Journal of Clinical and Translational Hepatology **2026** vol. 14(8) | 785–798
DOI: 10.14218/JCTH.2026.00447

Original Article

Early Prediction of Non-recovery in Drug-induced Liver Injury by Integrating Genetic Variants and Clinical Variables: An Interpretable Machine Learning Approach

Jingjing Jiang[#], Weiwei Lou[#], Qing Li, Ziqiang Li, Weiqian Lou, Xichen Zhu, Qing Xie[*] and Rongtao Lai[*]

Department of Infectious Diseases, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Received: June 05, 2026 | **Revised:** June 18, 2026 | **Accepted:** July 02, 2026 | **Published online:** August 05, 2026

Abstract

Background and Aims: Early predictors of 6-month non-recovery in drug-induced liver injury (DILI) remain limited. Genetic variants are stable host characteristics that may complement baseline clinical variables. We aimed to develop and validate an interpretable, clinical-genetic machine learning model for predicting 6-month non-recovery in patients with DILI. ***Methods:*** This retrospective, single-center study included 338 patients with DILI, who were classified as recovered (n = 171) or non-recovered (n = 167) at 6 months. Candidate single-nucleotide polymorphisms and baseline clinical variables were collected during initial hospitalization. Features were selected using complementary screening approaches. Multiple machine learning models were developed and compared. Model discrimination, calibration, clinical utility, the incremental value of genetic predictors, and interpretability using SHapley Additive exPlanations (SHAP) were assessed. ***Results:*** Five predictors were consistently retained for model development: rs72631567, rs28521457, alanine aminotransferase, monocyte percentage, and low-density lipoprotein. Among the candidate algorithms, the light gradient boosting machine model showed the best performance, with area under the receiver operating characteristic curve (AUC) values of 0.92 (95% confidence interval [CI] 0.89–0.95) in the training set and 0.81 (95% CI 0.70–0.91) in the validation set. The model showed acceptable calibration and favorable decision-curve performance. In ablation analysis, the clinical-only model showed limited discrimination (AUC 0.57, 95% CI 0.43–0.71). SHAP analysis identified rs72631567 as the most influential predictor. ***Conclusions:*** An interpretable model that integrates host genetic variants with baseline clinical variables demonstrated good internal performance for early prediction of 6-month non-recovery in DILI. These findings support external validation of genotype-informed risk stratification in patients with DILI.

Keywords: Chemical and drug-induced liver injury; Single-nucleotide polymorphism; Machine learning; Predictive model; Chronic disease; Prognosis.

[#]Contributed equally to this work.
[*]**Correspondence to:** Rongtao Lai and Qing Xie, Department of Infectious Diseases, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China. ORCID: https://orcid.org/0000-0003-1704-7086 (RL) and https://orcid.org/0000-0002-2582-8803 (QX). Tel: +86-21-64370045 (RL) and +86-21-64370045 (QX). E-mail: lairongtao1202@163.com (RL) and xieqingrjh@163.com (QX).

Citation of this article: Jiang J, Lou W, Li Q, Li Z, Lou W, Zhu X, *et al*. Early Prediction of Non-recovery in Drug-induced Liver Injury by Integrating Genetic Variants and Clinical Variables: An Interpretable Machine Learning Approach. J Clin Transl Hepatol 2026;14(8):785–798. doi: 10.14218/JCTH.2026.00447.

Introduction

Drug-induced liver injury (DILI) is a major challenge in hepatology and remains an important cause of abnormal liver biochemistry, post-marketing drug withdrawal, and acute liver failure.[1,2] Its clinical course is highly heterogeneous. Although most patients recover after withdrawal of the implicated agent, a substantial proportion develop persistent liver test abnormalities beyond 6 months, a phenotype commonly regarded as DILI chronicity.[3,4] Some patients who do not recover may remain at risk of prolonged morbidity and progressive liver disease.[2] Early identification of individuals at increased risk of non-recovery could therefore help guide the intensity of follow-up and clinical management.

Current reliable prediction tools for non-recovery after DILI are limited and have largely relied on clinical laboratory parameters, imaging features, or selected biomarkers.[5–7] However, these phenotypic markers are dynamic and may be influenced by the timing of assessment, disease stage, treatment, and comorbid conditions. Due to their relatively stable nature, genetic variants may provide complementary information for early risk prediction. In other disease settings, incorporating genetic information has improved the performance of clinical prediction models.[8,9]

Genetic studies of DILI have mainly focused on susceptibility rather than disease course.[10–12] In a large, international, multicenter genome-wide association study, Nicoletti *et al*[13] identified three non-HLA genetic signals (HLA, human leukocyte antigen) associated with DILI susceptibility: rs114577328, rs72631567, and rs28521457. Whether these variants are also associated with subsequent non-recovery after DILI onset remains unclear.

In the present study, we examined the associations of these candidate single-nucleotide polymorphisms (SNPs) with 6-month non-recovery in a real-world cohort. We then developed an interpretable machine learning model that integrated genetic variants with baseline clinical variables avail-

able during the initial hospitalization. We aimed to improve early prediction of non-recovery and provide a framework for future genotype-informed risk stratification of DILI.

Methods

Study design and patient selection

This retrospective, single-center cohort study was conducted at Ruijin Hospital, Shanghai, China. Patients hospitalized with suspected DILI between May 2018 and May 2023 were consecutively screened. The study was approved by the institutional ethics committee of Ruijin Hospital (Ethics Approval No. 11-2017) and conducted in accordance with the principles outlined in the Declaration of Helsinki.

Patients were eligible if they met all of the following criteria: age ≥18 years, documented exposure to a suspected drug, fulfillment of biochemical criteria for acute DILI, a Roussel Uclaf Causality Assessment Method (RUCAM) score ≥6, and available follow-up data for at least 6 months after DILI diagnosis. The biochemical criteria for acute DILI were defined as alanine aminotransferase (ALT) ≥5 × the upper limit of normal (ULN), alkaline phosphatase (ALP) ≥2 × ULN with bone disease excluded, or ALT ≥3 × ULN together with total bilirubin (TBil) ≥2 × ULN.[14] Patients were excluded if they were lost to follow-up or if liver injury could be attributed to other definite causes, including viral hepatitis, autoimmune hepatitis, liver tumors, obstructive jaundice, or liver injury secondary to systemic disease. The final cohort comprised 338 patients, including 171 in the recovery group and 167 in the non-recovery group.

Outcome definition

The primary outcome was 6-month non-recovery after DILI diagnosis, representing DILI chronicity in this study.[15] Non-recovery DILI was defined as persistent abnormalities in any of the following liver biochemical indices at 6 months after DILI diagnosis: ALT, AST, ALP, or TBil.

Clinical variable collection

Baseline clinical data were extracted from electronic medical records within the first 24 h of hospitalization. Candidate variables included demographic characteristics, lifestyle factors, comorbidity burden, biochemical indices, lipid and iron metabolism markers, coagulation parameters, hematologic indices, autoantibodies, immunoglobulin levels, and lymphocyte subsets. The collected variables included sex, age, BMI, smoking status, alcohol consumption history, Charlson Comorbidity Index (CCI), fasting glucose, ALT, AST, ALP, gamma-glutamyl transferase (GGT), TBil, direct bilirubin (DBil), total bile acids (TBA), albumin (ALB), prealbumin (PAB), creatinine (Cr), triglycerides (TG), total cholesterol (TC), high-density lipoprotein (HDL), low-density lipoprotein (LDL), free fatty acids (FFA), serum iron (SI), total iron-binding capacity (TIBC), ferritin (SF), transferrin saturation (TSAT), alpha-fetoprotein (AFP), prothrombin time (PT), activated partial thromboplastin time (APTT), international normalized ratio (INR), D-dimer, white blood cell (WBC) count, monocyte percentage (Mono), eosinophil percentage (Eos), hemoglobin (Hb), platelet (PLT) count, antinuclear antibody (ANA), immunoglobulin G (IgG), and lymphocyte subsets including $CD3^+$, $CD3^+CD4^+$, and $CD3^+CD8^+$ T cells. Culprit drugs were also recorded and descriptively categorized.

The pattern of liver injury was classified according to the R value as hepatocellular (R ≥ 5), mixed (2 < R < 5), or cholestatic (R ≤ 2). DILI severity was graded from Grade 1 to Grade 4 according to criteria established by the International DILI Expert Working Group.[16] Grade 1 DILI was defined as ALT ≥5 × ULN or ALP ≥2 × ULN with TBil <2 × ULN. Grade 2 DILI was defined as ALT ≥5 × ULN or ALP ≥2 × ULN with TBil ≥2 × ULN, or symptomatic hepatitis. Grade 3 DILI was defined as ALT ≥5 × ULN or ALP ≥2 × ULN with TBil ≥2 × ULN and at least one of the following: INR ≥1.5, ascites or hepatic encephalopathy within 26 weeks in the absence of cirrhosis, or other organ failure attributable to DILI. Grade 4 DILI was defined as death or liver transplantation due to DILI.

SNP genotyping

Three candidate SNPs, including rs114577328, rs72631567, and rs28521457, were selected from a previous genome-wide association study that identified host genetic variants associated with DILI susceptibility.[13] These loci were evaluated as candidate host genetic markers for association with 6-month recovery status after DILI onset. Genotyping was performed using Sanger sequencing. Genomic DNA was extracted from peripheral blood samples, and regions flanking each candidate locus were amplified by conventional PCR with locus-specific primers. The primer sequences are listed in Supplementary Table 1. Genotypes were assigned by comparing sequencing results with the corresponding reference sequences and identifying the nucleotide at each candidate SNP position.

Data preprocessing

The cohort was split into training and validation sets at an 8:2 ratio using stratified sampling according to non-recovery outcome status. After this split, feature selection, hyperparameter tuning, and model selection were performed exclusively on the training set, whereas the validation set was used solely for final model evaluation. Missing values were imputed using multiple imputation by chained equations (MICE) with predictive mean matching.[17] Five imputed datasets were generated with 10 iterations under a fixed random seed, and one completed dataset was used as the primary analytic dataset for model development and validation. SNPs were additively encoded as 0, 1, or 2 according to the number of effect alleles. Categorical variables were dummy-coded when required for model development.

Feature selection

To reduce redundancy and improve robustness, three complementary feature selection methods were applied to the training set: least absolute shrinkage and selection operator (LASSO),[18] support vector machine–recursive feature elimination (SVM-RFE),[19] and Boruta.[20] For LASSO, the optimal penalty parameter was selected using cross-validation according to the λ.1se criterion. For SVM-RFE, the optimal feature subset was defined as the set achieving the highest classification accuracy. Boruta, a random forest (RF)–based algorithm, was used to identify all relevant predictors by comparing observed variables with shadow features. Predictors consistently selected by all three methods were retained through intersection analysis for subsequent model construction.

Model development

Based on the selected predictors, six machine learning models were developed using the training set: Light Gradient Boosting Machine (LightGBM),[21] logistic regression (LR), RF,[22] support vector machine (SVM),[23] eXtreme Gradient Boosting (XGBoost),[24] and multilayer perceptron (MLP).[25] Model construction and hyperparameter tuning were performed using the "tidymodels" R package. Five-fold cross-validation was

conducted in the training set to optimize the model parameters and compare candidate models. The tuned models were then evaluated in the validation set.

Model performance evaluation

Model performance was assessed using receiver operating characteristic (ROC) curves and area under the curve (AUC). Additional metrics derived from confusion matrices included accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and F1 score. Decision curve analysis (DCA) was used to evaluate net clinical benefit across a range of threshold probabilities,[26] and calibration curves were used to assess agreement between predicted and observed probabilities. The final model was selected based on overall performance in discrimination, clinical utility, and calibration.

Ablation analysis for incremental predictive value of genetic variants

Ablation analyses were performed within the final LightGBM framework to evaluate the incremental predictive contribution of genetic variants. Five models were compared: a full model including rs72631567, rs28521457, ALT, Mono, and LDL; a genetic-only model including rs72631567 and rs28521457; a clinical-only model including ALT, Mono, and LDL; and two reduced models combining clinical variables with either rs72631567 or rs28521457. Model discrimination was assessed using AUC, and AUC differences were compared using DeLong's test.[27] In addition, log loss was calculated to evaluate the effect of removing individual genetic variables on probabilistic prediction performance.[28]

SHapley Additive exPlanations (SHAP)-based model interpretation

SHAP values were used to quantify the contribution of each predictor to model output.[29] Global model interpretation was performed using SHAP summary plots and mean absolute SHAP value rankings. SHAP dependence plots were used to evaluate relationships between individual predictor values and model output. Representative force plots were generated to illustrate how the final model produced predictions for individual recovery and non-recovery cases.

Statistical analysis

Baseline characteristics were summarized and compared between the recovery and non-recovery groups for descriptive analyses. Continuous variables were expressed as mean ± standard deviation or median (IQR), as appropriate. Categorical variables were expressed as counts and percentages. Normality was assessed using the Shapiro–Wilk test. Continuous variables were compared using Student's t-test or the Mann–Whitney U test, as appropriate, whereas categorical variables were analyzed using the chi-square test or Fisher's exact test, as appropriate. Baseline comparisons were performed for descriptive purposes and were not used as the sole basis for predictor selection. All statistical analyses were performed using R (version 4.5.1). All tests were two-sided, and $P < 0.05$ was considered statistically significant.

Results

Study cohort and baseline characteristics

Among the 430 patients with DILI initially screened, 338 admitted between 2018 and 2023 with a RUCAM score ≥6 were ultimately included in this study. The final cohort comprised 171 patients in the recovery group and 167 patients in the non-recovery group (Fig. 1). Herbal and dietary supplements constituted the largest category of culprit drugs (Supplementary Fig. 1).

Baseline characteristics are summarized in Table 1. Among the three candidate SNPs, rs28521457 ($P = 0.001$) and rs72631567 ($P < 0.001$) differed significantly between the groups, whereas rs114577328 showed no variation in this cohort and was excluded from subsequent analyses. Injury grade distribution also differed between the groups ($P = 0.035$). Patients with non-recovery DILI had lower baseline ALT levels ($P = 0.005$), lower LDL levels ($P = 0.023$), lower WBC counts ($P = 0.005$), and a higher Mono level ($P = 0.001$) than those with recovery DILI. No significant differences were observed in age, sex, BMI, clinical type, bilirubin-related indices, coagulation parameters, or most immune-related variables (all $P > 0.05$).

The cohort was split at an 8:2 ratio, stratified by non-recovery outcome, into a training set (n = 269) and a validation set (n = 69). Baseline characteristics were generally balanced between the two sets, except for the proportion of $CD3^+CD8^+$ T cells (Table 2). The distribution of missing values across candidate variables is summarized in Supplementary Tables 2 and 3.

Feature selection identified an integrated five-predictor panel

Boruta, SVM-RFE, and LASSO were applied to the training set for feature selection and retained 6, 8, and 10 candidate variables, respectively (Fig. 2A). Detailed outputs of the individual feature-selection procedures are provided in Supplementary Figure 2. Intersection analysis showed that rs72631567, rs28521457, ALT, Mono, and LDL were consistently selected by all three methods (Fig. 2B). These five variables were used as the final integrated predictors for subsequent model construction.

LightGBM demonstrated the best overall predictive performance among candidate algorithms

Using the five integrated predictors, six candidate algorithms (LightGBM, LR, RF, SVM, XGBoost, and MLP) were developed and compared. LightGBM achieved the highest discriminatory ability in both the training and validation sets, with AUCs of 0.92 (95% confidence interval (CI), 0.89–0.95) and 0.81 (95% CI, 0.70–0.91), respectively (Table 3; Fig. 3A and B). Although some models achieved comparable performance for individual metrics, LightGBM showed the most favorable overall balance. In the validation set, its accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and F1 score were 0.72, 0.79, 0.66, 0.69, 0.77, and 0.74, respectively (Table 3). Calibration analysis showed good agreement between predicted and observed probabilities for the LightGBM model (Fig. 3C). DCA showed that the LightGBM model provided a positive net benefit across a broad range of threshold probabilities in both the training and validation sets (Fig. 3D and E). In addition, the LightGBM model showed stable discrimination in cross-validation and good calibration in the training set (Supplementary Fig. 3).

Detailed hyperparameter settings and confusion matrices are provided in Supplementary Table 4 and Supplementary Figures 4–5. Given its superior overall performance, LightGBM was selected as the final model for further analyses.

Genetic variants provided incremental predictive value beyond clinical variables

To further assess the incremental predictive value of ge-

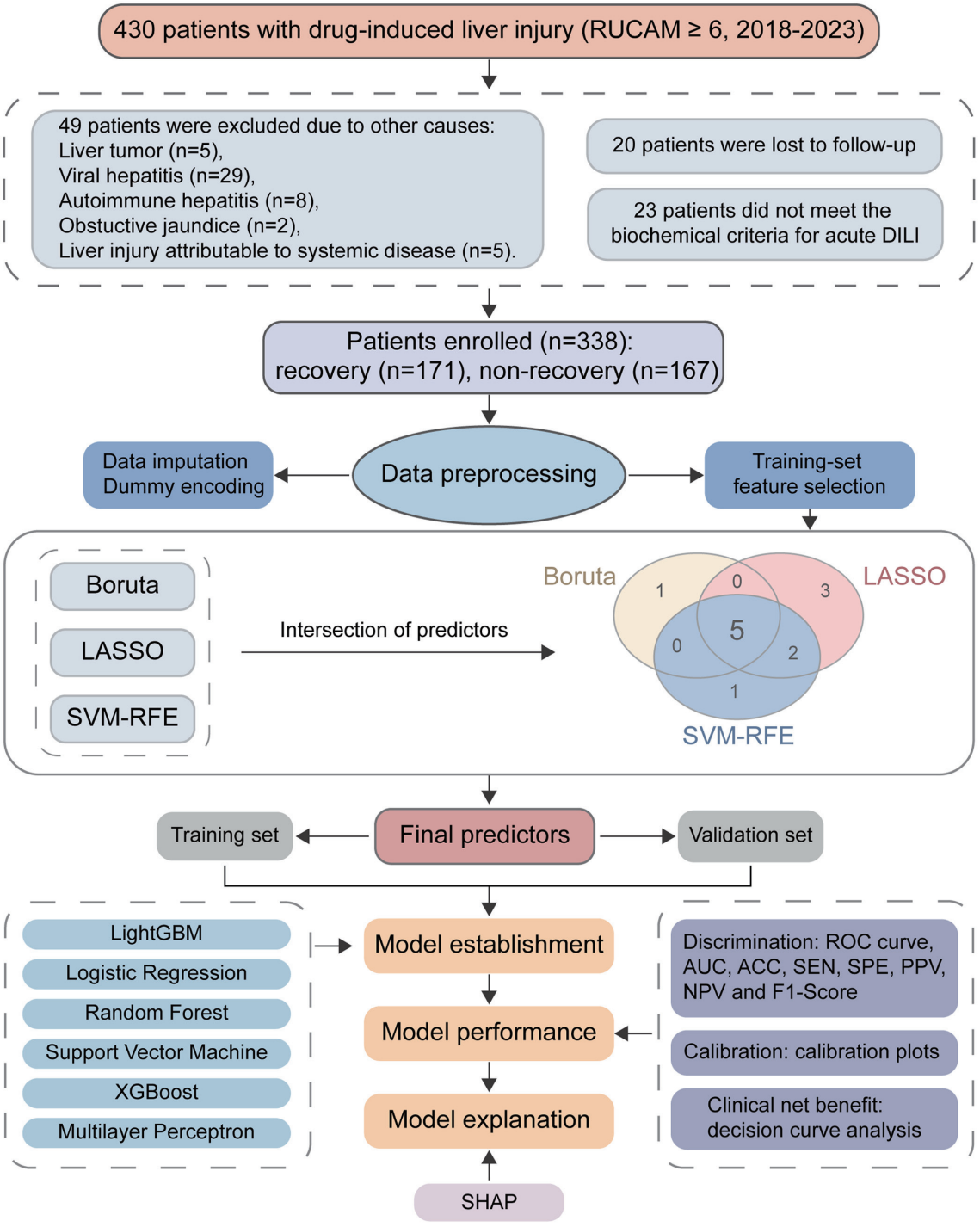

Fig. 1. Flowchart of the study. RUCAM, Roussel Uclaf Causality Assessment Method; LASSO, least absolute shrinkage and selection operator; SVM-RFE, support vector machine-recursive feature elimination; LightGBM, Light Gradient Boosting Machine; XGBoost, eXtreme Gradient Boosting; ROC, receiver operating characteristic; AUC, area under the curve; ACC, accuracy; SEN, sensitivity; SPE, specificity; PPV, positive predictive value; NPV, negative predictive value; SHAP, SHapley Additive exPlanations.

netic variants, we performed an ablation analysis within the LightGBM framework by comparing the full model with a genetic-only model, a clinical-only model, and two reduced models combining clinical variables with either rs72631567 or rs28521457 (Fig. 4A; Supplementary Table 5). In the validation set, the full model achieved the highest AUC (0.81, 95% CI 0.70–0.91), whereas the clinical-only model showed markedly poorer discrimination (AUC = 0.57, 95% CI 0.43–0.71). The genetic-only model retained good predictive ability (AUC = 0.79, 95% CI 0.69–0.90), approaching that of the full model. Consistently, DeLong's test showed that the full model significantly outperformed the clinical-only

Table 1. Baseline characteristics of patients with recovery and non-recovery DILI

Characteristics	Total (N = 338)	Recovery (N = 171)	Non-recovery (N = 167)	*P*-value
rs28521457, N (%)				0.001
GG	222 (66)	128 (75)	94 (56)	
AG	108 (32)	40 (23)	68 (41)	
AA	8 (2)	3 (2)	5 (3)	
rs72631567, N (%)				< 0.001
AA	151 (45)	105 (61)	46 (28)	
AG	152 (45)	64 (37)	88 (53)	
GG	35 (10)	2 (1)	33 (20)	
rs114577328, N (%)				1
GG	338 (100)	171 (100)	167 (100)	
Injury grade, N (%)				0.035
Grade 1	129 (38)	67 (39)	62 (37)	
Grade 2	182 (54)	84 (49)	98 (59)	
Grade 3	25 (7)	18 (11)	7 (4)	
Grade 4	2 (1)	2 (1)	0 (0)	
Gender, N (%)				0.395
Female	231 (68)	121 (71)	110 (66)	
Male	107 (32)	50 (29)	57 (34)	
Age (years)	55 (44.25, 64)	55 (43.5, 63)	55 (45, 65)	0.351
BMI (kg/m^2)	23.39 ± 3.28	23.42 ± 3.07	23.36 ± 3.49	0.867
Smoking, N (%)				0.08
NO	286 (85)	151 (88)	135 (81)	
YES	52 (15)	20 (12)	32 (19)	
Drinking, N (%)				0.081
NO	297 (88)	156 (91)	141 (84)	
YES	41 (12)	15 (9)	26 (16)	
CCI	2 (1, 4)	2 (1, 4)	2 (1, 3.5)	0.634
Clinical type, N (%)				0.48
Hepatocellular DILI	240 (71)	126 (74)	114 (68)	
Cholestatic DILI	55 (16)	24 (14)	31 (19)	
Mixed DILI	43 (13)	21 (12)	22 (13)	
Glucose (mmol/L)	4.81 (4.5, 5.51)	4.79 (4.47, 5.46)	4.82 (4.54, 5.54)	0.693
PAB (g/L)	112.5 (69.25, 171)	119 (70.5, 171)	108 (69, 171)	0.797
ALT (IU/L)	562 (380, 1,026.25)	671 (401.5, 1,199.5)	485 (372.5, 887.5)	0.005
AST (IU/L)	399.5 (172, 728.75)	408 (170, 749.5)	397 (173, 717.5)	0.439
ALP (IU/L)	155.5 (116, 222.5)	155 (123.5, 213)	156 (112.5, 237.5)	0.964
GGT (IU/L)	166 (92, 276.5)	161 (99, 261.5)	171 (85.5, 283)	0.566
TBil (μmol/L)	60.35 (21.33, 168.98)	66.8 (21.85, 148.05)	53.3 (20.5, 177.7)	0.75
DBil (μmol/L)	29.45 (6.3, 88.67)	31 (6.4, 80.85)	28.9 (5.75, 98.15)	0.985
ALB (g/L)	36 (32, 39)	36 (32, 39)	35 (32, 39)	0.476
TBA (μmol/L)	44.9 (12.3, 158.52)	41.8 (11.85, 156.75)	58 (14, 159.05)	0.45
Cr (μmol/L)	60 (54, 74.75)	60 (54, 71)	61 (54, 77)	0.313
TG (mmol/L)	1.54 (1.12, 2.26)	1.53 (1.15, 2.08)	1.55 (1.11, 2.37)	0.644

(continued)

Table 1. *(continued)*

Characteristics	Total (N = 338)	Recovery (N = 171)	Non-recovery (N = 167)	*P*-value
TC (mmol/L)	3.88 (3.12, 4.79)	3.97 (3.14, 4.84)	3.8 (3.09, 4.75)	0.424
HDL (mmol/L)	0.84 (0.51, 1.26)	0.86 (0.52, 1.23)	0.83 (0.48, 1.29)	0.823
LDL (mmol/L)	2.14 (1.46, 2.77)	2.33 (1.64, 2.79)	1.97 (1.07, 2.7)	0.023
FFA (mmol/L)	0.73 (0.52, 0.94)	0.76 (0.54, 0.99)	0.7 (0.51, 0.91)	0.125
SI (μmol/L)	27.1 (18.85, 38.55)	26 (18.75, 38.05)	28 (19, 39.15)	0.52
TSAT (%)	53.2 (36, 84.72)	53.4 (34.55, 82.4)	53 (36.6, 85.3)	0.535
TIBC (μmol/L)	51 (42.52, 58.1)	51 (42.95, 61.5)	51.1 (42.35, 56.9)	0.577
APTT (s)	31.8 (29.3, 34.5)	31.2 (28.9, 35.1)	32 (29.6, 34.35)	0.344
PT (s)	12.4 (11.6, 14)	12.3 (11.6, 14.4)	12.5 (11.7, 13.7)	0.612
INR	1.06 (0.99, 1.18)	1.04 (0.98, 1.25)	1.06 (1, 1.17)	0.667
D-dimer (mg/L FEU)	0.36 (0.25, 0.56)	0.36 (0.25, 0.65)	0.36 (0.25, 0.54)	0.264
ANA, N (%)				0.114
0	243 (72)	128 (75)	115 (69)	
1:80	48 (14)	19 (11)	29 (17)	
1:160	28 (8)	15 (9)	13 (8)	
1:320	15 (4)	5 (3)	10 (6)	
1:640	3 (1)	3 (2)	0 (0)	
1:1,280	1 (0)	1 (1)	0 (0)	
AFP (ng/mL)	5.62 (2.89, 16.13)	5.61 (2.89, 15.87)	5.64 (2.9, 16.28)	0.952
SF (ng/mL)	652.6 (263.38, 1,216.65)	678.9 (285.35, 1,374.95)	642.6 (239.65, 1,089.2)	0.232
WBC count ($\times 10^9$/L)	4.81 (3.83, 5.98)	5.13 (4, 6.38)	4.5 (3.53, 5.66)	0.005
Mono (%)	9 (6.8, 11.2)	8.1 (6.35, 10.6)	9.5 (7.5, 11.5)	0.001
Eos (%)	2.15 (0.83, 3.6)	2.1 (0.9, 4.05)	2.2 (0.8, 3.5)	0.348
Hb (g/L)	126 (118, 137)	126 (116.5, 137)	126 (119, 137.5)	0.724
PLT count ($\times 10^9$/L)	181 (137.25, 231)	182 (145, 236.5)	180 (127, 228)	0.223
IgG (mg/dL)	1,310 (1,110.25, 1,576)	1,310 (1,115.5, 1,654.5)	1,310 (1,108.5, 1,528)	0.29
$CD3^+$ T (%)	74.55 (66.8, 79.97)	74.8 (66.3, 81)	73.9 (67.3, 79.5)	0.702
$CD3^+CD4^+$ T (%)	40.37 ± 10.12	40.77 ± 9.97	39.96 ± 10.29	0.466
$CD3^+CD8^+$ T (%)	28.1 (22.22, 35.77)	28.3 (22.6, 35.95)	27.3 (21.75, 35)	0.599

Data are presented as mean ± standard deviation (SD), median (Interquartile range, IQR), or number (percentage). BMI, body mass index; CCI, Charlson Comorbidity Index; PAB, prealbumin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase; GGT, gamma-glutamyl transferase; TBil, total bilirubin; DBil, direct bilirubin; ALB, albumin; TBA, total bile acids; Cr, creatinine; TG, triglycerides; TC, total cholesterol; HDL, high-density lipoprotein; LDL, low-density lipoprotein; FFA, free fatty acids; SI, serum iron; TSAT, transferrin saturation; TIBC, total iron-binding capacity; APTT, activated partial thromboplastin time; PT, prothrombin time; INR, international normalized ratio; FEU: Fibrinogen Equivalent Units; ANA, antinuclear antibody; AFP, alpha-fetoprotein; SF, serum ferritin; WBC, white blood cell; Mono, monocyte percentage; Eos, eosinophil percentage; Hb, hemoglobin; PLT, platelet; IgG, immunoglobulin G.

model ($P = 1.8 \times 10^{-4}$), whereas no significant difference was observed between the full and genetic-only models ($P = 0.7965$) (Fig. 4B).

Detailed results of the reduced models are provided in Supplementary Table 5. Adding rs28521457 to the clinical variables increased the AUC from 0.57 to 0.65, whereas adding rs72631567 increased it from 0.57 to 0.76; inclusion of both variants further increased the AUC to 0.81. This pattern was consistent with the log loss analysis (Supplementary Table 6). The full model showed the lowest log loss (0.552), whereas omitting rs28521457 and rs72631567 increased the log loss to 0.608 and 0.662, respectively, again suggesting a greater contribution of rs72631567 to overall model performance.

DCA showed greater net benefit for the full and genetic-only models than for the clinical-only model across a broad range of threshold probabilities (Fig. 4C). Calibration analysis showed acceptable agreement between predicted and observed probabilities across models, with the full model maintaining a favorable balance between discrimination and calibration (Fig. 4D).

SHAP-based interpretation of the LightGBM model

To improve interpretability of the LightGBM model, SHAP analysis was performed to quantify each predictor's contribution to model output. The SHAP summary plot showed the effects of the predictors on model output (Fig. 5A and B). Features were ranked by mean absolute SHAP value, with rs72631567 being the most influential predictor, followed by

Table 2. Baseline characteristics of the training and validation sets

Characteristics	Training (n = 269)	Validation (n = 69)	*P*-value
Status, N (%)			1
Recovery	136 (51)	35 (51)	
Non-recovery	133 (49)	34 (49)	
rs28521457, N (%)			0.888
GG	177 (66)	45 (65)	
AG	86 (32)	22 (32)	
AA	6 (2)	2 (3)	
rs72631567, N (%)			0.201
AA	124 (46)	27 (39)	
AG	121 (45)	31 (45)	
GG	24 (9)	11 (16)	
Injury grade, N (%)			0.196
Grade 1	96 (36)	33 (48)	
Grade 2	152 (57)	30 (43)	
Grade 3	19 (7)	6 (9)	
Grade 4	2 (1)	0 (0)	
Gender, N (%)			0.631
Female	186 (69)	45 (65)	
Male	83 (31)	24 (35)	
Age (years)	55 (43, 64)	56 (49, 63)	0.332
BMI (kg/m^2)	23.31 ± 3.29	23.69 ± 3.24	0.387
Smoking, N (%)			0.146
NO	232 (86)	54 (78)	
YES	37 (14)	15 (22)	
Drinking, N (%)			0.088
NO	241 (90)	56 (81)	
YES	28 (10)	13 (19)	
CCI	2 (1, 4)	3 (1, 4)	0.23
Clinical type, N (%)			0.202
Hepatocellular DILI	196 (73)	44 (64)	
Cholestatic DILI	39 (14)	16 (23)	
Mixed DILI	34 (13)	9 (13)	
Glucose (mmol/L)	4.79 (4.49, 5.42)	4.95 (4.52, 6.68)	0.1
PAB (g/L)	112 (71, 166)	119 (62, 185)	0.705
ALT (IU/L)	594 (392, 1,086)	496 (349, 771)	0.053
AST (IU/L)	413 (188, 730)	301 (129, 658)	0.065
ALP (IU/L)	152 (113, 220)	168 (130, 242)	0.147
GGT (IU/L)	161 (90, 275)	176 (104, 277)	0.315
TBil (μmol/L)	66.2 (22.1, 160.5)	50.1 (16.8, 181.4)	0.444
DBil (μmol/L)	32.2 (6.6, 85.3)	13.4 (5.4, 94.8)	0.483
ALB (g/L)	36 (32, 39)	36 (32, 38)	0.478
TBA (μmol/L)	44.9 (13.3, 156.9)	45.4 (9, 164.9)	0.509

(continued)

Table 2. *(continued)*

Characteristics	Training (n = 269)	Validation (n = 69)	*P*-value
Cr (μmol/L)	60 (54, 73)	64 (54, 80)	0.085
TG (mmol/L)	1.53 (1.11, 2.23)	1.66 (1.23, 2.34)	0.315
TC (mmol/L)	3.87 (3.15, 4.78)	3.97 (3.01, 4.79)	0.905
HDL (mmol/L)	0.84 (0.54, 1.27)	0.77 (0.4, 1.15)	0.181
LDL (mmol/L)	2.14 (1.42, 2.78)	2.09 (1.59, 2.7)	0.859
FFA (mmol/L)	0.75 (0.52, 0.94)	0.68 (0.49, 0.95)	0.497
SI (μmol/L)	26.5 (18.2, 38.8)	29.3 (19.2, 35.2)	0.937
TSAT (%)	52.5 (35.6, 82.6)	57 (39.7, 88.8)	0.347
TIBC (μmol/L)	51.2 (44.4, 57.8)	49.4 (39.1, 58.2)	0.235
APTT (s)	31.8 (29.5, 34.3)	31.7 (28.7, 36)	0.833
PT (s)	12.3 (11.5, 13.9)	12.8 (11.8, 14.6)	0.203
INR	1.05 (0.98, 1.18)	1.09 (1, 1.25)	0.218
D-dimer (mg/L FEU)	0.35 (0.25, 0.53)	0.4 (0.28, 0.79)	0.139
ANA, N (%)			0.579
0	193 (72)	50 (72)	
1:80	41 (15)	7 (10)	
1:160	21 (8)	7 (10)	
1:320	10 (4)	5 (7)	
1:640	3 (1)	0 (0)	
1:1,280	1 (0)	0 (0)	
AFP (ng/mL)	5.61 (2.89, 13.08)	5.64 (2.81, 18.84)	0.231
SF (ng/mL)	615.4 (255.2, 1,189.6)	785.9 (352, 1,272.5)	0.099
WBC count ($\times10^9$/L)	4.72 (3.73, 5.89)	5.21 (4.06, 6.6)	0.123
Mono (%)	9 (6.7, 11.1)	8.4 (7.1, 11.4)	0.838
Eos (%)	2.1 (0.8, 3.6)	2.2 (0.9, 3.5)	0.773
Hb (g/L)	127 (119, 137)	124 (115, 137)	0.39
PLT count ($\times10^9$/L)	181 (136, 229)	180 (140, 239)	0.56
IgG (mg/dL)	1,292 (1,110, 1,524)	1,330 (1,138, 1,665)	0.229
$CD3^+$ T (%)	74.8 (66.8, 80.5)	73.6 (65.7, 79.4)	0.365
$CD3^+CD4^+$ T (%)	39.86 ± 9.79	42.35 ± 11.18	0.094
$CD3^+CD8^+$ T (%)	29.2 (22.5, 36.1)	24.8 (21.1, 30.4)	0.004

Data are presented as mean ± standard deviation (SD), median (Interquartile range, IQR), or number (percentage). BMI, body mass index; CCI, Charlson Comorbidity Index; PAB, prealbumin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase; GGT, gamma-glutamyl transferase; TBil, total bilirubin; DBil, direct bilirubin; ALB, albumin; TBA, total bile acids; Cr, creatinine; TG, triglycerides; TC, total cholesterol; HDL, high-density lipoprotein; LDL, low-density lipoprotein; FFA, free fatty acids; SI, serum iron; TSAT, transferrin saturation; TIBC, total iron-binding capacity; APTT, activated partial thromboplastin time; PT, prothrombin time; INR, international normalized ratio; FEU: Fibrinogen Equivalent Units; ANA, antinuclear antibody; AFP, alpha-fetoprotein; SF, serum ferritin; WBC, white blood cell; Mono, monocyte percentage; Eos, eosinophil percentage; Hb, hemoglobin; PLT, platelet; IgG, immunoglobulin G; DILI, drug-induced liver injury.

ALT, Mono, LDL, and rs28521457.

SHAP dependence plots further illustrated the relationships between individual feature values and model output (Supplementary Fig. 6). For both rs72631567 and rs28521457, the presence of effect alleles was associated with higher SHAP values, with a clearer allele-dose-dependent pattern observed for rs72631567. In contrast, ALT and LDL levels showed inverse associations with model output, with lower values contributing positively to the prediction of non-recovery DILI. Mono showed a positive association with SHAP values, with higher values corresponding to a higher predicted probability of non-recovery DILI.

To further illustrate individual predictions, SHAP force plots were generated for two representative patients. In a representative recovery case, the absence of effect alleles at rs72631567 and rs28521457, together with relatively high ALT, higher LDL, and lower Mono levels, contributed negatively to model output, thereby reducing the predicted probability of non-recovery DILI (Fig. 5C). In contrast, in a representative non-recovery case, carriage of one effect allele at rs72631567 and two effect alleles at rs28521457, together with lower ALT and higher Mono, contributed positively to the model output and jointly increased the predicted probability of non-recovery DILI (Fig. 5D).

A

Methods	Predictor variables
Boruta	rs72631567, rs28521457, ALT, Mono, LDL, Injury grade
SVM-RFE	rs72631567, rs28521457, ALT, Mono, LDL, WBC count, D-Dimer, Hb
LASSO	rs72631567, rs28521457, ALT, Mono, LDL, WBC count, Eos, Hb, IgG, Smoking

B

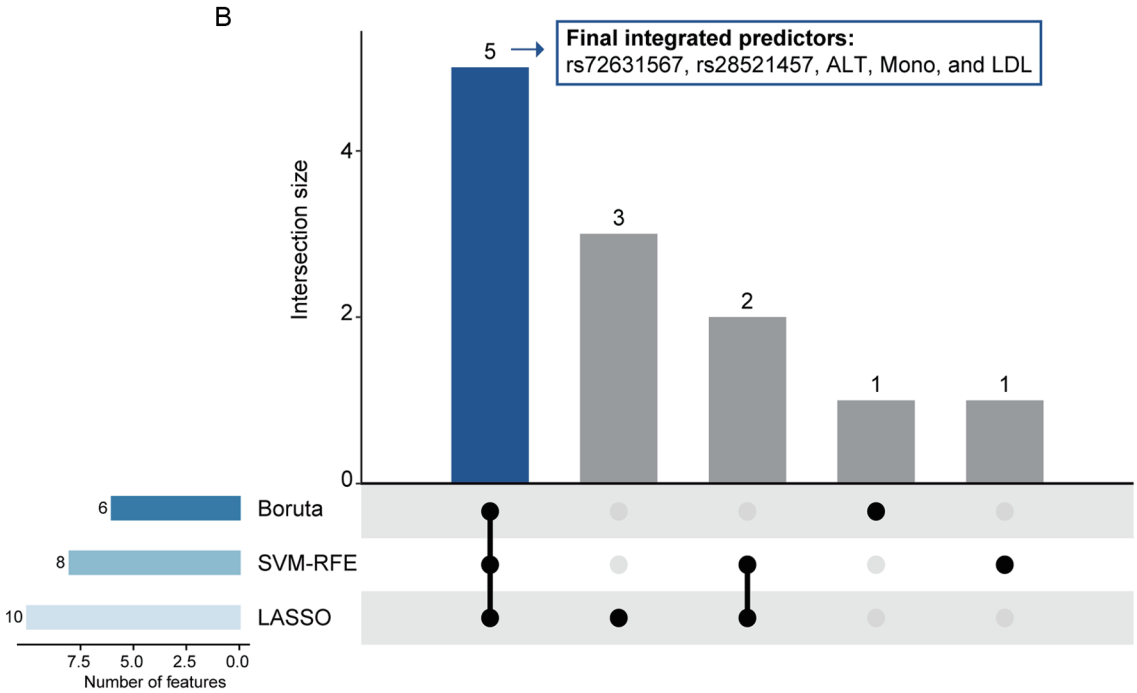

Fig. 2. Selection of candidate predictors. (A) Candidate variables retained by Boruta, SVM-RFE, and LASSO. (B) UpSet plot showing the overlap among predictors identified by the three feature-selection methods. SVM-RFE, support vector machine-recursive feature elimination; LASSO, least absolute shrinkage and selection operator; ALT, alanine aminotransferase; Mono, monocyte percentage; LDL, low-density lipoprotein; WBC, white blood cell; Eos, eosinophil percentage; Hb, hemoglobin; IgG, immunoglobulin G.

Table 3. Predictive performance metrics of six machine learning algorithms in the training and validation sets

Model	AUC (95% CI)	ACC	SEN	SPE	PPV	NPV	F1 score
Training set							
LightGBM	0.92 (0.89–0.95)	0.85	0.90	0.80	0.82	0.89	0.86
LR	0.78 (0.73–0.83)	0.71	0.67	0.75	0.72	0.70	0.70
RF	0.89 (0.86–0.93)	0.81	0.89	0.74	0.77	0.87	0.83
SVM	0.78 (0.73–0.83)	0.72	0.67	0.77	0.74	0.70	0.70
XGBoost	0.79 (0.74–0.85)	0.74	0.74	0.74	0.73	0.74	0.73
MLP	0.78 (0.73–0.84)	0.72	0.83	0.61	0.68	0.79	0.75
Validation set							
LightGBM	0.81 (0.70–0.91)	0.72	0.79	0.66	0.69	0.77	0.74
LR	0.80 (0.69–0.91)	0.72	0.76	0.69	0.7	0.75	0.73
RF	0.79 (0.68–0.90)	0.74	0.85	0.63	0.69	0.81	0.76
SVM	0.79 (0.68–0.90)	0.72	0.74	0.71	0.71	0.74	0.72
XGBoost	0.76 (0.64–0.88)	0.75	0.82	0.69	0.72	0.80	0.77
MLP	0.75 (0.64–0.87)	0.68	0.94	0.43	0.62	0.88	0.74

LightGBM, Light Gradient Boosting Machine; LR, logistic regression; RF, random forest; SVM, support vector machine; XGBoost, eXtreme Gradient Boosting; MLP, multilayer perceptron; AUC, area under the curve; ACC, accuracy; SEN, sensitivity; SPE, specificity; PPV, positive predictive value; NPV, negative predictive value.

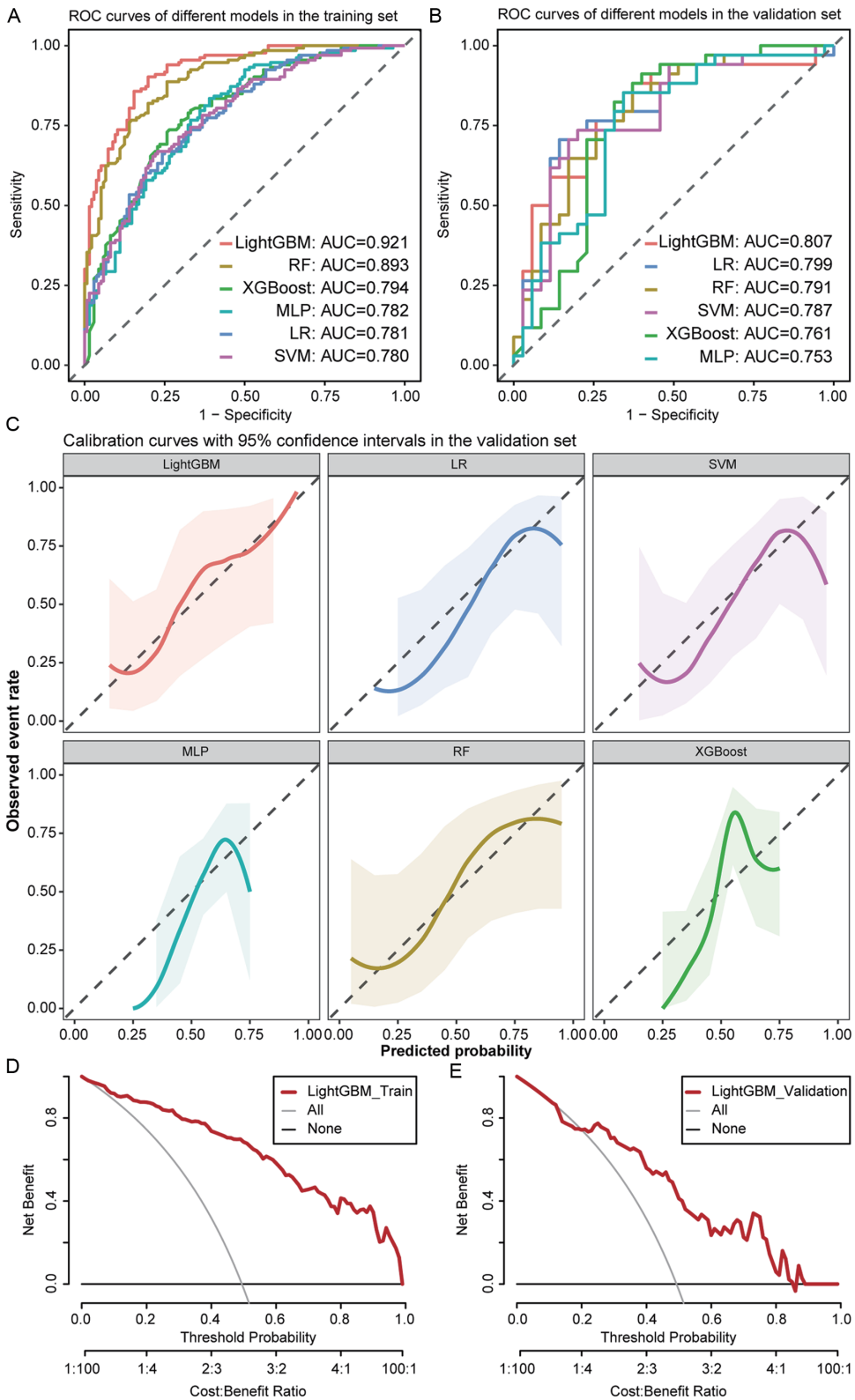

Fig. 3. Performance comparison of candidate machine learning models in the training and validation sets. (A, B) ROC curves in the training and validation sets. (C) Calibration curves of candidate machine learning models in the validation set. Shaded areas indicate 95% confidence intervals. (D, E) DCA for the final LightGBM model in the training and validation sets. ROC, receiver operating characteristic; AUC, area under the curve; LightGBM, Light Gradient Boosting Machine; RF, random forest; XGBoost, eXtreme Gradient Boosting; MLP, multilayer perceptron; LR, logistic regression; SVM, support vector machine; DCA, decision curve analysis.

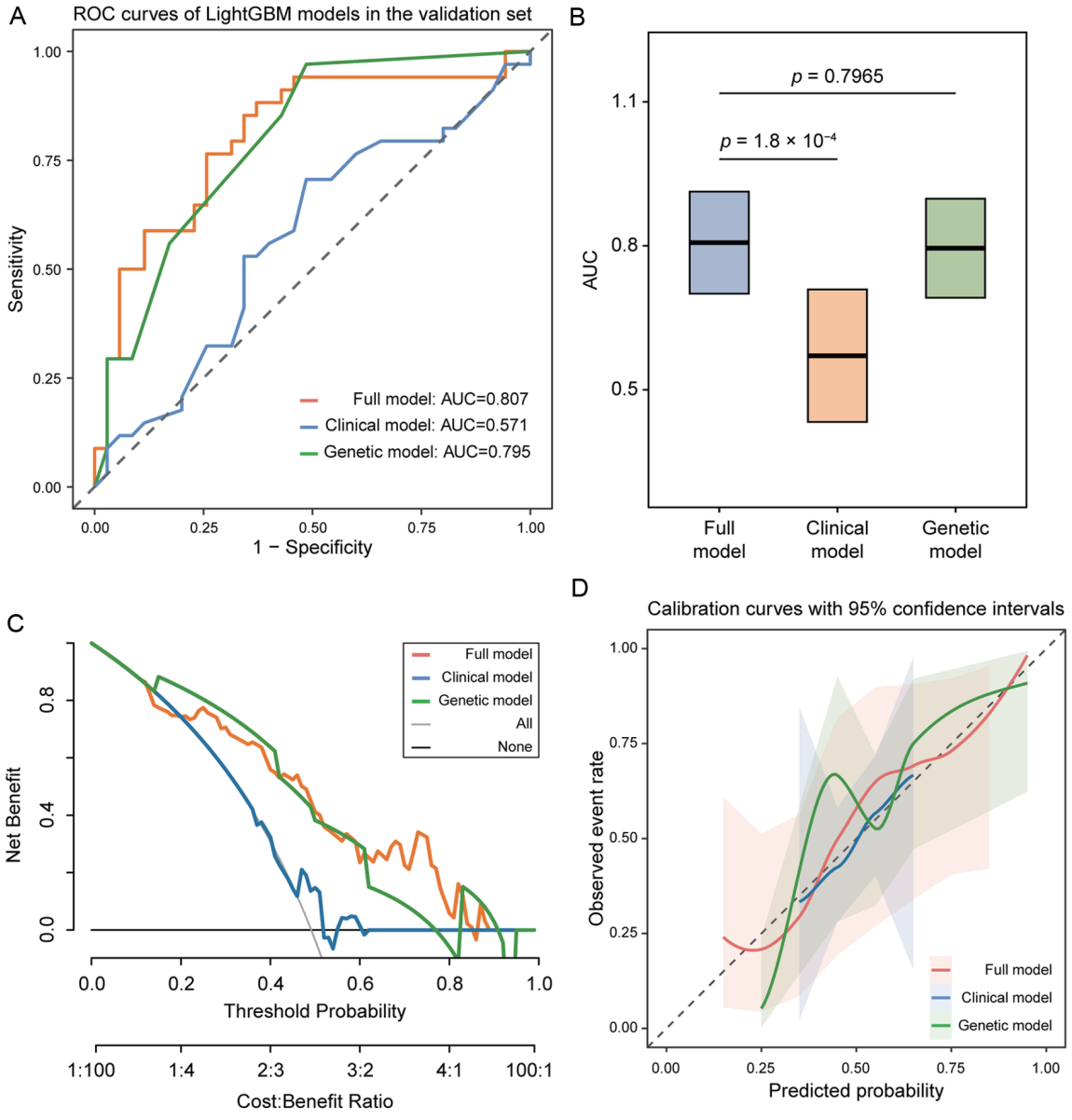

Fig. 4. Ablation analysis showing the incremental predictive value of genetic variants beyond clinical variables. (A) ROC curves of the full model (rs72631567, rs28521457, ALT, Mono, and LDL), genetic-only model (rs72631567 and rs28521457), and clinical-only model (ALT, Mono, and LDL) in the validation set. (B) Comparison of AUCs between models using DeLong's test. Horizontal lines indicate AUC values and boxes represent 95% confidence intervals. (C) DCA of the full, genetic-only, and clinical-only models in the validation set. (D) Calibration curves of the ablation models in the validation set. Shaded areas indicate 95% confidence intervals. ROC, receiver operating characteristic; LightGBM, Light Gradient Boosting Machine; ALT, alanine aminotransferase; Mono, monocyte percentage; LDL, low-density lipoprotein; AUC, area under the curve; DCA, decision curve analysis.

Discussion

In this retrospective, single-center cohort of patients with DILI, we developed an interpretable model to predict 6-month non-recovery by combining host genetic variants with baseline clinical variables available at initial hospitalization. We identified rs72631567, rs28521457, ALT, LDL, and Mono as features associated with 6-month non-recovery in DILI. Using these predictors, we constructed and compared multiple machine learning models. The LightGBM model showed the best overall performance among the algorithms examined. Notably, the addition of genetic information materially improved discrimination over the clinical-only model, suggesting that host genetic background may capture prognostic information not adequately reflected by baseline clinical features alone.

Early identification of patients at risk of non-recovery after DILI is clinically important, as persistent liver injury may delay treatment of the underlying condition and increase the risk of adverse long-term liver-related outcomes. However, reliable early prediction tools remain limited.[30] A baseline clinical model identified lymphocyte count and cholinesterase as predictors of chronic herbal DILI, but its discriminatory performance was modest (AUC = 0.663).[5] In our previous work, we found that noninvasive biomarker-based models may improve early risk stratification in DILI, including magnetic resonance imaging (MRI) radiomics- and metabolite

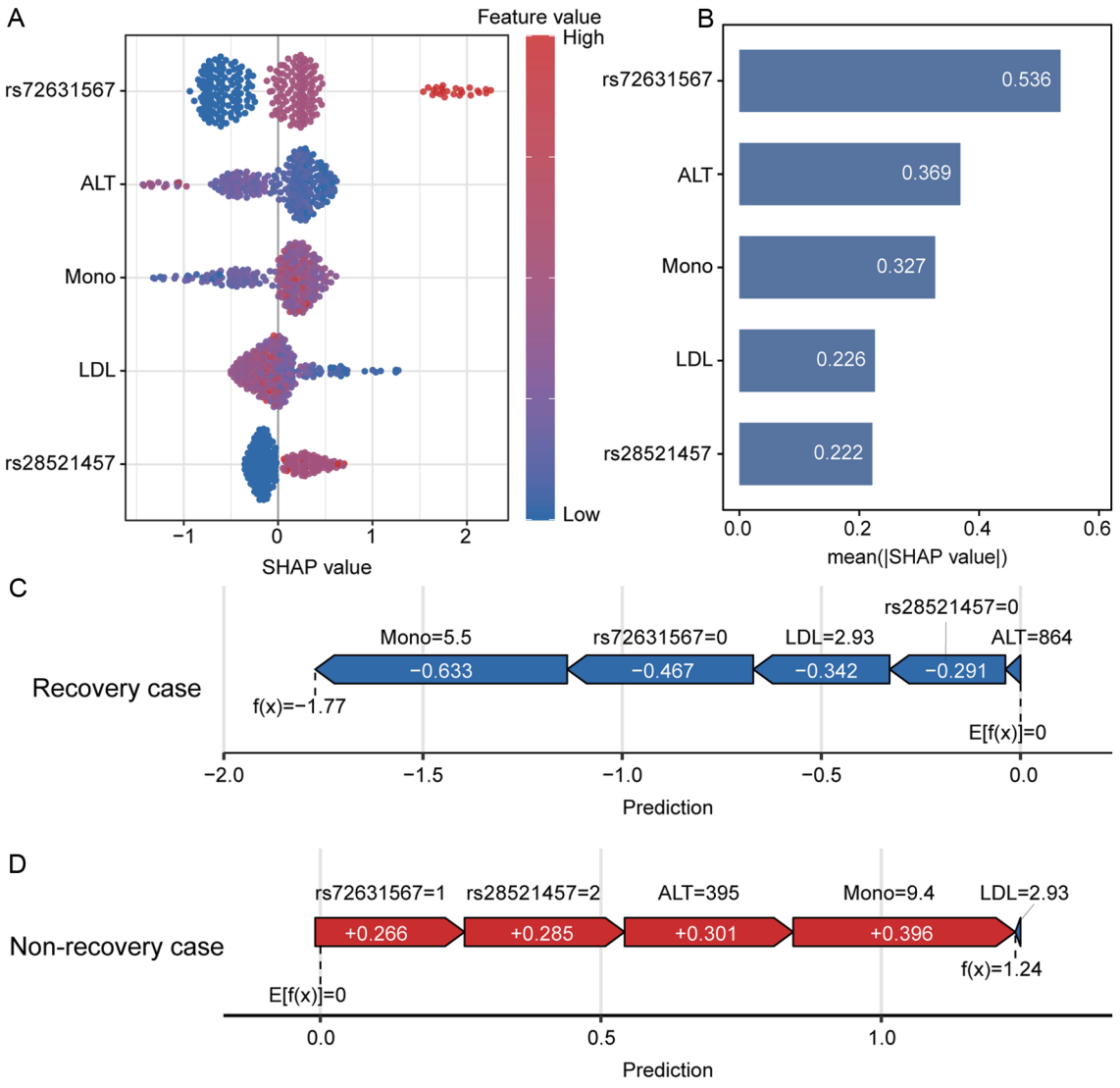

Fig. 5. SHAP-based interpretation of the LightGBM model. (A) Beeswarm plot showing the contribution of the included predictors to model output. (B) Bar plot of mean absolute SHAP values, ranking predictors by their overall contribution to the model. (C, D) SHAP force plots illustrating feature contributions to model predictions for representative recovery (C) and non-recovery (D) cases. SHAP, SHapley Additive exPlanations; LightGBM, Light Gradient Boosting Machine; ALT, alanine aminotransferase; LDL, low-density lipoprotein; Mono, monocyte percentage.

panel-based approaches with AUCs of 0.88 and 0.89, respectively.[6,7] However, these approaches remain influenced by disease stage and timing of assessment, which may limit their generalizability in early clinical decision-making. In contrast, genetic variants are stable throughout the disease course and may therefore provide complementary prognostic information at an early stage. Machine learning approaches may further enhance prediction by capturing nonlinear relationships and interactions among heterogeneous variables.

The five selected predictors may capture complementary aspects of DILI progression and recovery. Among the clinical variables, lower baseline ALT and LDL levels and higher Mono were associated with non-recovery in our cohort. Although ALT is commonly regarded as a marker of hepatocellular injury, this finding suggests that the long-term outcome of DILI may not be determined solely by the severity of hepatocellular injury at presentation. Instead, lower ALT may partly reflect differences in the injury phenotype, including a possible cholestatic or mixed component, which has been associated with slower resolution and a higher likelihood of persistent liver injury.[31] Higher Mono may indicate sustained innate immune activation or a persistent inflammatory milieu, which is biologically plausible given the recognized role of recruited monocytes in chronic hepatic inflammation and fibrogenic progression.[32,33] Lower LDL may reflect altered hepatic metabolic status or impaired liver synthetic capacity, although the biological basis of this association remains unclear. Although the genetic-only model performed comparably to the full model, adding ALT, Mono, and LDL modestly improved the AUC, yielded the lowest log loss, and produced a calibration curve closer to the ideal calibration line. These findings suggest that the clinical variables provided complementary information on baseline injury phenotype, inflammatory status, and metabolic context, thereby improving probability estimation and calibration, despite not being the dominant discriminatory signals in this cohort.

The genetic variables included in our model were selected based on a large international multicenter genome-wide association study reported by Nicoletti *et al*., in which rs28521457 and rs72631567 emerged as candidate non-HLA susceptibility loci for DILI.[13] Biologically, rs28521457 is an intronic variant within *LRBA*, a gene encoding a protein involved in vesicle trafficking and immune regulation. *LRBA* deficiency has been linked to reduced cytotoxic T lymphocyte-associated antigen-4 expression, regulatory T-cell dysfunction, and systemic immune dysregulation.[34,35] Variation within *LRBA* may plausibly influence not only susceptibility to DILI onset but also subsequent resolution of liver injury, as impaired immune homeostasis or weakened inhibitory checkpoint control could favor sustained inflammatory activation and delayed termination of the injury response. By contrast, rs72631567 lies in an intergenic region approximately 800 kb upstream of *SOX11*, and its proximity to *SOX11* suggests it may reside within regulatory elements that affect *SOX11* or neighboring genes. *SOX11* is a developmental transcription factor with established roles in cell fate determination, chromatin remodeling, and tissue remodeling, but its relevance to liver injury or DILI remains unclear.[36,37] Experimental evidence indicates that *SOX11* deficiency leads to upregulation of Mlxipl (also known as carbohydrate response element-binding protein, ChREBP), a key regulator of hepatic glucose and lipid metabolism, accompanied by mitochondrial dysfunction.[38,39] Given the importance of metabolic adaptation and mitochondrial integrity in injury resolution, rs72631567 may influence DILI chronicity through regulatory effects on *SOX11*-related pathways; however, this interpretation remains speculative. Notably, in our cohort, rs72631567 was the dominant genetic signal in the model, ranking first in SHAP analysis and contributing the largest incremental gain in performance in ablation analysis. Together, these findings suggest that rs72631567 may serve as a candidate host genetic marker associated with DILI recovery, although its functional relevance remains to be established.

These candidate SNPs were selected from previous DILI studies conducted mainly in European populations, and their relevance to Chinese patients should be interpreted with caution. In our cohort, rs114577328 showed no variation, whereas rs72631567 and rs28521457 were polymorphic and differed between the recovery and non-recovery groups, suggesting potential prognostic value. However, population differences in allele frequency, linkage disequilibrium patterns, and effect size may have affected the consistency of these associations. Therefore, further validation is required in independent Chinese cohorts and DILI populations from diverse ethnic groups.

This model has potential clinical value because it integrates variables available at the time of initial evaluation, enabling early risk stratification for non-recovery in DILI. In addition to showing acceptable discrimination, calibration, and decision curve performance, it also demonstrated favorable interpretability. SHAP analysis quantified the relative contribution of each predictor to the model output and illustrated how different combinations of genetic and clinical factors shaped individual risk predictions, thereby improving model transparency and supporting future clinical translation.

This study has several limitations. First, the cohort was drawn from a single center and consisted predominantly of Asian patients. Therefore, the allele frequencies of rs72631567 and rs28521457, as well as the model's performance, may differ in other populations and will require careful external validation. Second, only hospitalized patients were included, which may have introduced selection bias and enriched the cohort for more severe or clinically complex cases. Third, the outcome was defined at 6 months in this study, whereas 12-month outcome definitions have also been proposed in some studies; 12-month follow-up data were not uniformly available. Future studies should validate this model in independent, multicenter cohorts with extended follow-up periods and explore the functional relevance of these variants in DILI persistence and recovery.

Conclusions

In this study, we developed an interpretable model that integrates baseline clinical and genetic information to predict 6-month non-recovery in patients with DILI. The model showed encouraging performance in this cohort and highlighted the potential prognostic contribution of host genetic variants. However, independent multicenter validation is needed to confirm its generalizability and clinical utility.

Supporting information

Supplementary material for this article is available at https://doi.org/10.14218/JCTH.2026.00447.

Funding

This work was supported by the National Natural Science Foundation of China (81970514 to RL); the Shanghai Municipal Key Clinical Specialty (shslczdzk01103 to RL); and the Shanghai Magnolia Talent Award Program (24PJD066 to RL).

Conflict of interest

The authors have no conflicts of interest related to this publication.

Author contributions

Study concept and design (JJ, QX, RL), acquisition of clinical data (WwL, WqL, XZ), performance of SNP genotyping and genotype interpretation (QL, ZL), data curation and statistical analysis (JJ), interpretation of data (JJ, WwL, QX, RL), drafting of the manuscript (JJ, WwL), critical revision of the manuscript for important intellectual content (WwL, QX, RL), technical or material support (QL, ZL), funding acquisition (RL), administrative support and study supervision (QX, RL). All authors have made significant contributions to this study and have approved the final manuscript.

Ethical statement

This study was conducted in accordance with the Declaration of Helsinki (as revised in 2024) and was approved by the Ethics Committee of Ruijin Hospital (Ethics Approval No. 11–2017). The individual consent for this retrospective analysis was waived.

Data sharing statement

The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

References

[1] Suk KT, Kim DJ. Drug-induced liver injury: present and future. Clin Mol Hepatol 2012;18(3):249–257. doi:10.3350/cmh.2012.18.3.249, PMID:23091804.

[2] Hoofnagle JH, Björnsson ES. Drug-Induced Liver Injury - Types and Phenotypes. N Engl J Med 2019;381(3):264–273. doi:10.1056/NEJMra1816149, PMID:31314970.
[3] Fontana RJ, Hayashi PH, Gu J, Reddy KR, Barnhart H, Watkins PB, *et al*. Idiosyncratic drug-induced liver injury is associated with substantial morbidity and mortality within 6 months from onset. Gastroenterology 2014;147(1):96–108.e4. doi:10.1053/j.gastro.2014.03.045, PMID:24681128.
[4] Shen T, Liu Y, Shang J, Xie Q, Li J, Yan M, *et al*. Incidence and Etiology of Drug-Induced Liver Injury in Mainland China. Gastroenterology 2019;156(8):2230–2241.e11. doi:10.1053/j.gastro.2019.02.002, PMID:30742832.
[5] Zeng Z, Yi W, Dong JP, Chen QQ, Sun FF, Lu HH, *et al*. Baseline lymphocyte and cholinesterase levels may be the predictors of chronic herbal drug-induced liver injury. Front Pharmacol 2022;13:962480. doi:10.3389/fphar.2022.962480, PMID:35991883.
[6] Fu H, Shen Z, Lai R, Zhou T, Huang Y, Zhao S, *et al*. Clinic-radiomics model using liver magnetic resonance imaging helps predict chronicity of drug-induced liver injury. Hepatol Int 2023;17(6):1626–1636. doi:10.1007/s12072-023-10539-4, PMID:37188998.
[7] Zhao S, Fu H, Zhou T, Cai M, Huang Y, Gan Q, *et al*. Alteration of Bile Acids and Omega-6 PUFAs Are Correlated With the Progression and Prognosis of Drug-Induced Liver Injury. Front Immunol 2022;13:772368. doi:10.3389/fimmu.2022.772368, PMID:35493499.
[8] Hou T, Liu K, Fa W, Liu C, Zhu M, Liang X, *et al*. Association of polygenic risk scores with Alzheimer's disease and plasma biomarkers among Chinese older adults: A community-based study. Alzheimers Dement 2024;20(10):6669–6681. doi:10.1002/alz.13924, PMID:39171679.
[9] Mars N, Koskela JT, Ripatti P, Kiiskinen TTJ, Havulinna AS, Lindbohm JV, *et al*. Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. Nat Med 2020;26(4):549–557. doi:10.1038/s41591-020-0800-0, PMID:32273609.
[10] Stephens C, Andrade RJ. Genetic Predisposition to Drug-Induced Liver Injury. Clin Liver Dis 2020;24(1):11–23. doi:10.1016/j.cld.2019.08.003, PMID:31753244.
[11] Nicoletti P, Dellinger A, Li YJ, Barnhart HX, Chalasani N, Fontana RJ, *et al*. Identification of Reduced ERAP2 Expression and a Novel HLA Allele as Components of a Risk Score for Susceptibility to Liver Injury Due to Amoxicillin-Clavulanate. Gastroenterology 2023;164(3):454–466. doi:10.1053/j.gastro.2022.11.036, PMID:36496055.
[12] Li YJ, Phillips EJ, Dellinger A, Nicoletti P, Schutte R, Li D, *et al*. Human Leukocyte Antigen B*14:01 and B*35:01 Are Associated With Trimethoprim-Sulfamethoxazole Induced Liver Injury. Hepatology 2021;73(1):268–281. doi:10.1002/hep.31258, PMID:32270503.
[13] Nicoletti P, Aithal GP, Bjornsson ES, Andrade RJ, Sawle A, Arrese M, *et al*. Association of Liver Injury From Specific Drugs, or Groups of Drugs, With Polymorphisms in HLA and Other Genes in a Genome-Wide Association Study. Gastroenterology 2017;152(5):1078–1089. doi:10.1053/j.gastro.2016.12.016, PMID:28043905.
[14] Mao Y, Ma S, Liu C, Liu X, Su M, Li D, *et al*. Chinese guideline for the diagnosis and treatment of drug-induced liver injury: an update. Hepatol Int 2024;18(2):384–419. doi:10.1007/s12072-023-10633-7, PMID:38402364.
[15] European Association for the Study of the Liver. EASL Clinical Practice Guidelines: Drug-induced liver injury. J Hepatol 2019;70(6):1222–1261. doi:10.1016/j.jhep.2019.02.014, PMID:30926241.
[16] Aithal GP, Watkins PB, Andrade RJ, Larrey D, Molokhia M, Takikawa H, *et al*. Case definition and phenotype standardization in drug-induced liver injury. Clin Pharmacol Ther 2011;89(6):806–815. doi:10.1038/clpt.2011.58, PMID:21544079.
[17] van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. J Stat Softw 2011;45(3):1–67. doi:10.18637/jss.v045.i03.
[18] Tibshirani R. Regression Shrinkage and Selection Via the Lasso. J R Stat Soc Series B Stat Methodol 1996;58(1):267–288. doi:10.1111/j.2517-6161.1996.tb02080.x.
[19] Guyon I, Weston J, Barnhill S, Vapnik V. Gene Selection for Cancer Classification using Support Vector Machines. Mach Learn 2002;46(1-3):389–422. doi:10.1023/a:1012487302797.
[20] Kursa MB, Rudnicki WR. Feature Selection with the Boruta Package. J Stat Softw 2010;36(11):1–13. doi:10.18637/jss.v036.i11.
[21] Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In: Guyon I, Von Luxburg U, Bengio S, Wallach H, Fergus R, Vishwanathan S, *et al* (eds). Advances in Neural Information Processing Systems. Vol. 30. Curran Associates, Inc; 2017. doi:10.5555/3294996.3295074.
[22] Breiman L. Random Forests. Mach Learn 2001;45(1):5–32. doi:10.1023/a:1010933404324.
[23] Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20(3):273–297. doi:10.1007/bf00994018.
[24] Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016:785–794. doi:10.1145/2939672.2939785.
[25] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature 1986;323(6088):533–536. doi:10.1038/323533a0.
[26] Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26(6):565–574. doi:10.1177/0272989X06295361, PMID:17099194.
[27] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44(3):837–845. PMID:3203132.
[28] Chu J, Li Y, Wang X, Xu Q, Xu Z. Development of a Longitudinal Model for Disability Prediction in Older Adults in China: Analysis of CHARLS Data (2015-2020). JMIR Aging 2025;8:e66723. doi:10.2196/66723, PMID:40247464.
[29] Li H, Li Y, Gong C, Xiao Y, Su Z, Han L, *et al*. Study of a Kawasaki disease diagnostic prediction model based on the LightGBM machine learning algorithm. Front Artif Intell 2026;9:1776007. doi:10.3389/frai.2026.1776007, PMID:42266873.
[30] Wang Q, Huang A, Wang JB, Zou Z. Chronic Drug-Induced Liver Injury: Updates and Future Challenges. Front Pharmacol 2021;12:627133. doi:10.3389/fphar.2021.627133, PMID:33762948.
[31] Ashby K, Zhuang W, González-Jimenez A, Alvarez-Alvarez I, Lucena MI, Andrade RJ, *et al*. Elevated bilirubin, alkaline phosphatase at onset, and drug metabolism are associated with prolonged recovery from DILI. J Hepatol 2021;75(2):333–341. doi:10.1016/j.jhep.2021.03.021, PMID:33845060.
[32] Tacke F. Functional role of intrahepatic monocyte subsets for the progression of liver inflammation and liver fibrosis in vivo. Fibrogenesis Tissue Repair 2012;5(Suppl 1):S27. doi:10.1186/1755-1536-5-S1-S27, PMID:23259611.
[33] Liaskou E, Zimmermann HW, Li KK, Oo YH, Suresh S, Stamataki Z, *et al*. Monocyte subsets in human liver disease show distinct phenotypic and functional characteristics. Hepatology 2013;57(1):385–398. doi:10.1002/hep.26016, PMID:22911542.
[34] Janman D, Hinze C, Kennedy A, Halliday N, Waters E, Williams C, *et al*. Regulation of CTLA-4 recycling by LRBA and Rab11. Immunology 2021;164(1):106–119. doi:10.1111/imm.13343, PMID:33960403.
[35] Lo B, Zhang K, Lu W, Zheng L, Zhang Q, Kanellopoulou C, *et al*. AUTOIMMUNE DISEASE. Patients with LRBA deficiency show CTLA4 loss and immune dysregulation responsive to abatacept therapy. Science 2015;349(6246):436–440. doi:10.1126/science.aaa1663, PMID:26206937.
[36] Wegner M. From head to toes: the multiple facets of Sox proteins. Nucleic Acids Res 1999;27(6):1409–1420. doi:10.1093/nar/27.6.1409, PMID:10037800.
[37] Dodonova SO, Zhu F, Dienemann C, Taipale J, Cramer P. Nucleosome-bound SOX2 and SOX11 structures elucidate pioneer factor function. Nature 2020;580(7805):669–672. doi:10.1038/s41586-020-2195-y, PMID:32350470.
[38] Wu R, Tang W, Li Y, Deng Z, Zhang J, Li X, *et al*. Sox11 deficiency induces paravertebral muscle injury and scoliosis via Mlxipl upregulation. Genes Dis 2026;13(3):101970. doi:10.1016/j.gendis.2025.101970, PMID:41624909.
[39] Iizuka K. The Roles of Carbohydrate Response Element Binding Protein in the Relationship between Carbohydrate Intake and Diseases. Int J Mol Sci 2021;22(21):12058. doi:10.3390/ijms222112058, PMID:34769488.