Drug Design, Development and Therapy

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Risk Prediction of Liver Injury in Pediatric Tuberculosis Treatment: Development of an Automated Machine Learning Model

Ying Zeng[1,*], Hong Lu[1,*], Sen Li[2,*], Qun-Zhi Shi[1], Lin Liu[1], Yong-Qing Gong[1], Pan Yan[1]

[1]Department of Pharmacy, The Affiliated Changsha Central Hospital, Hengyang Medical School, University of South China, Changsha, 410004, People's Republic of China; [2]Department of Pharmacy, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, 430030, People's Republic of China

*These authors contributed equally to this work

Correspondence: Pan Yan, Department of Pharmacy, The Affiliated Changsha Central Hospital, Hengyang Medical School, University of South China, Changsha, People's Republic of China, Email 2022050025@usc.edu.cn

Drug Design, Development and Therapy downloaded from https://www.dovepress.com/
For personal use only.

Purpose: Drug-induced liver injury (DILI) is one of the most common and serious adverse drug reactions related to first-line anti-tuberculosis drugs in pediatric tuberculosis patients. This study aims to develop an automatic machine learning (AutoML) model for predicting the risk of anti-tuberculosis drug-induced liver injury (ATB-DILI) in children.

Methods: A retrospective study was performed on the clinical data and therapeutic drug monitoring (TDM) results of children initially treated for tuberculosis at the affiliated Changsha Central Hospital of University of South China. After the features were screened by univariate risk factor analysis, AutoML technology was used to establish predictive models. The area under the receiver operating characteristic curve (AUC) was used to evaluate model's performance, and then the TreeShap algorithm was employed to interpret the variable contributions.

Results: A total of 184 children were enrolled in this study, of whom 19 (10.33%) developed ATB-DILI. Univariate analysis showed that seven variables were risk factors for ATB-DILI, including the plasma peak concentration (C_{max}) of rifampicin, body mass index (BMI), alanine aminotransferase, total bilirubin, total bile acids, aspartate aminotransferase and creatinine. Among the numerous predictive models constructed by the "H2O" AutoML platform, the gradient boost machine (GBM) model exhibited the superior performance with AUCs of 0.838 and 0.784 on the training and testing sets, respectively. The TreeShap algorithm showed that C_{max} of rifampicin and BMI were important features that affect the AutoML model's performance.

Conclusion: The GBM model established by AutoML technology shows high predictive accuracy and interpretability for ATB-DILI in children. The prediction model can assist clinicians to implement timely interventions and mitigation strategies, and formulate personalized medication regimens, thereby minimizing potential harm to high-risk children of ATB-DILI.

Keywords: anti-tuberculosis drug-induced liver injury, children, retrospective study, automatic machine learning, gradient boost machine

Introduction

Tuberculosis (TB), a chronic infectious disease caused by *Mycobacterium tuberculosis*, is a major global public health challenge and has become the leading cause of death among single infectious diseases.[1] The currently accepted first-line anti-TB drugs, including isoniazid, rifampin, pyrazinamide and ethambutol, play a central role in TB treatment and, when used in combination, can improve cure rates and reduce the development of drug resistance.[2–4] However, anti-tuberculosis drug-induced liver injury (ATB-DILI) has become the most common and serious adverse effect of TB treatment.[5] ATB-DILI can cause patients to discontinue treatment with single or combination anti-TB drugs, and discontinuation and rechallenge of anti-TB drugs can lead to the emergence of multidrug-resistant or even extensively drug-resistant strains, increasing the risk of treatment failure.[6] According to the World Health Organization's Global

Received: 10 September 2024
Accepted: 9 January 2025
Published: 13 January 2025

© 2025 Zeng et al. This work is published and licensed by Dove Medical Press Limited. The full terms of this license are available at https://www.dovepress.com/terms.php and incorporate the Creative Commons Attribution – Non Commercial (unported, v3.0) License (http://creativecommons.org/licenses/by-nc/3.0/). By accessing the work you hereby accept the Terms. Non-commercial uses of the work are permitted without any further permission from Dove Medical Press Limited, provided the work is properly attributed. For permission for commercial use of this work, please see paragraphs 4.2 and 5 of our Terms (https://www.dovepress.com/terms.php).

Tuberculosis Report 2022, there were approximately 10.6 million diagnosed TB cases worldwide,[7] with approximately 11% of all TB cases in children and adolescents under the age of 15.[8] Tuberculosis in children is increasingly recognized as an important component of the global tuberculosis burden, especially in resource-limited settings. The top five countries with the highest absolute burden were India, Nigeria, China, Indonesia and the Democratic Republic of Congo, and per capita childhood TB mortality was highest in countries in sub-Saharan Africa with high annual TB incidence.[9] Liver and kidney function are not fully developed in children, resulting in limited ability to metabolize drugs and increased risk of adverse reactions.[10] Given the significant impact of ATB-DILI on the treatment course and prognosis of pediatric TB patients, it is particularly important to identify and implement early preventive measures for high-risk patients.

The application of machine learning technology has shown great potential in the detection, diagnosis, prognosis and efficacy evaluation of diseases.[6] In the process of diagnosing and treating TB, a large amount of objective and quantifiable data has been accumulated through routine laboratory testing, including not only basic information such as blood counts and liver and kidney function tests, but also pharmacogenomics and drug concentration monitoring data.[11,12] Based on these diverse clinical data, predictive models built using machine learning techniques can quickly and accurately screen for high risk patients of ATB-DILI. A study using multiple logistic regression, support vector machine, and random forest to predict the risk of liver toxicity in patients treated with nilotinib, has demonstrated that the areas under the receiver operating characteristic curves (AUC) values of these machine learning models ranged from 0.61 to 0.65.[13] In addition, a nomogram prediction model established to risk of DILI based on the dataset of 1979 TB patients has shown that the AUC of the model in the training group, validation group I and validation group II were 0.833, 0.668 and 0.753, respectively.[14] Notably, several pieces of evidence suggest that machine learning algorithms facilitate the identification of different risks for children, such as pressure injuries in children undergoing living donor liver transplantation, chemotherapy-induced myelosuppression in children with Wilms' tumor and tacrolimus-induced tubular toxicity in children with nephrotic syndrome.[15–17]

Automated machine learning (AutoML) is an emerging approach to machine learning that automates the development of more accurate models. It is efficient and user-friendly to use, allowing cross-disciplinary experts without specialist knowledge of artificial intelligence to benefit from advances, including clinical doctors and pharmacists.[18] While few risk factors have been found to be associated with ATB-DILI in children,[19,20] to our knowledge, no study has used machine learning or automated machine learning methods to identify the high-risk children of ATB-DILI, due to the relatively low incidence and paucity of research data on ATB-DILI studies in children.

In summary, this study aims to build an ATB-DILI risk prediction model for children specifically designed for children with TB by integrating advanced machine learning techniques with rich clinical data resources, which will fill the current gaps in ATB-DILI risk assessment tools. Through the implementation of this study, not only can provide clinicians with a new tool to quickly and accurately assess the risk of ATB-DILI in children, but also is expected to provide strong support for optimizing the treatment plan of childhood tuberculosis, improve the treatment effect, reduce the incidence of adverse reactions, and achieve the goal of optimizing medical resources in the region with tuberculosis burden.

Materials and Methods

Inclusion and Exclusion of Patients

Study population: From June 2020 to June 2024, children diagnosed with TB and aged 1–17 years were recruited from the affiliated Changsha Central Hospital of University of South China. All children were treated with hepatoprotective drugs such as bicyclol, glutathione and compound diisopropylamine dichloroacetate.

The inclusion criteria were children: (1) receiving initial treatment with anti-tuberculosis drugs; (2) being treated with 3 to 4 first-line anti-tuberculosis drugs [isoniazid + rifampicin + pyrazinamide (or + ethambutol)] for the intensification period, while accepting isoniazid + rifampicin treatment for the consolidation period;[10] (3) whose plasma peak concentrations (C_{max}) of isoniazid and rifampicin (2h) were monitored for 1 week to 1 month after treatment with anti-

tuberculosis drugs; (4) whose biochemical indicators such as liver and kidney function were remeasured within 2 to 3 months of intensive treatment.

The exclusion criteria were children: (1) not treated with the standard regimen of first-line antituberculosis drugs [isoniazid + rifampicin + pyrazinamide (or + ethambutol)]; (2) whose clinical data were incomplete; (3) who had hepatic dysfunction or combined viral hepatitis, steatotic liver disease, Epstein-Barr virus infection, and cytomegalovirus infection prior to enrollment; (4) whose isoniazid and rifampicin C_{max} (2h) was assessed after the onset of drug-induced liver injury.

Diagnosis of DILI

The diagnosis of DILI is based on guidelines published by the Chinese Medical Association and literature.[21–24] The criteria for the diagnosis of DILI are detailed in Table S1.

Data Collection

The baseline clinical data of the children were collected before the initial treatment with anti-TB drugs. Demographic and clinical data mainly included sex, age, height, weight, allergy history, tuberculosis exposure, disease diagnosis, liver function, kidney function, and blood routine. The C_{max} of isoniazid and rifampicin (2h) was measured after the drugs had reached steady state and before the onset of liver injury. The C_{max} of pyrazinamide was not available for most children and was therefore not collected in the study. Finally, information on the combination of drugs, including hepatotoxic drugs, proton pump inhibitors, steroids, and antibiotics, during anti-TB treatment was collected for all children.

Factor Identification

The multi-collinearity among variables was meticulously assessed by the variance inflation factor (VIF). Univariate analysis was used to screen for significant variables between the DILI group and the control group ($P < 0.05$). We identified optimal cut-off points based on Youden index of the receiver operating characteristic (ROC) curve derived from univariate logistic analysis, and then these cut-off points were used to transform continuous variables into categorical variables.[25] The optimal cut-off point associated with the Youden index was defined as the maximum value of (sensitivity + specificity −1), and it was often used as the diagnostic threshold in clinical studies.[26] For instance, the calculation process of cut-off point of the rifampicin C_{max} was display in Table S2. The variables that showed significant differences ($P < 0.05$) between the two groups were used to establish machine learning models.

Construction of Predictive Models

Compared with traditional statistical methodologies, machine learning algorithms are capable to capture multifaceted nonlinear relationships and complex interactions between variables, and have advantages on handling high-dimensional and complex data sources beyond numerical sources.[18,27] AutoML, an efficient and easy-to-deploy solution, was chosen to develop the predictive models in this study.

All patients were randomly divided into a training set (80%) and a testing set (20%), which were used to construct the prediction model and validate the model's performance, respectively.[28] The AutoML technology was used to establish predictive models by the H2O AutoML platform.[29] This platform automates the process of training and tuning a large selection of candidate models. The H2O AutoML platform incorporates six primary machine learning algorithms, which include generalized linear models (GLM), distributed random forest (DRF), gradient boosting machines (GBM), fully connected deep neural network (DeepLearning), eXtremely randomized trees, stacked ensemble, and XGBoost. The main metrics for model performance included AUC, logloss, area under the precision-recall curve (AUCPR), mean_per_class_error, root mean squared error (RMSE), and mean squared error (MSE). A "leaderboard" was generated, ranking the candidate models based on their performance scores obtained from 5-fold cross-validation, which facilitated the filtering out of the best predictive models. The 5-fold cross-validation divided the training dataset into 5 sets, iterated the training process on 4-folds, and validated the model on the remaining 1 set, ensuring that the model performs well on unseen data and minimizing the risk of overfitting and underfitting.[30] Based on the model's performance score of cross-validation, H2O AutoML automated the hyperparameter tuning by grid search improving the robustness of the selected model.[31]

The main methods for addressing class-imbalance data at both the data level and algorithm level include random under-sampling, random over-sampling, loss functions, cost-sensitive learning, and threshold moving.[32] The H2O AutoML platform allows the user to select the appropriate decision thresholds based on criteria such as max F1, max F2, max accuracy and max precision. Further details are available on the web at https://docs.h2o.ai/h2o/latest-stable/h2o-docs/index.html.

In this study, the H2O open-source package (version 3.46.0.4) was installed via Python (version 3.10.9), and then implemented by Jupyter Notebook.

Model Explainability

To enhance the interpretability of machine learning model, we utilized a combination of analytical tools: variable importance plot and variable importance heatmap for visualizing the importance of variables, and SHAP Summary and local interpretable model-agnostic explanations (LIME) for understanding the impact of each variable on model predictions. SHAP analysis can identify key features that significantly affect model predictions and assesses the extent to which these features contribute to overall model performance. LIME is able to show the contribution of each feature to the prediction for a selected example.[33]

Statistical Analysis

Continuous variables were analyzed using the *t*-test, while categorical variables were analyzed using the chi-squared (χ^2) test or Fisher's exact test. Statistical analyses were performed using GraphPad Prism 9 (GraphPad Software, Inc., San Diego, CA, USA). A p-value of less than 0.05 was considered statistically significant.

Results

Patient Characteristics and Factor Identification

The study initially enrolled a total of 426 children who were treated with anti-tuberculosis drugs at Changsha Central Hospital from June 2020 to June 2024 and whose C_{max} of isoniazid and rifampicin had been measured. A total of 184 children with TB were included in the study, while a total of 242 children with TB were excluded from the study. The excluded cases included 74 patients receiving a second-line antituberculosis regimen, including linezolid, cycloserine, levofloxacin, moxifloxacin, streptomycin sulfate, rifapentin, and others, 103 patients with a single hospitalization record, 46 patients with incomplete clinical data, and 19 children under one year of age. Of the cases, 19 cases (10.33%) developed DILI following the administration of anti-tuberculosis therapy, including 17 cases classified as grade 1 (mild liver injury) and 2 cases classified as grade 2 (moderate liver injury), respectively.

All 24 variables, including demographic, clinical characteristics, and medication information are listed in Table 1. The VIF for all variables was less than 10, indicating an acceptable level of multi-collinearity. The sole continuous variable that showed a statistically significant correlation with DILI was the C_{max} of rifampicin. Therefore, following the conversion of the continuous variable into categorical variable using the Youden's index, it was demonstrated that aspartate aminotransferase (AST), alanine aminotransferase (ALT), total bile acids, creatinine, body mass index (BMI), total bilirubin and the C_{max} of rifampicin were significantly associated with ATB-DILI ($P < 0.05$). The C_{max} of isoniazid ($P = 0.0508$) and diagnosis of tuberculous meningitis ($P = 0.0528$) were identified as potential factors influencing the development of ATB-DILI. However, other variables, including sex, age, allergy history, etiology, co-medication information, blood routine and renal function data, were not found to be significantly different between the DILI and No-DILI groups. That is, categorical variables including BMI, total bilirubin, AST, ALT, total bile acids, creatinine, and C_{max} of rifampicin were used to build the machine learning model in the study.

Development and Validation of Predictive Model

Numerous machine learning models were automatically constructed utilizing the H2O AutoML platform. Thereafter, the performances of these models were rigorously evaluated and systematically ranked based on the outcomes of cross-validation. The top 20 candidate models with details of their scores are shown in Table 2.

Table 1 Comparison of the Characteristics of Patients Using Univariate Analysis

Characteristics	Total (n = 184)	No-DILI (n = 165)	DILI (n =19)	*P*-value
Age (years), median (IQR)	12 (6–15)	12 (6–15)	12 (5–15)	0.6910
Age ≥ 6 years[#], n (%)	147 (79.89)	134 (81.21)	13 (68.42)	0.2248[a]
Sex (male/female)	94/90	83/83	11/8	0.5307[b]
BMI, median (IQR)	16.85 (15.09–18.77)	16.96 (15.23–18.82)	16.05 (14.61–18.28)	0.2124
BMI ≥ 15.06[#], n (%)	139 (75.54)	129 (78.18)	10 (52.63)	0.0224[a]
History of drug allergy	30 (16.22)	29 (17.58)	1 (5.26)	0.3208[a]
Pathogeny (Essential/Secondary)	50/134	43/122	7/12	0.3171[b]
Tuberculous meningitis, n (%)	9 (4.86)	6 (3.64)	3 (15.79)	0.0528
Albumin (g/L), median (IQR)	40 (37–42)	40 (37–42)	40 (35–44)	0.6795
Albumin ≥ 29 g/L[#], n (%)	182 (98.38)	163 (98.79)	19 (100.00)	NA
Total bilirubin (μmol/L), median (IQR)	8.90 (6.23–12.58)	8.94 (6.31–12.48)	7.70 (5.53–13.07)	0.7307
Total bilirubin ≥ 21.2 μmol/L[#], n (%)	9 (4.89)	5 (3.03)	4 (21.05)	0.0075[a]
Direct bilirubin (μmol/L), median (IQR)	4.23 (2.90–5.92)	4.20 (2.92–5.90)	4.50 (2.00–6.68)	0.9829
Direct bilirubin ≥ 6.68 μmol/L[#], n (%)	30 (16.30)	25 (15.15)	5 (26.32)	0.2037[a]
Alanine aminotransferase (U/L), median (IQR)	10 (7–14)	10 (7–14)	13 (7–17)	0.4340
Alanine aminotransferase ≥ 13 U/L[#], n (%)	59 (32.07)	49 (29.70)	10 (52.63)	0.0425[b]
Aspartate aminotransferase (U/L), median (IQR)	22 (17–27)	22 (17–27)	23 (19–37)	0.2968
Aspartate aminotransferase ≥ 19 U/L[#], n (%)	126 (68.48)	109 (66.06)	17 (89.47)	0.0375[b]
Total biliary acid (μmol/L), median (IQR)	5.20 (3.10–8.07)	5.20 (3.16–7.94)	4.70 (2.60–11.10)	0.9865
Total biliary acid ≥ 19.32 μmol/L[#], n (%)	12 (6.52)	8 (4.85)	4 (21.05)	0.0238[a]
Creatinine (μmol/L), median (IQR)	37.5 (28.0–48.0)	38.0 (28.0–48.0)	34.0 (28.0–51.0)	0.9829
Creatinine ≥ 61 μmol/L[#], n (%)	14 (7.61)	10 (6.06)	4 (21.05)	0.0419[a]
Urea nitrogen (mmol/L), median (IQR)	3.75 (3.09–4.47)	3.75 (3.07–4.62)	3.66 (3.11–4.24)	0.7580
Urea nitrogen ≥ 4.69 mmol/L[#], n (%)	41 (22.28)	39 (23.64)	2 (10.53)	0.2532[a]
White blood cell (10^9/L), median (IQR)	6.73 (5.68–8.73)	6.72 (5.59–8.71)	6.73 (5.76–9.42)	0.5264
White blood cell ≥ 5.70*10^9/L[#], n (%)	137 (74.46)	120 (72.73)	17 (89.47)	0.1644[a]
Red blood cell (10^9/L), median (IQR)	4.51 (4.20–4.80)	4.51 (4.19–4.81)	4.51 (4.23–4.76)	0.8558
Red blood cell ≥ 4.50*10^9/L[#], n (%)	100 (54.35)	88 (53.33)	12 (63.16)	0.4156[b]
Hemoglobin (g/L), median (IQR)	122 (112–134)	122 (112–133)	128 (111–136)	0.4138
Hemoglobin ≥ 121 g/L[#], n (%)	99 (53.80)	85 (51.52)	14 (73.68)	0.0664[b]
Blood platelet (10^9/L), median (IQR)	332.5 (272.8–394.0)	332.0 (275.5–393.0)	350.0 (259.0–418.0)	0.8153
Blood platelet ≥ 413*10^9/L[#], n (%)	38 (20.65)	32 (19.39)	6 (31.58)	0.2339[a]
Combination with hepatotoxic drugs, n (%)	75 (40.76)	69 (41.82)	6 (31.59)	0.3897[b]
Combination with proton pump inhibitor, n (%)	3 (1.63)	2 (1.21)	1 (5.26)	0.2803[a]
Combination with steroids, n (%)	1 (0.54)	1 (0.61)	0 (0.00)	NA
Combination with antibiotics, n (%)	111 (60.33)	101 (61.21)	10 (52.63)	0.4691[b]
C_{max} of isoniazid (μg/mL), median (IQR)	6.17 (4.45–8.58)	6.09 (4.33–8.45)	7.24 (5.18–11.05)	0.1344
C_{max} of isoniazid ≥ 8.29 μg/mL[#], n (%)	52 (28.26)	43 (26.06)	9 (47.37)	0.0508[b]
C_{max} of rifampicin (μg/mL), median (IQR)	8.22 (5.88–11.43)	8.10 (5.84–10.81)	9.96 (6.46–15.04)	0.0280
C_{max} of rifampicin ≥ 13.48 μg/mL[#], n (%)	26 (14.13)	17 (10.30)	9 (47.37)	0.0002[a]

Notes: [#], the cut-off value determined by the Youden index; [a], Fisher's exact test; [b], Chi-squared test.
Abbreviations: IQR, interquartile range; C_{max}, plasma peak concentration.

We discovered that the top 3 machine learning models on the leaderboard are all GBM models. These models possess higher AUC scores and lower logloss values, indicating superior performance in classification accuracy and probabilistic prediction quality, respectively. Among these machine learning models, the optimal GBM model (Model_ID: GBM_lr_annealing_selection_AutoML_1_20240904_124545_select_model) has the highest AUC score. This model was selected for systematic evaluation and in-depth analysis as it particularly excelled in terming of balance between error rate and prediction accuracy. The AUC, AUCPR, logloss, and mean per-class error for the optimal GBM model on both the training set and testing set are displayed in Table 3. The optimal GBM model showed an AUC of 0.838 for the

Table 2 The Leaderboard of Machine Learning Models Ranked by Cross-Validated Performance

Model_id	AUC	logloss	AUCPR	mean_per_class_error	rmse	mse
GBM_lr_annealing_selection_AutoML_1_20240904_124545_select_model	0.776	0.286	0.306	0.308	0.286	0.082
GBM_grid_1_AutoML_1_20240904_124545_model_18	0.773	0.287	0.304	0.308	0.287	0.082
GBM_grid_1_AutoML_1_20240904_124545_model_15	0.762	0.285	0.315	0.327	0.284	0.081
StackedEnsemble_BestOfFamily_4_AutoML_1_20240904_124545	0.761	0.395	0.228	0.275	0.319	0.101
DeepLearning_grid_1_AutoML_1_20240904_124545_model_8	0.754	0.342	0.229	0.334	0.308	0.095
GLM_1_AutoML_1_20240904_124545	0.754	0.290	0.360	0.242	0.286	0.082
GBM_grid_1_AutoML_1_20240904_124545_model_7	0.740	0.287	0.415	0.345	0.281	0.079
GBM_grid_1_AutoML_1_20240904_124545_model_12	0.721	0.303	0.271	0.348	0.287	0.083
StackedEnsemble_BestOfFamily_3_AutoML_1_20240904_124545	0.715	0.318	0.214	0.252	0.303	0.092
GBM_grid_1_AutoML_1_20240904_124545_model_9	0.713	0.298	0.317	0.312	0.287	0.082
GBM_grid_1_AutoML_1_20240904_124545_model_2	0.713	0.294	0.293	0.323	0.283	0.080
GBM_grid_1_AutoML_1_20240904_124545_model_10	0.710	0.299	0.342	0.308	0.288	0.083
StackedEnsemble_AllModels_3_AutoML_1_20240904_124545	0.710	0.316	0.182	0.317	0.303	0.092
GBM_grid_1_AutoML_1_20240904_124545_model_22	0.706	0.299	0.256	0.316	0.291	0.084
StackedEnsemble_BestOfFamily_1_AutoML_1_20240904_124545	0.702	0.318	0.254	0.330	0.300	0.090
GBM_grid_1_AutoML_1_20240904_124545_model_21	0.696	0.319	0.249	0.356	0.298	0.089
GBM_grid_1_AutoML_1_20240904_124545_model_3	0.694	0.310	0.259	0.298	0.294	0.087
GBM_2_AutoML_1_20240904_124545	0.694	0.302	0.285	0.316	0.287	0.082
XRT_1_AutoML_1_20240904_124545	0.691	0.352	0.264	0.408	0.312	0.097
StackedEnsemble_AllModels_4_AutoML_1_20240904_124545	0.688	0.311	0.195	0.310	0.298	0.089

Abbreviations: AUC, area under the curve; logloss, logarithmic loss; AUCPR, area under the curve for precision–recall; mean_per_class_error, mean per-class error; rmse, root mean squared error; mse, mean squared error; GBM, gradient boost machine; GLM, generalized linear models; XRT, extremely randomized trees.

Table 3 The Classification Performance of Optimal GBM Model

	AUC	AUCPR	Logloss	Mean Per-Class Error
Training set	0.838	0.476	0.252	0.252
Testing set	0.784	0.635	0.264	0.250

Abbreviations: AUC, area under the curve; logloss, logarithmic loss; AUCPR, area under the curve for precision–recall; mean_per_class_error, mean per-class error.

training set (Figure 1A) and 0.784 for the test set (Figure 1B). In this model, at the decision threshold corresponding to max F2, the accuracy, sensitivity, and specificity were 0.844, 0.733, and 0.856, respectively.

Model Interpretation

As shown in Figure 2A, the variables are ranked by their relative importance within the optimal GBM model. The C_{max} of rifampicin emerges as the pivotal variable, followed by BMI, total bilirubin, creatinine, ALT, total bile acid, and AST. Furthermore, a variable importance heatmap was utilized to visualize the significance of variables across multiple models (Figure 2B). The C_{max} of rifampicin was identified as the most crucial feature across multiple GBM, XRT, and GLM models. Additionally, BMI and ALT were recognized as the secondarily significant variables within certain GBM models.

The TreeSHAP algorithm, a specialized version of SHAP designed for tree-based machine learning models, was employed to elucidate the model predictions.[34] The SHAP summary plot, generated by the optimal GBM, is illustrated in Figure 3A. The feature ranking of the SHAP plot showed that the C_{max} of rifampicin, BMI, and ALT were the top three features associated with the development of ATB-DILI in children. In addition, creatinine, AST, total bilirubin, and total bile acid were also correlated with the development of ATB-DILI. In the study, the input variables for machine learning are binary variables, with "1" corresponding to greater than or equal to the cut-off point and "0" corresponding to less than the cut-off point. We found that C_{max} of rifampicin ≥13.8 μg/mL corresponded to red points with positive SHAP values, indicating that children with higher C_{max} of rifampicin were more susceptible to ATB-DILI. Similarly, we found

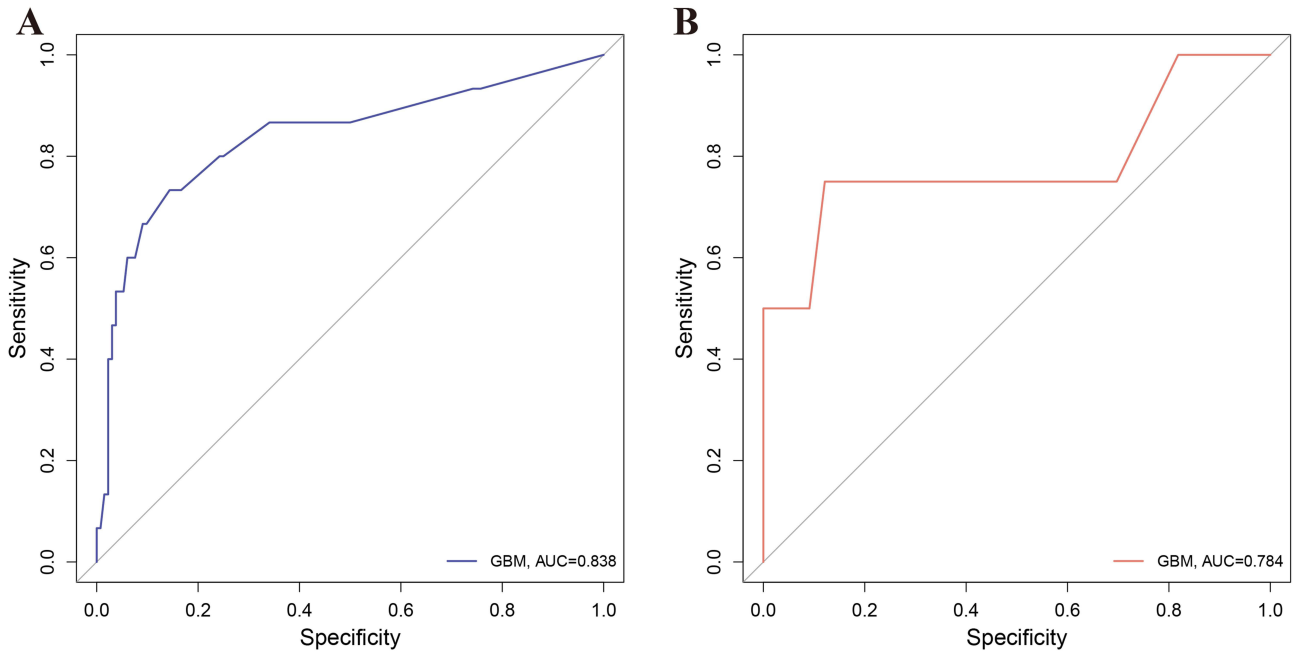

Figure 1 ROC curves for the training set and the testing set of the optimal GBM model. (**A**) Training set, (**B**) Testing set.

that most red dots corresponding to BMI ≥ 15.06 had negative SHAP values, suggesting that children with higher BMI values were less likely to develop ATB-DILI.

The LIME plot was shown in Figure 3B, which can illustrate the specific impact of each variable on the prediction of anti-tuberculosis DILI for a given patient. For example, for the pediatric patient (No. 35), who was predicted to develop ATB-DILI, the local explanation graph confirmed the critical role of BMI and C_{max} of rifampicin as predictive factors. Variables such as creatinine, total bile acids, and total bilirubin had lower SHAP values, indicating a lesser impact on the risk prediction for this patient.

Discussion

ATB-DILI is the leading cause of treatment interruption in tuberculosis patients. Predicting patients at high-risk for ATB-DILI enables clinicians to proactively adjust medication protocols and promptly institute preventive measures. In this study, the AutoML was executed to identify the optimal model for predicting ATB-DILI. To our knowledge, this is the first time that a machine learning model has been applied to the prediction of ATB-DILI in children.

The results of univariate analysis revealed that C_{max} of rifampicin, BMI, total bilirubin, ALT, AST, total bile acids and creatinine were significant risk factors for ATB-DILI in children. These important predictors were utilized in the construction of the machine learning models. Our findings revealed that C_{max} of rifampicin ≥ 13.48 μg/mL was positively associated with the risk of ATB-DILI. However, a study reported a cut-off value of 12.50 μg/mL for rifampicin-induced hepatotoxicity in adults.[35] A pharmacokinetics and safety/tolerability study demonstrated that children who experienced DILI had higher $AUC_{0\text{-}24}$ and C_{max} levels of rifampicin on day 10 compared to children without DILI.[36] The mechanism of rifampicin-induced liver injury may involve triggering endoplasmic reticulum stress through various pathways, such as the accumulation of bile acids and the generation of toxic drug metabolites by cytochrome p450 enzymes, processes regulated by the pregnane X receptor.[37] We found that BMI ≥ 15.06 was negatively associated with ATB-DILI risk in children. Previous report showed that the malnourished children have a higher risk of ATB-DILI,[38] which may be attributed to reduced clearance rates of xenobiotics.[39] A study proposed that BMI < 18.5 is an independent risk factor for ATB-DILI in children, with 2.3-fold higher incidence than in normal children.[10] However, another study found that nutritional status was not associated with ATB-DILI in children,[19] which may be attributed to the fact that the study only included children in the intensification period. In this study, creatinine ≥ 61 μmol/L was positively associated with the

https://doi.org/10.2147/DDDT.S495555

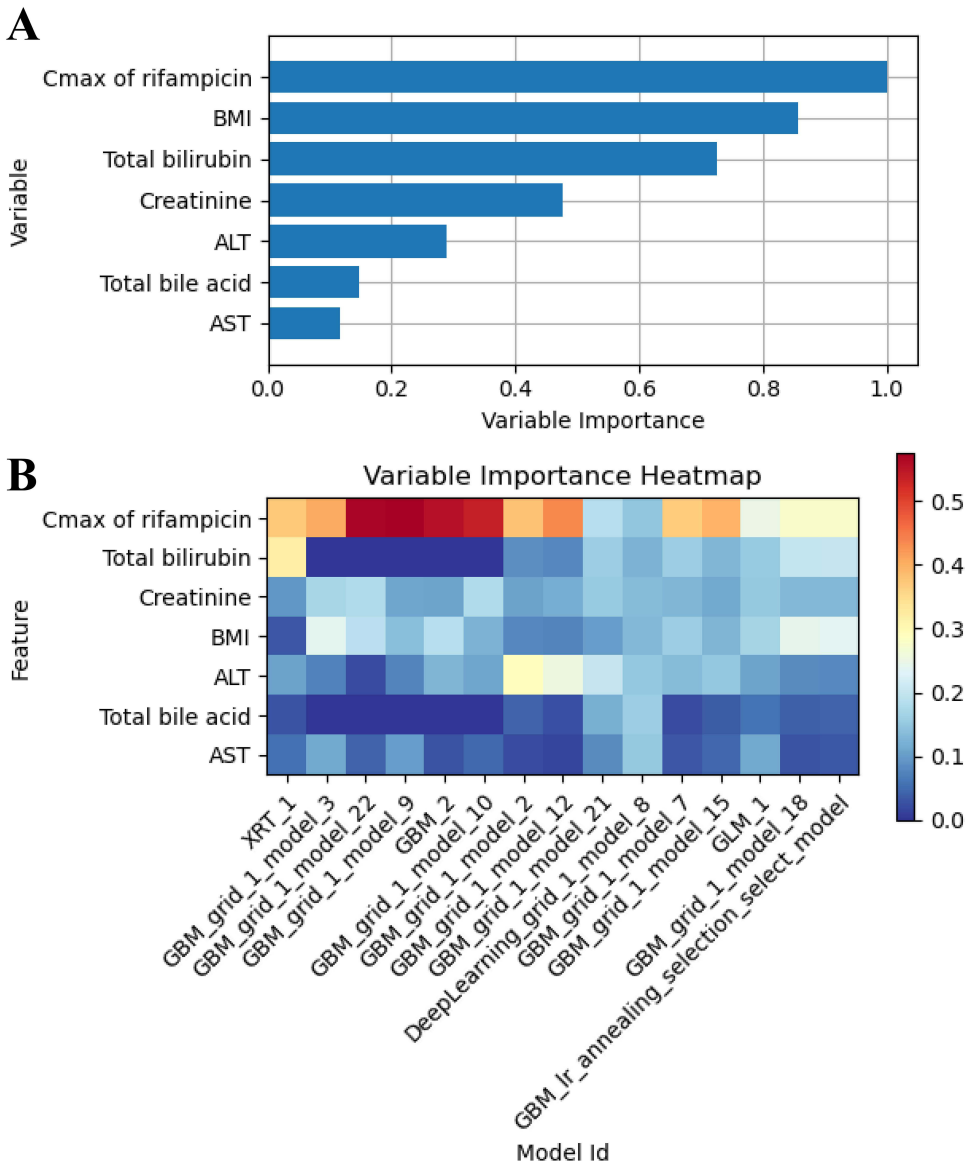

Figure 2 The variable importance plot and variable importance heatmap. (**A**) Importance score of variables in optimal GBM model, (**B**) Variable importance across multiple models.

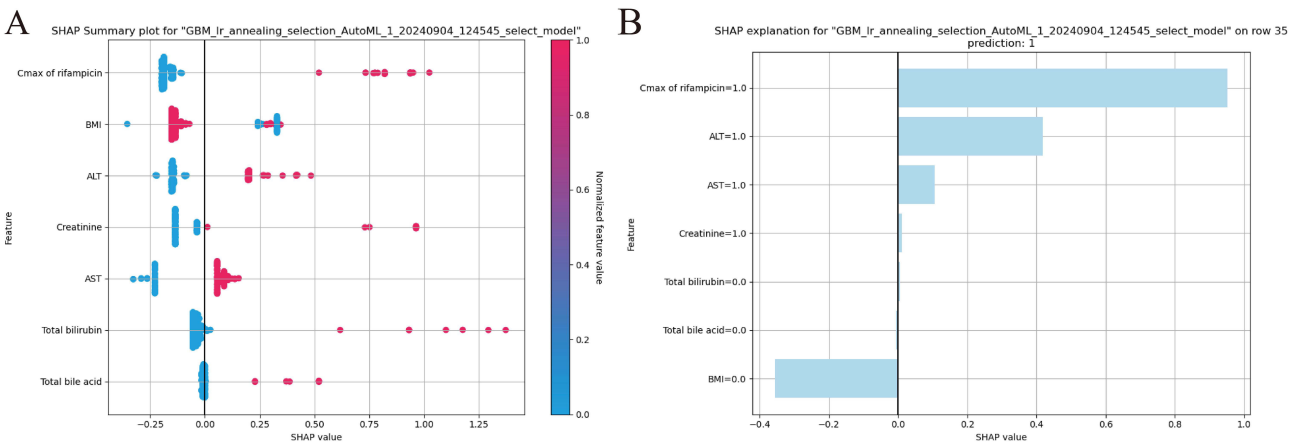

Figure 3 The SHAP Explanation of optimal GBM model in the training set. (**A**) SHAP summary plot, (**B**) The local explanation of SHAP.

 https://doi.org/10.2147/DDDT.S495555

risk of DILI. The decline in renal function may lead to increased drug concentration in the liver, potentially increasing the risk of hepatotoxicity.[40,41] Baseline total bilirubin, total bile acids, ALT and AST are all critical indicators of liver function. Several studies have confirmed that higher levels of these markers are significant risk factors for the development of ATB-DILI.[14,42–44] Elevated levels of these markers might indicate underlying hepatic metabolic disturbances, that predispose patients to ATB-DILI upon initiation of anti-tuberculosis treatment.

Currently, several studies have employed machine learning models to predict DILI. At the molecular level, a study combined transcriptomic profiles of human cell lines with a deep neural network model to predict DILI, which achieved an AUC of 0.802 and 0.798 for training and an independent validation set, respectively.[45] In the clinical stage, a time-series deep learning model to predict the risk of DILI in patients taking angiotensin receptor blockers, that demonstrated excellent classification performance for telmisartan, losartan and irbesartan, achieving AUCs of 0.93, 0.92 and 0.90, respectively.[46] Another study proposed a decision tree model for predicting the risk of carbapenem-induced liver injury. The model suggested that doripenem might pose a higher risk in patients with ALT > 22 IU/L and albumin-bilirubin score > −1.87.[47] In our study, the prediction models for ATB-DILI were developed using AutoML technology. The optimal model, selected through 5-fold cross-validation, demonstrated excellent predictive performance, achieving an AUC of 0.838 and 0.734 on the training set and testing set, respectively. Therefore, the risk of ATB-DILI in children can be accurately predicted by an optimal GBM model, enabling clinicians to develop personalized treatment regimens for these patients that significantly reduce the probability of ATB-DILI occurrence. In addition, the AutoML technology employed in this study is more accessible to experts from non-machine learning fields to facilitate the development of high-quality machine learning models. Furthermore, the AutoML technology employed in this study is more accessible to experts from non-machine learning fields to facilitate the development of high-quality machine learning models.

At present, the Tuberculosis Branch of the Chinese Medical Association has formulated guidelines for the diagnosis and treatment of ATB-DILI. The diagnostic criteria mainly involve comprehensive medical history collection, liver biochemical tests, and routine abdominal imaging should be performed for ATB-DILI patients, as well as liver histopathological examination.[48] Nevertheless, our predictive model established by clinical data and TDM test data can timely and accurately identify high-risk children of ATB-DILI. This will change the clinical management pattern from passive to active, benefiting pediatricians and children. This proactive approach allows healthcare professionals to implement timely interventions and mitigation strategies, thereby minimizing potential harm to high-risk children with ATB-DILI, such as reducing the physiological burden and discomfort compared to invasive detection methods. In addition, the high-risk children of ATB-DILI can be identified by the machine learning model so that tuberculosis-endemic and resource-limited regions can optimize the allocation of medical resources and reduce their medical burden, improve the healthcare of high-risk children of ATB-DILI, and promote the establishment of a sustainable medical security system.

This study had a few limitations. Firstly, the data utilized to construct the machine learning model were sourced exclusively from a single medical center. Secondly, although the plasma concentrations of isoniazid and rifampicin were covered, medication information on the plasma concentration of pyrazinamide was missing because this testing program was not performed in our hospital at the early stage. In future research, we will collaborate with multiple centers to obtain external validation data to ensure the model's generalization ability. Additionally, the pyrazinamide and genetic data will be incorporated to enhance the predictive performance of the machine learning model.

Conclusions

This study is the first application of AutoML technology on predicting the risk of ATB-DILI in children. Based on seven key features, a series of predictive models were constructed using AutoML technology. Among these models, the GBM model significantly outperformed others in predicting ATB-DILI, showing superior performance with AUCs of 0.838 and 0.784 on the training and testing sets respectively. The predictive model can provide early warning signals before ATB-DILI occurs and generate interpretable information, which could potentially be developed into a professional application to help clinicians adjust the medication plan in time and potentially reduce the likelihood of ATB-DILI in children. In addition, this study will facilitate the optimization of medical resource allocation and alleviate the medical burden in tuberculosis-endemic and resource-limited areas, and it provides new perspectives and methodological guidance for the

study of other adverse effects of anti-tuberculosis drugs. Due to the single-center nature of this study, data sources are limited. In future research, genomics information will be integrated to enhance model performance, and the case will be expanded through multi-center collaboration.

Data Sharing Statement

The machine learning model generated during the current study are available from the corresponding author by the Email address: 2022050025@usc.edu.cn

Ethics Approval

This study has been approved by the Research Ethics Committee of the University of South China (approval number: 2024-107) and were performed according to the Declaration of Helsinki's ethical principles. Owing to the retrospective nature of the study, informed consent from participants was waived.

Funding

This work was supported by Changsha Central Hospital Affiliated to University of South China Foundation of key Program (YNKY202305 and YNKY202204), National Natural Science Foundation of China (82003870), Natural Science Foundation of Changsha (kq2014012), Scientific Research Project of Education Department of Hunan Province (23B0438), and Scientific and technological innovation project of Hunan Province (2021SK53407).

Disclosure

The authors declare no conflicts of interest in this work.

References

1. Zhong T, Zhuang Z, Dong X, et al. Predicting antituberculosis drug–induced liver injury using an interpretable machine learning method: model development and validation study. *JMIR Med Inform*. 2021;9(7):e29226. doi:10.2196/29226
2. Isa SE, Ebonyi AO, Shehu NY, et al. Antituberculosis drugs and hepatotoxicity among hospitalized patients in Jos, Nigeria. *Int J Mycobact*. 2016;5 (1):21–26. doi:10.1016/j.ijmyco.2015.10.001
3. Shakya R, Rao BS, Shrestha B. Incidence of hepatotoxicity due to antitubercular medicines and assessment of risk factors. *Ann Pharmacother*. 2004;38(6):1074–1079. doi:10.1345/aph.1D525
4. Dheda K, Barry CE 3rd, Maartens G. Tuberculosis. *Lancet*. 2016;387(10024):1211–1216. doi:10.1016/S0140-6736(15)00151-8
5. Zhong T, Fan Y, Dong X-L, et al. An investigation of the risk factors associated with anti-tuberculosis drug-induced liver injury or abnormal liver functioning in 757 patients with pulmonary tuberculosis. *Front Pharmacol*. 2021;12:708522. doi:10.3389/fphar.2021.708522
6. Lai N-H, Shen W-C, Lee C-N, et al. Comparison of the predictive outcomes for anti-tuberculosis drug-induced hepatotoxicity by different machine learning techniques. *Comput Methods Programs Biomed*. 2020;188:105307. doi:10.1016/j.cmpb.2019.105307
7. Huang C-K, Huang J-Y, Chang C-H, et al. The effect of statins on the risk of anti-tuberculosis drug-induced liver injury among patients with active tuberculosis: a cohort study. *J Microbiol Immunol Infect*. 2024;57(3):498–508. doi:10.1016/j.jmii.2024.04.002
8. WHO Guidelines Approved by the Guidelines Review Committee. *WHO Consolidated Guidelines on Tuberculosis: Module 5: Management of Tuberculosis in Children and Adolescents*. World Health Organization; 2022.
9. Dodd PJ, Yuen CM, Sismanidis C, Seddon JA, Jenkins HE. The global burden of tuberculosis mortality in children: a mathematical modelling study. *Lancet Glob Health*. 2017;5(9):e898–e906. doi:10.1016/s2214-109x(17)30289-9
10. Chen F, Zhang X, Zhou H, Zhang F, Wang M. Analysis of status and influencing factors associated with anti-tuberculosis drug-related liver injury in children. *Chin J Antituberculosis*. 2023;45(1):45–51. doi:10.19982/j.issn.1000-6621.20220290
11. Mei YM, Zhang WY, Sun JY, et al. Genomic characteristics of mycobacterium tuberculosis isolates of cutaneous tuberculosis. *Front Microbiol*. 2023;14:1165916. doi:10.3389/fmicb.2023.1165916
12. Peloquin CA, Davies GR. The treatment of tuberculosis. *Clin Pharmacol Ther*. 2021;110(6):1455–1466. doi:10.1002/cpt.2261
13. Kim JS, Han JM, Cho YS, Choi KH, Gwak HS. Machine learning approaches to predict hepatotoxicity risk in patients receiving nilotinib. *Molecules*. 2021;26(11):3300. doi:10.3390/molecules26113300
14. Ji S, Lu B, Pan X. A nomogram model to predict the risk of drug-induced liver injury in patients receiving anti-tuberculosis treatment. *Front Pharmacol*. 2023;14:1153815. doi:10.3389/fphar.2023.1153815
15. Li M, Wang Q, Lu P, et al. Development of a machine learning-based prediction model for chemotherapy-induced myelosuppression in children with wilms' tumor. *Cancers*. 2023;15(4):1078. doi:10.3390/cancers15041078
16. Mo X, Chen X, Ieong C, et al. Early prediction of tacrolimus-induced tubular toxicity in pediatric refractory nephrotic syndrome using machine learning. *Front Pharmacol*. 2021;12:638724. doi:10.3389/fphar.2021.638724
17. Chen X, Tang S, Qin Y, et al. A predictive model of pressure injury in children undergoing living donor liver transplantation based on machine learning algorithm. *J Adv Nurs*. 2024:1–10. doi:10.1111/jan.16449
18. SenthilKumar G, Madhusudhana S, Flitcroft M, et al. Automated machine learning (AutoML) can predict 90-day mortality after gastrectomy for cancer. *Sci Rep*. 2023;13(1):11051. doi:10.1038/s41598-023-37396-3

19. Gafar F, Arifin H, Jurnalis YD, et al. Antituberculosis drug-induced liver injury in children: incidence and risk factors during the two-month intensive phase of therapy. *Pediatr Infect Dis J*. 2019;38(1):50–53. doi:10.1097/inf.0000000000002192
20. Mehra A, Semwal P, Bhat NK, Bolia R. A prospective observational study of hepatic dysfunction in children on antitubercular drugs. *Indian J Pediatr*. 2022;89(11):1126–1128. doi:10.1007/s12098-022-04317-7
21. Yu Y-C, Mao Y-M, Chen C-W, et al. CSH guidelines for the diagnosis and treatment of drug-induced liver injury. *Hepatol Int*. 2017;11:221–241. doi:10.1007/s12072-017-9793-2
22. Lu H, Zeng Y, Shi Q-Z, et al. Low albumin combined with low-molecular-weight heparin as risk factors for liver injury using azvudine: evidence from an analysis of COVID-19 patients in a national prospective pharmacovigilance database. *Int J Clin Pharmacol Ther*. 2024;62(5):222–228. doi:10.5414/CP204544
23. Xiao X, Tang J, Mao Y, et al. Guidance for the clinical evaluation of traditional Chinese medicine-induced liver injuryIssued by China food and drug administration. *Acta Pharm Sin B*. 2019;9(3):648–658. doi:10.1016/j.apsb.2018.12.003
24. Association CM, House CMJP, Gastroenterology CSo, et al. Guideline for primary care of drug-induced liver injury: practice version(2019).. *Chin J Gen Pract*. 2020;19(10):876–881. doi:10.3760/cma.j.cn114798-20200812-00901
25. Imai S, Takekuma Y, Kashiwagi H, et al. Validation of the usefulness of artificial neural networks for risk prediction of adverse drug reactions used for individual patients in clinical practice. *PLoS One*. 2020;15(7):e0236789. doi:10.1371/journal.pone.0236789
26. Nakas CT, Alonzo TA, Yiannoutsos CT. Accuracy and cut-off point selection in three-class classification problems using a generalization of the youden index. *Stat Med*. 2010;29(28):2946–2955. doi:10.1002/sim.4044
27. Wong JE, Yamaguchi M, Nishi N, Araki M, Wee LH. Predicting overweight and obesity status among Malaysian working adults with machine learning or logistic regression: retrospective comparison study. *JMIR Form Res*. 2022;6(12):e40404. doi:10.2196/40404
28. Maray I, Rodríguez-Ferreras A, Álvarez-Asteinza C, Alaguero-Calero M, Valledor P, Fernández J. Linezolid induced thrombocytopenia in critically ill patients: risk factors and development of a machine learning-based prediction model. *J Infect Chemother*. 2022;28(9):1249–1254. doi:10.1016/j.jiac.2022.05.004
29. LeDell E, Poirier S. H2o automl: scalable automatic machine learning. In: *ICML*. San Diego, CA, USA; 2020.
30. Park Y. Automated machine learning with R: autoML tools for beginners in clinical research. *J Minim Invasive Surg*. 2024;27(3):129–137. doi:10.7602/jmis.2024.27.3.129
31. Jiao Z, Zhang X, Xuan Y, et al. Leveraging cfDNA fragmentomic features in a stacked ensemble model for early detection of esophageal squamous cell carcinoma. *Cell Rep Med*. 2024;5(8):101664. doi:10.1016/j.xcrm.2024.101664
32. Johnson JM, Khoshgoftaar TM. Survey on deep learning with class imbalance. *J Big Data*. 2019;6(1):27. doi:10.1186/s40537-019-0192-5
33. Liu L, Zhang R, Shi D, et al. Automated machine learning to predict the difficulty for endoscopic resection of gastric gastrointestinal stromal tumor. *Front Oncol*. 2023;13:1190987. doi:10.3389/fonc.2023.1190987
34. Bifarin OO, Fernández FM. Automated machine learning and explainable AI (AutoML-XAI) for metabolomics: improving cancer diagnostics. *J Am Soc Mass Spectrom*. 2024;35(6):1089–1100. doi:10.1021/jasms.3c00403
35. Satyaraddi A, Velpandian T, Sharma SK, et al. Correlation of plasma anti-tuberculosis drug levels with subsequent development of hepatotoxicity. *Int J Tuberc Lung Dis*. 2014;18(2):188–195. doi:10.5588/ijtld.13.0128
36. Ruslami R, Gafar F, Yunivita V, et al. Pharmacokinetics and safety/tolerability of isoniazid, rifampicin and pyrazinamide in children and adolescents treated for tuberculous meningitis. *Arch Dis Child*. 2022;107(1):70–77. doi:10.1136/archdischild-2020-321426
37. Hou W, Nsengimana B, Yan C, Nashan B, Han S. Involvement of endoplasmic reticulum stress in rifampicin-induced liver injury. *Front Pharmacol*. 2022;13:1022809. doi:10.3389/fphar.2022.1022809
38. Gaude GS, Chaudhury A, Hattiholi J. Drug-induced hepatitis and the risk factors for liver injury in pulmonary tuberculosis patients. *J Family Med Prim Care*. 2015;4(2):238–243. doi:10.4103/2249-4863.154661
39. Tostmann A, Boeree MJ, Aarnoutse RE, De Lange WC, Der Ven AJ V, Dekhuijzen R. Antituberculosis drug-induced hepatotoxicity: concise up-to-date review. *J Gastroenterol Hepatol*. 2008;23(2):192–202. doi:10.1111/j.1440-1746.2007.05207.x
40. Sun H, Frassetto L, Benet LZ. Effects of renal failure on drug transport and metabolism. *Pharmacol Ther*. 2006;109(1–2):1–11. doi:10.1016/j.pharmthera.2005.05.010
41. Soedarsono S, Riadi ARW. Tuberculosis drug-induced liver injury. *J Respirasi*. 2020;6(2):49–54. doi:10.20473/jr.v6-I.2.2020.49-54
42. Jiang F, Yan H, Liang L, et al. Incidence and risk factors of anti-tuberculosis drug induced liver injury (Dili): large cohort study involving 4652 Chinese adult tuberculosis patients. *Liver Int*. 2021;41(7):1565–1575. doi:10.1111/liv.14896
43. Liu Q, Huang L, Yan H, et al. Clinical risk factors for moderate and severe antituberculosis drug-induced liver injury. *Front Pharmacol*. 2024;15:1406454. doi:10.3389/fphar.2024.1406454
44. Zhang J, Zhou W, Ma S, et al. Combined electronic medical records and gene polymorphism characteristics to establish an anti-tuberculosis drug-induced hepatic injury (ATDH) prediction model and evaluate the prediction value. *Ann Transl Med*. 2022;10(20):1114. doi:10.21037/atm-22-4551
45. Li T, Tong W, Roberts R, Liu Z, Thakkar S. Deep learning on high-throughput transcriptomics to predict drug-induced liver injury. *Front Bioeng Biotechnol*. 2020;8:562677. doi:10.3389/fbioe.2020.562677
46. Heo S, Yu JY, Kang EA, et al. Development and verification of time-series deep learning for drug-induced liver injury detection in patients taking angiotensin ii receptor blockers: a multicenter distributed research network approach. *Healthc Inform Res*. 2023;29(3):246–255. doi:10.4258/hir.2023.29.3.246
47. Asai Y, Ooi H, Sato Y. Risk evaluation of carbapenem-induced liver injury based on machine learning analysis. *J Infect Chemother*. 2023;29(7):660–666. doi:10.1016/j.jiac.2023.03.007
48. Branch CMAT. Guidelines for diagnosis and management of drug-induced liver injury caused by anti-tuberculosis drugs (2024 version). *Chin J Tuberc Respir Dis*. 2024;47(11):1069–1090. doi:10.3760/cma.j.cn112147-20240614-00338

Drug Design, Development and Therapy

Dovepress
Taylor & Francis Group

Publish your work in this journal

Drug Design, Development and Therapy is an international, peer-reviewed open-access journal that spans the spectrum of drug design and development through to clinical applications. Clinical outcomes, patient safety, and programs for the development and effective, safe, and sustained use of medicines are a feature of the journal, which has also been accepted for indexing on PubMed Central. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit http://www.dovepress.com/testimonials.php to read real quotes from published authors.

Submit your manuscript here: https://www.dovepress.com/drug-design-development-and-therapy-journal