.472 and specificity is 0.849. Gender has limited discriminatory ability (AUC = 0.595, 95% CI: 0.529–0.660), with a sensitivity of

0.679 and specificity of 0.509 at a cut-off value of 0.500. The AUC for TC is 0.588 (95% CI: 0.510–0.666), with a sensitivity of 0.283 and specificity of 1.000 at a cut-off value of 5.190. The predictive value of GGT is low, with an AUC of 0.566 (95% CI: 0.488–0.643). Notably, because the lower bound of the 95% CI fell below 0.5, GGT lacks significant discriminatory power.

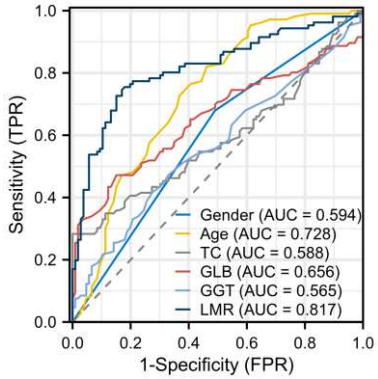


Figure 4: The ROC curve performance of independent predictors in patients with rotator cuff injuries.

Analysis of the diagnostic accuracy of each predictive indicator using ROC curves

Table 2: Results of ROC Curve Analysis for Independent Predictors in Patients with Rotator Cuff Injuries.

variables	Cut-off value	Sensitivity	Specificity	95% Confidence Interval (95% CI)
Gender	0.500	0.679	0.509	0.529 – 0.660
Age	49.500	0.821	0.538	0.660 – 0.797
TC	5.190	0.283	1.000	0.510 – 0.666
GLB	24.650	0.472	0.849	0.580 – 0.731
GGT	21.500	0.472	0.670	0.488 – 0.643
LMR	5.016	0.755	0.821	0.759 – 0.876

Diagnostic Nomogram

Based on the results of the multivariate regression analysis, we construct a predictive model for diagnosing rotator cuff injury, with the formula: Logit(P) = 0.96529 -1.1038 * gender (male=0, female=1) + 0.073718 * age +0.8269 * TC -0.20686 * GLB + 0.035459 * GGT -0.3781 * LMR. Subsequently, we use the "rms" package in R software (version 4.2.1) to visualize this predictive model as a nomogram, as presented in Figure 5. The uppermost section of the nomogram features the 'Points' scale, which quantifies the contribution of each independent predictor. The subsequent rows represent these predictors: Gender, Age, GLB, GGT, LMR and TC. Vertically projecting the value of each factor onto the upper 'Points' scale yields its corresponding individual score. Projecting the sum of all factor scores (Total Points) onto the bottom 'Risk' scale determines the individualized predicted risk. A higher value indicates a greater probability of rotator cuff injury.

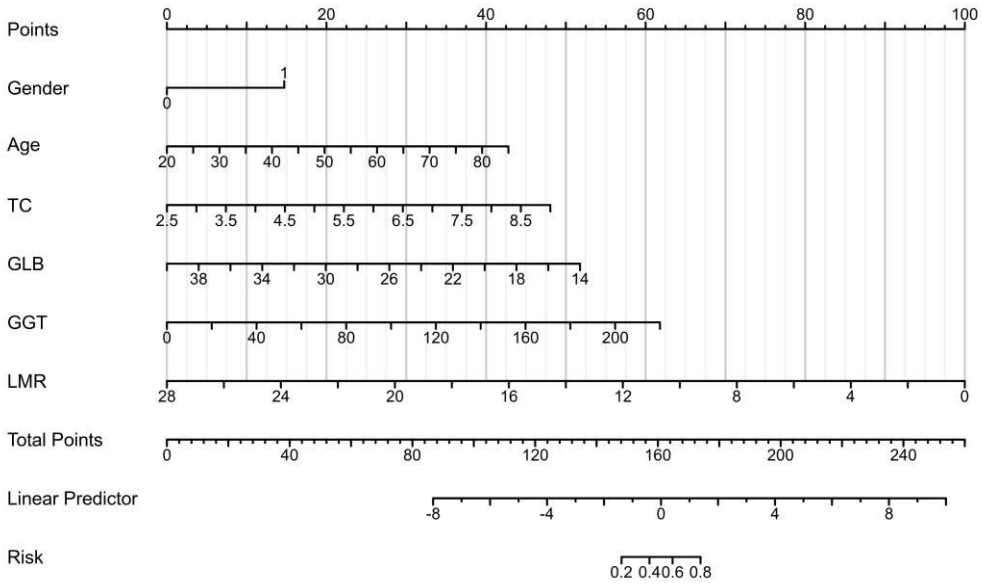


Figure 5: Nomogram for the diagnostic prediction model in patients with rotator cuff tears. For the gender variable, 0 denotes male and 1 denotes female.

ROC Curve of the Diagnostic Model

To systematically evaluate the discriminatory ability of the model, we conduct ROC curve analysis for the diagnostic model using the Bootstrap resampling method (B = 1,000). As shown in Figure 6A, the original C-statistic (AUC) for the training set is 0.879672 (95% CI: 0.833636–0.923759). After Bootstrap correction, the C-statistic is 0.879052, with an optimism of only 0.000620 (n = 212). The extremely low optimism value indicates minimal model overfitting, good discrimination and stability. As shown in Figure 6B, in the validation set, the original C-statistic is 0.921488 (95% CI: 0.865116–0.968949), and the corrected C-statistic is 0.921524, with an Optimism of -0.000036. This negligible negative value, resulting from the observation that the corrected AUC slightly exceeded the original, is attributable to normal sampling variability (n = 88). The AUC of the validation set is slightly higher than that of the training set, further suggesting good generalization ability of the model. In summary, this diagnostic model exhibits high discriminatory ability in both the training and validation sets (AUC > 0.87), and the results are robust after Bootstrap internal validation, indicating its potential value for clinical application.

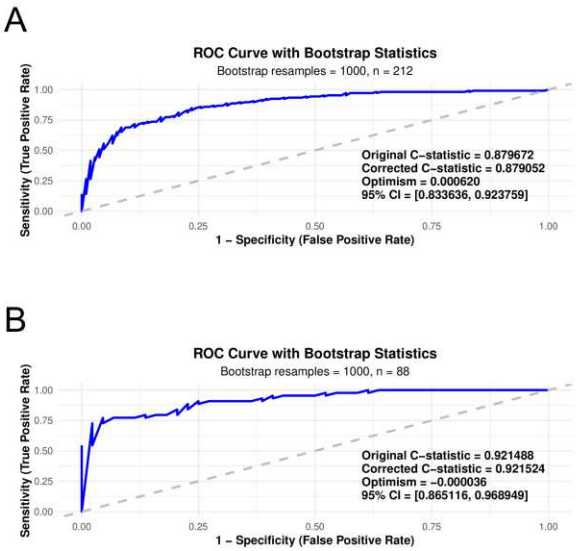


Figure 6: ROC curves of the diagnostic nomogram for patients with rotator cuff tears using the bootstrap resampling method (based on 1,000 bootstrap resamples). Figure 6A displays the ROC curve for the training set, while Figure 6B illustrates the ROC curve for the validation set.

Calibration Curve of the Diagnostic Model

To evaluate the calibration of the model, we use the Bootstrap resampling method (B=1000 times) to plot calibration curves, demonstrating the consistency between the model's predicted probabilities and the actual probabilities. As shown in Figure 7, the apparent curve represents the uncorrected predictions, the bias-corrected curve corresponds to the calibration results after bias correction, and the ideal curve represents the ideal reference line. The closer the curve is to the diagonal, the better the model's fit. The calibration curve for the training set demonstrates strong agreement between predicted probabilities and observed probabilities (Figure 7A). The Hosmer-Lemeshow test results indicate a good model fit ($\chi^2 = 2.893$, $P = 0.941$), suggesting no significant difference between predicted and observed probabilities. The calibration curve is close to the ideal diagonal, and after correction with 1000 Bootstrap resamples (n = 212), it remains closely aligned with the apparent curve, further validating the model's calibration performance in the training set. In the validation set, the calibration curve also demonstrates good calibration performance (Figure 7B). The Hosmer-Lemeshow test results show $\chi^2 = 12.11$, $P = 0.15$, indicating no statistically significant discrepancy between the model's predicted probabilities and the observed outcomes. The calibration curve closely approximated the ideal curve, and after correction with Bootstrap resampling (n=88), it exhibits minimal deviation, indicating good generalization ability of the model in external validation.

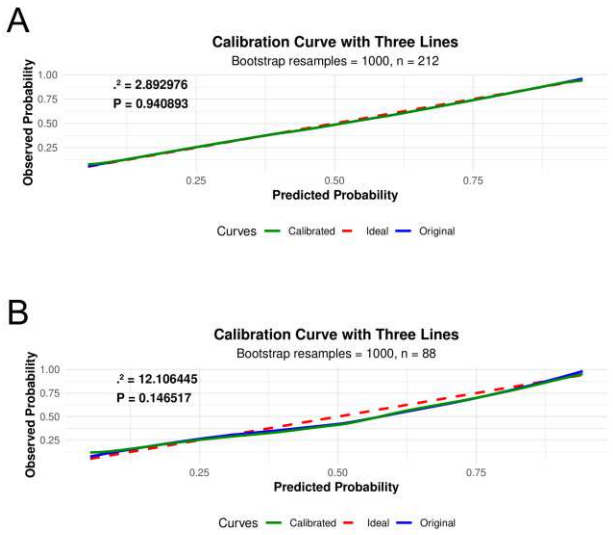


Figure 7: Calibration curves for the rotator cuff injury diagnostic nomogram using the bootstrap resampling method (B=1000 replications). Figure 7A displays the calibration curve for the training set, while Figure 7B displays the calibration curve for the validation set.

Decision Curve Analysis (DCA) Plot

The Decision Curve Analysis (DCA) plot is used to illustrate the variation in net benefit associated with model-based interventions as the risk probability threshold varies, as shown in Figure 8. The DCA plot for the training set shows that across the risk threshold range of 0.08 to 0.95, the model demonstrates a positive net benefit (Figure 8A). Similarly, the validation set exhibits a positive net benefit across the threshold range of 0.05 to 0.94 (Figure 8B).

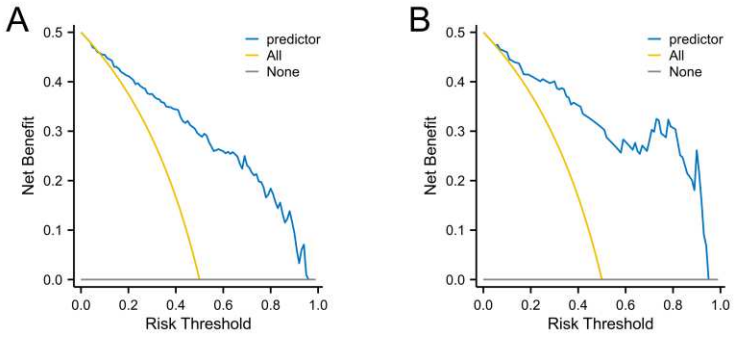


Figure 8: Decision curve analysis (DCA) plots of the diagnostic nomogram in patients with rotator cuff tears. Figure 8A displays results for the training set and Figure 8B for the validation set.

Discussion

Rotator cuff injury is one of the most common tendinopathies and a leading cause of shoulder pain and functional impairment. Accurate diagnosis is essential for establishing treatment plans and improving prognosis. Although MRI has high

diagnostic sensitivity, its high cost and limited availability in some primary healthcare institutions underscore the need for convenient and economical diagnostic tools based on routine clinical indicators[29]. Previous studies have developed diagnostic prediction models for rotator cuff injury using demographic characteristics, physical examination findings, and imaging parameters[30]. Other researchers have incorporated acromial morphological parameters into machine learning models to predict the risk of rotator cuff tears[31]. However, these models often rely on specialized imaging measurements or physical examination maneuvers, which limits their widespread clinical application. Moreover, existing studies have predominantly focused on prognostic prediction after rotator cuff repair[32-34], whereas serological diagnostic models specifically for rotator cuff injury remain scarce. In recent years, systemic energy metabolism and inflammatory status have received increasing attention in the context of degenerative tendon diseases[35]. Dyslipidemia and abnormal immune-inflammatory indicators may contribute to the pathogenesis and progression of rotator cuff injury, providing a biological basis for diagnostic models based on serological markers.

The lymphocyte-to-monocyte ratio (LMR), a readily accessible and cost-effective serological indicator, reflects the dynamic balance between systemic inflammation and immune status. In recent years, low LMR levels have been shown to predict poorer functional outcomes in various diseases, particularly in the tumor microenvironment and inflammatory conditions[17]. In the present study, LMR exhibited the highest univariate diagnostic efficacy among the candidate serological markers for rotator cuff injury, with an AUC superior to that of other traditional indicators, suggesting a strong discriminative ability to distinguish patients from healthy controls. Multivariate analysis further confirmed that low LMR was significantly and independently associated with the risk of rotator cuff injury, indicating that inflammation–immune imbalance may play a pivotal role in the pathogenesis and progression of this condition.

The pathological mechanism of rotator cuff degeneration involves multi-level cellular and molecular events, among which tenocyte apoptosis has been identified as a key link. Previous studies using TUNEL staining and ssDNA detection have revealed abundant apoptotic cells in the degenerated rotator cuff tendons of patients with subacromial impingement syndrome, with the apoptotic index significantly higher than that in normal controls[36]. Clinical biopsy studies have further shown that the proportion of apoptotic cells in rotator cuff tear specimens is significantly higher than that in healthy tissue, and the degree of apoptosis is closely related to the histological grade of tendon degeneration, suggesting that apoptosis participates in the early pathological process of rotator cuff injury. Animal experiments have confirmed that rotator cuff impingement can induce mitochondrial dysfunction and endoplasmic reticulum stress, thereby activating synergistic apoptotic pathways, leading to excessive tenocyte apoptosis and metabolic imbalance, and ultimately promoting the development of degenerative tendon disease[37]. Additionally, in partial-thickness rotator cuff tears, the non-torn articular-side tendon also exhibits significantly increased apoptotic activity, indicating that degeneration is not limited to the tear site

but is widespread throughout the entire rotator cuff tissue[38]. In patients with chronic rotator cuff tears, the expression of apoptosis-related markers is further upregulated and correlates with tear severity. Notably, rotator cuff injury not only involves tenocyte apoptosis but is also accompanied by changes in the local inflammatory microenvironment, such as increased expression of pro-inflammatory cytokines like IL-1β and recruitment of myofibroblasts[39]. Together, these factors constitute a complex pathological network of tendon degeneration.

The pathogenesis of rotator cuff injury involves complex interactions among multiple factors. Previous studies have mostly focused on traditional risk factors such as anatomical abnormalities, traumatic mechanisms, and degenerative changes[40]. The degenerative process of the rotator cuff tendon, however, is not only influenced by mechanical load and age-related changes but is also increasingly regulated by systemic metabolic status and the immune-inflammatory microenvironment. There is growing interest in the role of systemic inflammation in tendon degeneration. The neutrophil-to-lymphocyte ratio (NLR) and platelet-to-lymphocyte ratio (PLR), as classic systemic inflammatory indicators, have been shown to be strongly associated with the incidence and progression of rotator cuff injury. For example, preoperative NLR can predict acute pain after arthroscopic rotator cuff repair and is significantly correlated with postoperative analgesic consumption and nerve block duration[41]. Nevertheless, these studies mainly focused on postoperative outcomes, and direct evidence linking NLR and PLR to the risk of rotator cuff injury remains limited.

Epidemiological evidence on the relationship between routine biochemical indicators such as serum globulin and lipid profile and rotator cuff injury remains sparse. Globulin, as a serological indicator reflecting the overall level of immunoglobulins and inflammatory status, is closely related to the occurrence and development of various chronic degenerative diseases. Total cholesterol, a core marker of lipid metabolism, may also adversely affect tendon tissue homeostasis through the systemic low-grade inflammatory state associated with its abnormal elevation. Based on a rigorously designed case-control study, multivariate logistic regression analysis identified decreased serum globulin levels and increased total cholesterol levels as independent risk factors for rotator cuff injury. This finding provides new clinical epidemiological evidence for the involvement of “metabolic-immune interactions” in the pathophysiology of rotator cuff tendon degeneration. Previous studies have shown that rotator cuff degeneration is not merely the result of mechanical wear but rather a biological process driven by cellular senescence, extracellular matrix degradation, and inflammatory mediator infiltration. In this context, decreased serum globulin levels may reflect changes in humoral immune function or increased immunoglobulin consumption under chronic inflammatory conditions, thereby weakening local immune protection of tendon tissue. On the other hand, abnormal cholesterol metabolism can exacerbate tenocyte apoptosis and degradation of extracellular matrix components by promoting oxidative stress and activating inflammatory signaling pathways. Clinical studies have also observed a correlation between dyslipidemia and an increased risk of tendon pathology, supporting the hypothesis that lipid metabolism disorders promote tendon degeneration. In this study, decreased globulin and

increased cholesterol were simultaneously included in the risk assessment model, suggesting that they may participate in the pathogenesis of rotator cuff injury through synergistic effects: the decline in immune surveillance indicated by reduced globulin and the deterioration of the inflammatory microenvironment induced by elevated cholesterol may form a mutually reinforcing pathological loop during tendon degeneration. This finding expands the etiological understanding of rotator cuff injury beyond mechanical factors and anatomical abnormalities, implying that clinicians should consider patients' systemic metabolic and immune status. However, these findings require validation in larger prospective studies, and the underlying biological mechanisms need further exploration through basic experiments.

Gender and age have previously been well documented as independent predictors of rotator cuff injury, highlighting their epidemiological significance. Specifically, a cross-sectional study of patients with spinal cord injury and paraplegia showed that the risk of rotator cuff tear increased by 11% for each one-year increase in age, similar in magnitude to the effect observed here. Previous research suggests that age-related tendon degeneration may represent a common pathogenic pathway independent of mechanical load[42]. Furthermore, studies of patients with massive irreparable rotator cuff tears confirmed age as an independent risk factor for limited active forward shoulder elevation (OR=1.08), reinforcing the concept that cumulative age-related damage to tendon quality and muscle atrophy exacerbates rotator cuff pathology[43]. In studies examining postoperative healing of small-to-medium-sized rotator cuff tears, age also independently predicted tendon integrity; ROC analysis identified 63 years as the optimal cutoff for successful healing. This indicates that age is not only a risk factor for the onset of rotator cuff injury but also a critical determinant throughout disease progression and repair[44]. These lines of evidence from different populations and disease stages collectively reinforce the epidemiological basis of age as a core risk factor for rotator cuff injury.

Previous explorations of risk prediction for rotator cuff injury have mostly focused on single serological markers or isolated imaging signs. For example, some studies examined the associations of fasting blood glucose levels and insulin resistance with degenerative rotator cuff tears and found that higher fasting blood glucose levels within the normal range are an independent risk factor; however, these studies were limited to glucose metabolism indicators and did not include multi-dimensional systemic parameters. In the imaging field, multiple studies have assessed the risk of rotator cuff tears based on MRI or three-dimensional CT, confirming that anatomical parameters such as the critical shoulder angle and glenoid bone morphology are significantly associated with rotator cuff tears[45]. Other studies developed models to predict re-tear risk after rotator cuff repair, but these mainly rely on intraoperative findings and imaging features such as tendon retraction and fatty infiltration grade, often overlooking the systemic metabolic and inflammatory status reflected by serological indicators. Furthermore, studies predicting the efficacy of rotator cuff repair have primarily focused on structural factors such as age, tear size, and fatty degeneration[46,47]. Although a few studies considered systemic conditions such as metabolic syndrome, research systematically integrating multidimensional indicators

such as serum biochemistry and inflammation-related ratios to construct a comprehensive diagnostic model for rotator cuff injury remains scarce. This single-dimensional research paradigm limits the ability to fully characterize the complex pathophysiology of rotator cuff injury. In recent years, machine learning methods have shown potential in the diagnostic prediction of rotator cuff injury. Some studies have used motion capture and machine learning algorithms to classify rotator cuff tears with an accuracy exceeding 90%, but these methods rely on specialized equipment or technology, making them difficult to implement in routine clinical practice[48]. In contrast, routine serological indicators are easily obtainable, cost-effective, and highly standardized; however, previous studies have rarely integrated these indicators through rigorous multi-step feature selection strategies. LASSO regression, as a regularized linear model, can effectively handle high-dimensional data and shrink the coefficients of irrelevant variables, while the Boruta algorithm, based on random forest feature importance assessment, can robustly distinguish true predictive variables from non-informative ones. Combining these two approaches offers methodological advantages in variable selection. The present study therefore employed a three-step screening strategy involving LASSO regression, the Boruta algorithm, and multivariate logistic regression. Compared with the commonly used method of directly entering univariable-screened variables into multivariate regression, this approach more effectively controls multicollinearity and reduces the risk of overfitting. Consequently, it is methodologically more rigorous, yielding more robust and reproducible variable selection.

Notably, the nomogram model developed in this study is entirely based on routine clinical laboratory indicators and basic demographic information, which facilitates clinical translation. Unlike risk prediction tools that rely on imaging examinations or invasive procedures, this model imposes no additional financial burden on patients and requires no extra medical resources. Studies have shown that clinical prediction models based on routine laboratory indicators are practically feasible and have important auxiliary value for early disease identification and screening of high-risk populations[49]. Classic tools such as the Fagan nomogram have long demonstrated the practical value of rationally integrating laboratory results into clinical decision-making, allowing simple probability calculations to aid clinical judgment. The variables included in this model are all routinely collected during hospital admission or outpatient visits; the detection platforms are widely available, operational procedures are well standardized, and results are reproducible across different levels of medical institutions. Previous literature has also confirmed that nomograms combining CT imaging assessment with routine laboratory indicators show good discriminative performance and clinical utility in predicting the risk of complications from portal hypertension, suggesting that prediction tools based on routine data have broad prospects for optimizing clinical decision-making pathways[50]. Because this model does not require special equipment or advanced imaging, clinicians in primary hospitals and outpatient settings can quickly perform individualized risk assessments using only patients' routine laboratory reports and basic information. This not only helps identify high-risk individuals for rotator cuff

injury but also provides a simple, effective quantitative tool for optimizing referral strategies and resource allocation.

Internal validation of a prediction model is a critical step in assessing its clinical applicability, as it aims to quantify the model's optimism in fitting the data and correct for potential overfitting bias. Although the model developed in this study demonstrated good discriminative performance in the training set, resampling techniques for internal validation are essential to obtain an unbiased estimate of true performance, because model performance is often overestimated in the development sample[51]. Bootstrap resampling, as an effective and widely used internal validation strategy, can robustly estimate the optimism of model performance metrics by repeatedly drawing samples with replacement from the original dataset, thereby providing a quantitative basis for model stability[52]. Previous studies have also shown that the optimism-corrected C-statistic obtained through bootstrap correction provides a more reliable reference for model generalizability[53]. In this study, the bootstrap internal validation results showed that the C-statistic before and after correction was almost identical, with a very low optimism value. This indicates that the model exhibits structural stability, minimal risk of overfitting, and reliable, robust predictive performance. This result is consistent with findings from other studies employing similar validation strategies, in which models maintaining high discriminative ability after internal validation have more prominent clinical potential. Additionally, the low optimism suggests that the combination of predictors included in the model has good internal consistency and that the model is robust to data fluctuations, providing a solid foundation for subsequent external validation in independent populations.

The model's discriminative performance in the validation set was comparable to, or marginally better than, that in the training set. Although this phenomenon is not typical in prediction model studies, similar results have been reported in the literature. Previous studies have shown that when the training set sample size is limited or the data distribution exhibits some heterogeneity, moderate fluctuations or slight increases in validation-set AUC may be attributable to distribution differences caused by random sampling, rather than an indication of overfitting. Furthermore, a slightly higher validation AUC may also be related to a more uniform distribution of baseline characteristics between cases and controls in the validation set, which is particularly common in small validation cohorts.

Results of decision curve analysis further validated the model's value from a clinical utility perspective. Unlike assessments relying solely on discrimination (e.g., AUC), decision curve analysis quantifies the net benefit of a model at different threshold probabilities, thereby more intuitively reflecting its clinical application potential[54,55]. Previous studies have shown that a prediction model with good discrimination may not necessarily yield clinical utility in decision-making, whereas decision curve analysis, by balancing the trade-offs between true positives and false positives, provides a critical basis for evaluating clinical utility[56,57]. In this study, the model demonstrated a positive net benefit over a wide range of threshold probabilities in both the training and validation sets. This result has significant clinical implications. A wide risk threshold interval means that regardless of whether clinicians tend to adopt

conservative intervention strategies (corresponding to lower thresholds) or consider intervention only when the risk is high (corresponding to higher thresholds), the model can provide effective risk stratification information for decision-making. This characteristic is consistent with the decision-curve features reported for risk prediction tools developed in the fields of cardiovascular disease and oncology, which also emphasize the importance of maintaining a robust net benefit across diverse clinical scenarios. For example, a validated risk score for acute coronary syndrome was shown through decision curve analysis to have superior clinical net benefit compared with traditional scores, supporting its application in clinical practice[58]. Moreover, the strong concordance of the net-benefit threshold range between the training and validation sets (0.08–0.95 vs. 0.05–0.94) suggests robust generalizability and stability, supporting potential application across diverse patient populations. Notably, the starting threshold for net benefit was slightly lower in the validation set than in the training set. While this may indicate enhanced decision support in low-risk intervals within the validation cohort, such a minor difference could also result from random fluctuations associated with the smaller validation sample size. Overall, the decision curve analysis provides strong support for the diagnostic model from the perspective of clinical decision utility, suggesting that it has the potential to become a practical tool for assisting clinicians in risk stratification for rotator cuff injury.

As a visualized risk prediction tool, a nomogram's core value lies in its ability to transform complex multifactorial regression models into intuitive, individualized risk scoring systems. In recent years, such tools have demonstrated promising clinical utility in the prediction of rotator cuff injury. For instance, researchers developed a nomogram based on shoulder MRI anatomical features that exhibited high discriminative performance in predicting non-full-thickness rotator cuff injuries, highlighting its potential to avoid unnecessary invasive diagnostic procedures and align with the principles of precision medicine[59]. In predicting rotator cuff lesions related to bone abnormalities, other scholars have used nomograms to integrate bone density and metabolic indicators, showing good calibration and clinical utility[60]. The diagnostic nomogram constructed in this study, based on routine serological indicators and demographic characteristics, follows a similar methodological approach by presenting individualized risk probabilities in an intuitive format. This tool enables clinicians to rapidly estimate a patient's risk of rotator cuff injury in an outpatient setting by summing the scores of individual predictors. The process requires no complex computing equipment or software support, offering significant advantages in simplicity and real-time application, particularly suited to the fast-paced outpatient environment. Furthermore, although imaging examinations are important for diagnosing rotator cuff injury, their application requires consideration of cost-effectiveness and resource optimization. MRI, as a reliable diagnostic method, has high sensitivity and specificity but is relatively expensive[61]. Compared with invasive procedures such as arthrography, MRI is non-invasive, but imaging resources remain relatively limited. Even in settings with abundant medical resources, rationally triaging suspected patients and avoiding unnecessary imaging examinations have important health-economic implications. Evidence has shown that risk stratification

using non-invasive indicators can help identify high-risk individuals for rotator cuff injury, thereby guiding more targeted subsequent examination strategies. The development of this nomogram is consistent with this goal: through risk stratification, it identifies individuals who truly need imaging examinations, improves the utilization efficiency of advanced imaging resources such as MRI, and reduces excessive testing in low-risk individuals. At the same time, this tool can assist clinicians in communicating risk with patients during the initial consultation and provide more individualized diagnostic and treatment recommendations. For high-risk individuals identified through screening, timely referral for imaging will facilitate early detection of rotator cuff injury, preventing worsening of injury due to delayed diagnosis and treatment. As previous studies have pointed out, early intervention is crucial for improving patient prognosis. Therefore, this nomogram has the potential to be integrated into the outpatient diagnostic workflow as an effective auxiliary tool for imaging referral decisions.

This study, using a retrospective case-control design and strictly controlling for the confounding factors of age and gender, preliminarily revealed independent associations of gender, age, total cholesterol, globulin, gamma-glutamyl transferase, and lymphocyte-to-monocyte ratio with rotator cuff injury, and constructed a diagnostic prediction model with high discriminative performance. However, the inherent limitations of the study design constrain the generalizability of our conclusions. Retrospective case-control studies are inherently weaker in causal inference than prospective cohort studies. Their data sources rely on existing medical records and physical examination files, which may introduce information bias, particularly in the measurement of exposure factors, where the degree of standardization is difficult to match the strict quality control of prospective designs. Although the two groups were comparable in age and gender composition, this was a single-center study with geographic limitations and population specificity, so selection bias cannot be fully avoided, potentially limiting the model's applicability to populations of different ethnicities and socioeconomic backgrounds.

The control group consisted of healthy individuals who underwent routine physical examinations during the same period. Although individuals presenting with shoulder-related symptoms were excluded through physical examination at inclusion, and rotator cuff abnormalities were excluded via imaging (such as ultrasound or MRI) when available, not all controls systematically underwent MRI. This design choice may have affected the homogeneity of the control group to some extent. Epidemiological studies have clearly indicated that the prevalence of full-thickness rotator cuff tears increases significantly with age, with the detection rate of asymptomatic rotator cuff tears reaching 20%–40% in individuals over 50 years old. It is worth noting that although asymptomatic and symptomatic rotator cuff tears share overlapping pathological features, the former is typically not accompanied by significant dysfunction or pain complaints. Because this study did not perform MRI screening on all controls, some individuals with possible asymptomatic rotator cuff degeneration or partial tears may have been included in the control group. This could theoretically attenuate the observed differences in inflammation and

metabolism-related serological indicators between the case and control groups, thereby introducing null bias in effect size estimation. Similar issues have been raised in other musculoskeletal disease studies using healthy populations as controls, where "clinical health" with negative imaging findings does not equate to complete histological or structural normality. Future studies that implement standardized MRI screening for control populations would help more accurately define the health status of controls, thereby improving the precision of effect size estimates and the robustness of study conclusions.

Although the diagnostic model for rotator cuff injury constructed in this study based on routine serological indicators achieved satisfactory discriminative performance and calibration, its limitations must be objectively acknowledged. First, in terms of predictor coverage, this model only included routine tests such as complete blood count and biochemistry, without quantitative imaging parameters. Previous studies have shown that quantitative MRI indicators, such as the degree of tendon degeneration, tear retraction distance, and the proportion of muscle fatty infiltration, can provide direct anatomical information reflecting the structural integrity and functional status of rotator cuff tissue. Combining these parameters with serological indicators could significantly improve the accuracy of disease diagnosis and the precision of staging assessment. Second, genetic susceptibility markers are also absent from this model. Established literature confirms that specific gene polymorphisms and epigenetic changes are strongly associated with tendon degeneration and repair capacity. For example, differences in DNA methylation profiles have been shown to be involved in the pathological processes of arthritis and tendon disorders. Such biological information could reveal intrinsic individual differences in susceptibility to rotator cuff injury and provide a molecular-level complement to the model's individualized predictions. Additionally, the impact of rotator cuff injury on patients' functional status and quality of life is multidimensional, and relying solely on laboratory indicators cannot fully reflect the overall disease burden. Studies have shown that serological markers have inherent limitations in predicting disease outcomes, and their association with changes in peripheral blood mononuclear cell subsets and cytokine networks requires further elucidation. Meanwhile, the relationship between structural imaging parameters and functional outcome indicators is often confounded by multiple factors, making it difficult for single-dimensional predictors to capture the full heterogeneity of the disease. Therefore, future research urgently needs to incorporate quantitative imaging parameters, genetic risk scores, and patient-reported outcome measures into the prediction framework, and to leverage multimodal data fusion strategies to construct more comprehensive and robust diagnostic models. The integration of anatomical, molecular, and functional data is anticipated to enhance both the accuracy of rotator cuff injury risk prediction and the specificity of clinical decision-making.

Although the nomogram constructed in this study demonstrated good discrimination and calibration in internal validation, it is important to note that the external validity of the model has not been fully confirmed. The current validation of the model relies solely on an internal validation set generated by random partitioning of the original

dataset. This validation strategy remains an internal evaluation of model fit, and its results tend to overestimate the model's predictive performance in entirely new populations. Previous studies have shown that prediction models built on single-center data, even those demonstrating robust internal validation, may exhibit a significant decline in predictive accuracy when applied to external populations from different geographic regions or healthcare institutions[62,63]. For example, in the development and validation of a lung cancer prediction model, researchers found that although the model had a high concordance index in internal validation, its performance in an independent external cohort showed some attenuation, highlighting the irreplaceable role of external validation in assessing generalizability. Similarly, in fields such as obstetrics/gynecology and urology, differences between internal and external validation performance have been observed in prediction model studies, further emphasizing the necessity of independent external validation[64,65]. Collectively, this evidence suggests that relying solely on an internally partitioned validation set may not fully and accurately assess model performance in diverse clinical practice settings.

Conclusion

In this study, we develop and internally validate a diagnostic nomogram for rotator cuff injury based on routinely available clinical and serological indicators. Through a sequential variable selection strategy combining univariate logistic regression, LASSO regression and the Boruta algorithm, six variables—gender, age, total cholesterol, globulin, gamma-glutamyl transferase and the lymphocyte-to-monocyte ratio—are identified as independent predictors and integrated into the final model. The nomogram demonstrates robust discriminative ability in both the training cohort (AUC = 0.880, bootstrap-corrected 0.879) and the validation cohort (AUC = 0.921), together with satisfactory calibration (Hosmer–Lemeshow P = 0.941 and 0.147, respectively) and a favorable net benefit across a wide range of threshold probabilities on decision curve analysis. These findings suggest that systemic inflammatory and metabolic alterations reflected by routine blood tests carry meaningful diagnostic information for rotator cuff injury, and that a low lymphocyte-to-monocyte ratio may serve as a particularly informative indicator of the underlying inflammation–immune imbalance. Because all predictors are derived from basic demographic data and standard laboratory tests, the model is inexpensive, non-invasive, and readily applicable in primary care and outpatient settings, where it may assist clinicians in the early screening of suspected patients and in optimizing referral for MRI. Nevertheless, given the retrospective, single-center design and the absence of external validation, the present results should be interpreted as exploratory. Prospective, multicenter studies with independent external cohorts are warranted to confirm the generalizability of this model before its integration into routine clinical practice.

Funding

This study is supported by the Science and Technology Project of Bengbu Medical University (Observation on the Efficacy of Acupuncture with Three Shoulder Points

Combined with Rehabilitation Training in the Treatment of Acute Rotator Cuff Injury) (No. 2024byzd491).

Author contributions
All authors contributed to planning, analysis of various papers and writing and concluding the paper.

Availability of data and materials
No datasets were generated or analysed during the current study.

Ethics, Consent to Participate and Consent to Publish declarations
This study is conducted in accordance with the Declaration of Helsinki. The study protocol is reviewed and approved by the Medical Ethics Committee of the Third People's Hospital of Bengbu (Approval No. [2025] K44). Owing to the retrospective nature of the study, the requirement for informed consent is waived by the ethics committee. All authors have read and approved the publication of this article.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Clinical trial number
Not applicable.

References
[1] Yuan T, Qian H, Yu X, et al. Proteomic analysis reveals rotator cuff injury caused by oxidative stress[J]. Ther Adv Chronic Dis, 2021, 12: 20406223209870 57.

[2] Bolam S M, Konar S, Gamble G, et al. Ethnicity, sex, and socioeconomic disparities in the treatment of traumatic rotator cuff injuries in Aotearoa/New Zealand[J]. J Shoulder Elbow Surg, 2023, 32(1): 121–132.

[3] Vasquez-Bolanos L S, Gibbons M C, Ruoss S, et al. Transcriptional Time Course After Rotator Cuff Tear[J]. Front Physiol, 2021, 12: 707116.

[4] Sidiropoulos K, Samundeeswari S, Giannatos V, et al. Partial Cuff Repair in Rotator Cuff Tears: Current Concepts and Clinical Considerations[J]. Indian J Orthop, 2025, 59(6): 743–755.

[5] Vidt M E, Santago A C, 2nd, Marsh A P, et al. The effects of a rotator cuff tear on activities of daily living in older adults: A kinematic analysis[J]. J Biomech, 2016, 49(4): 611–617.

[6] Koukos C, Montoya F, Samundeeswari S, et al. Are we misunderstanding rotator cuff lesions? - Age-based patterns of rotator cuff tears and their implications[J]. Int Orthop, 2025, 49(7): 1723–1730.

[7] Gowda C S, Mirza K, Galagali D A. Rotator Cuff Tears: Correlation Between Clinical Examination, Magnetic Resonance Imaging and Arthroscopy[J]. Cureus, 2024, 16(3): e56065.

[8] Etancelin-Jamet M, Bouilleau L, Martin A, et al. Diagnostic value of angled oblique sagittal images of the supraspinatus tendon for the detection of rotator cuff tears on MR imaging[J]. Diagn Interv Imaging, 2017, 98(2): 161–169.

[9] Duque D F, Montoya Restrepo J L, Ayala Torres J D, et al. Diagnostic Performance of Articular Ultrasound Versus Magnetic Resonance Imaging in the Approach to Rotator Cuff Lesions at a Referral Hospital[J]. Cureus, 2025, 17(3): e81149.

[10] Toh Y. Ultrasound Versus Magnetic Resonance Imaging as First-Line Imaging Strategies for Rotator Cuff Pathologies: A Comprehensive Analysis of Clinical Practices, Economic Efficiency, and Future Perspectives[J]. Cureus, 2024, 16(4): e59231.

[11] Sarıtaş T B, Ertürk C, Arıcı V, et al. Serum (1)H-NMR Metabolomic Profiling of Full-Thickness Rotator Cuff Tears: A Hypothesis-Generating Case-Control Pilot Study[J]. NMR Biomed, 2026, 39(7): e70318.

[12] Lee Y S, Kim J Y, Ki S Y, et al. Influence of Smoking on the Expression of Genes and Proteins Related to Fat Infiltration, Inflammation, and Fibrosis in the Rotator Cuff Muscles of Patients With Chronic Rotator Cuff Tears: A Pilot Study[J]. Arthroscopy, 2019, 35(12): 3181–3191.

[13] Ruderman L, Leinroth A, Rueckert H, et al. Histologic Differences in Human Rotator Cuff Muscle Based on Tear Characteristics[J]. J Bone Joint Surg Am, 2022, 104(13): 1148–1156.

[14] Lee H J, Hong O K, Kwak D H, et al. Metabolic profiling of serum and tissue from the rotator interval and anterior capsule in shoulder stiffness: a preliminary study[J]. BMC Musculoskelet Disord, 2019, 20(1): 364.

[15] Xiang L, Deng H, Zhou S. Effects of TNF-α on Behaviour and Inflammation in Rats with Rotator Cuff Injury through NGF[J]. Discov Med, 2024, 36(185): 1241–1249.

[16] Pandey V, Jaap Willems W. Rotator cuff tear: A detailed update[J]. Asia Pac J Sports Med Arthrosc Rehabil Technol, 2015, 2(1): 1–14.

[17] Sun Y, Lin J, Luo Z, et al. Preoperative Lymphocyte to Monocyte Ratio Can Be a Prognostic Factor in Arthroscopic Repair of Small to Large Rotator Cuff Tears[J]. Am J Sports Med, 2020, 48(12): 3042–3050.

[18] Wang Y, Xu B, Ma C, et al. Association between inflammatory proteins and rotator cuff tears: a bidirectional Mendelian randomization study[J]. Sci Rep, 2025, 15(1): 26987.

[19] Abboud J A, Kim J S. The effect of hypercholesterolemia on rotator cuff disease[J]. Clin Orthop Relat Res, 2010, 468(6): 1493–1497.

[20] Longo U G, Franceschi F, Spiezia F, et al. Triglycerides and total serum cholesterol in rotator cuff tears: do they matter?[J]. Br J Sports Med, 2010, 44(13): 948–951.

[21] Huang H, Jiesisibieke D, Zhou X, et al. A lipid metabolite lipidomics assay for prediction and severity evaluation of rotator cuff injury[J]. Front Nutr, 2022, 9: 1000947.

[22] Teng X, Han K, Jin W, et al. Development and validation of an early diagnosis model for bone metastasis in non-small cell lung cancer based on serological characteristics of the bone metastasis mechanism[J]. EClinicalMedicine, 2024, 72: 102617.

[23] Zhao M, Li P, Chen F, et al. Development and evaluation of machine learning models for predicting relapse in idiopathic nephrotic syndrome[J]. Front Endocrinol (Lausanne), 2026, 17: 1687315.

[24] Lai C, Zhang C, Lv H, et al. A novel prognostic model predicts overall survival in patients with nasopharyngeal carcinoma based on clinical features and blood biomarkers[J]. Cancer Med, 2021, 10(11): 3511–3523.

[25] Weng H, Zhao Y, Xu Y, et al. A Diagnostic Model for Breast Lesions With Enlarged Enhancement Extent on Contrast-Enhanced Ultrasound Improves Malignancy Prediction[J]. Ultrasound Med Biol, 2023, 49(7): 1535–1543.

[26] Leerink J M, Feijen E A M, de Baat E C, et al. A Biomarker-Based Diagnostic Model for Cardiac Dysfunction in Childhood Cancer Survivors[J]. JACC CardioOncol, 2024, 6(2): 236–247.

[27] Gu Y, Li C, Yan J, et al. Development of a diagnostic model focusing on nutritional indicators for frailty classification in people with chronic heart failure[J]. Eur J Cardiovasc Nurs, 2022, 21(4): 356–365.

[28] Cao X, Xiao X, Jiang P, et al. Construction and evaluation of a diagnostic model for metabolic dysfunction-associated steatotic liver disease based on advanced glycation end products and their receptors[J]. Front Med (Lausanne), 2025, 12: 1539708.

[29] Jain N B, Fan R, Higgins L D, et al. Does My Patient With Shoulder Pain Have a Rotator Cuff Tear?: A Predictive Model From the ROW Cohort[J]. Orthop J Sports Med, 2018, 6(7): 2325967118784897.

[30] Águila-Ledesma I R, Córdova-Fonseca J L, Medina-Pontaza O, et al. [Diagnostic value of a predictive model for complete ruptures of the rotator cuff associated to subacromial impingement][J]. Acta Ortop Mex, 2017, 31(3): 108–112.

[31] Hou Z, Wan Q, Wan L. Machine learning-based prediction of rotator cuff tears using anatomical parameters: a retrospective cohort study[J]. BMC Musculoskelet Disord, 2026, 27(1).

[32] Frangiamore S, Dornan G J, Horan M P, et al. Predictive Modeling to Determine Functional Outcomes After Arthroscopic Rotator Cuff Repair[J]. Am J Sports Med, 2020, 48(7): 1559–1567.

[33] Audigé L, Aghlmandi S, Grobet C, et al. Prediction of Shoulder Stiffness After Arthroscopic Rotator Cuff Repair[J]. Am J Sports Med, 2021, 49(11): 3030–3039.

[34] Hu Y, Chen S, Zhang H, et al. Development and validation of a risk prediction model for re-tear after arthroscopic rotator cuff repair[J]. Sci Rep, 2025, 15(1): 38778.

[35] Ding X, Zhang Y, Chen Y, et al. Predicting rotator cuff retear after arthroscopic repair: a lipoprotein-based model[J]. J Shoulder Elbow Surg, 2025, 34(9): 2107–2113.

[36] Tuoheti Y, Itoi E, Pradhan R L, et al. Apoptosis in the supraspinatus tendon with stage II subacromial impingement[J]. J Shoulder Elbow Surg, 2005, 14(5): 535–541.

[37] Yu S, Wei K, Deng X H, et al. Rotator Cuff Impingement Damages the Organelles of Tenocytes, Resulting in Excessive Tenocyte Apoptosis and Tendinopathy[J]. Am J Sports Med, 2026, 54(5): 1103–1113.

[38] Yang H, Lee H J, Lee Y G, et al. Integrity of the Untorn Articular-Sided Tendon in Bursal-Sided Partial-Thickness Rotator Cuff Tear: A Comparative Study of Apoptotic Activity in Torn and Untorn Layers[J]. Am J Sports Med, 2018, 46(10): 2478–2485.

[39] Ko J Y, Wang F S, Huang H Y, et al. Increased IL-1beta expression and myofibroblast recruitment in subacromial bursa is associated with rotator cuff lesions with shoulder stiffness[J]. J Orthop Res, 2008, 26(8): 1090–1097.

[40] Testa E J, Katz L, Zhang H, et al. Rotator Cuff Tears to Shoulder Instability: The Relationship Between Acromial Morphology and Shoulder Pathology[J]. JBJS Rev, 2024, 12(1).

[41] Öner K, Okutan A E, Ayas M S, et al. Predicting postoperative pain with neutrophil/ lymphocyte ratio after arthroscopic rotator cuff repair[J]. Asia Pac J Sports Med Arthrosc Rehabil Technol, 2020, 20: 24–27.

[42] Pepke W, Brunner M, Abel R, et al. [Risk factors for the development of rotator cuff tears in individuals with paraplegia : A cross-sectional study][J]. Orthopade, 2018, 47(7): 561–566.

[43] Reddy R P, Sabzevari S, Karimi A, et al. Pseudoparesis and pseudoparalysis in the setting of massive irreparable rotator cuff tear: demographic, anatomic, and radiographic risk factors[J]. J Shoulder Elbow Surg, 2024, 33(2): e88–e96.

[44] Erşen A, Şahin K, Albayrak M O. Older age and higher body mass index are independent risk factors for tendon healing in small- to medium-sized rotator cuff tears[J]. Knee Surg Sports Traumatol Arthrosc, 2023, 31(2): 681–690.

[45] Nakamura Y, Yokoya S, Matsubara Y, et al. Relationship between scapular body morphology and rotator cuff tears: a threedimensional computed tomography study[J]. Acta Orthop Traumatol Turc, 2024, 58(3): 167–170.

[46] Lee Y S, Jeong J Y, Park C D, et al. Evaluation of the Risk Factors for a Rotator Cuff Retear After Repair Surgery[J]. Am J Sports Med, 2017, 45(8): 1755–1761.

[47] Takeuchi N, Kozono N, Nishii A, et al. Stump classification was correlated with retear in the suture-bridge and double-row repair techniques for arthroscopic rotator cuff repair[J]. Knee Surg Sports Traumatol Arthrosc, 2021, 29(8): 2587–2594.

[48] Darevsky D M, Hu D A, Gomez F A, et al. Algorithmic assessment of shoulder function using smartphone video capture and machine learning[J]. Sci Rep, 2023, 13(1): 19986.

[49] Zhang S, Wang H, Lin S, et al. Construction of a prediction model for pneumonia in children undergoing neurosurgery based on regular clinical laboratory tests and baseline information[J]. Front Pediatr, 2025, 13: 1638012.

[50] Cheng W, Wang K Y, Li W Q, et al. CT-based nomogram predicts esophageal gastric variceal bleeding in noncirrhotic portal hypertension caused by hepatic schistosomiasis[J]. BMC Med Inform Decis Mak, 2025, 25(1): 8.

[51] Kirton L, Riley R, Gaunt P, et al. Development and internal validation of prognostic models for event-free survival in newly diagnosed Ewing sarcoma patients based on routinely collected clinical characteristics from the European randomised controlled trial, EE2012[J]. BMJ Open, 2024, 14(12): e082941.

[52] Buccheri S, Capodanno D, Barbanti M, et al. A Risk Model for Prediction of 1-Year Mortality in Patients Undergoing MitraClip Implantation[J]. Am J Cardiol, 2017, 119(9): 1443–1449.

[53] Chang C F, Huang P W, Chen J S, et al. Prognostic Factors for Advanced Pancreatic Cancer Treated with Gemcitabine Plus S-1: Retrospective Analysis and Development of a Prognostic Model[J]. Cancers (Basel), 2019, 11(1).

[54] Tiberi D, Rodrigues G, Pickles T, et al. External validation of the ProCaRS nomograms and comparison of existing risk-stratification tools for localized prostate cancer[J]. Can Urol Assoc J, 2017, 11(3-4): 94–100.

[55] Foerster B, Abufaraj M, Matin S F, et al. Pretreatment Risk Stratification for Endoscopic Kidney-sparing Surgery in Upper Tract Urothelial Carcinoma: An International Collaborative Study[J]. Eur Urol, 2021, 80(4): 507–515.

[56] Wu C, Yu S, Zhang Y, et al. CT-Based Radiomics Nomogram Improves Risk Stratification and Prediction of Early Recurrence in Hepatocellular Carcinoma After Partial Hepatectomy[J]. Front Oncol, 2022, 12: 896002.

[57] Yang D, Su Y, Zhao F, et al. A Practical Nomogram and Risk Stratification System Predicting the Cancer-Specific Survival for Patients With Advanced Hepatocellular Carcinoma[J]. Front Oncol, 2022, 12: 914192.

[58] Molenaar M A, Selder J L, Schmidt A F, et al. Validation of machine learning-based risk stratification scores for patients with acute coronary syndrome treated with percutaneous coronary intervention[J]. Eur Heart J Digit Health, 2024, 5(6): 702–711.

[59] Wu H, Zuo Z, Li Y, et al. Anatomic characteristics of shoulder based on MRI accurately predict incomplete rotator cuff injuries in patients: relevance for predictive, preventive, and personalized healthcare strategies[J]. Epma j, 2023, 14(3): 553–570.

[60] Liu G, Li W, Zhang L, et al. The role of vitamin D on rotator cuff tear with osteoporosis[J]. Front Endocrinol (Lausanne), 2022, 13: 1017835.

[61] Fazal Gafoor H, Jose G A, Mampalli Narayanan B. Role of Magnetic Resonance Imaging (MRI) in the Diagnosis of Rotator Cuff Injuries and Correlation With Arthroscopy Findings[J]. Cureus, 2023, 15(12): e50103.

[62] Zhang D, Zhou M, Zhang M, et al. Development and external validation of clinical predictive model for stress urinary incontinence in Chinese women : a multicenter retrospective study[J]. BMC Womens Health, 2024, 24(1): 532.

[63] Piccinelli M L, Barletta F, Tappero S, et al. Development and External Validation of a Novel Nomogram Predicting Cancer-specific Mortality-free Survival in Surgically Treated Papillary Renal Cell Carcinoma Patients[J]. Eur Urol Focus, 2023, 9(5): 799–806.

[64] Chen Y, Jiang J, He M, et al. Nomogram for predicting difficult total laparoscopic hysterectomy: a multi-institutional, retrospective model development and validation study[J]. Int J Surg, 2024, 110(6): 3249–3257.

[65] Giganti F, Coppola A, Ambrosi A, et al. Apparent diffusion coefficient in the evaluation of side-specific extracapsular extension in prostate cancer: Development and external validation of a nomogram of clinical use[J]. Urol Oncol, 2016, 34(7): 291.e299–291.e217.