Journal of Inflammation Research

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Routine Laboratory Markers-Based Machine Learning Model for Predicting Severe Kawasaki Disease in Pediatric Patients

Meng Wu[1,*], Jinlong Chen[2,*], Ya Gao[3,*], Hongbing Chen[1], Wei Li[4]

[1]Department of Clinical Laboratory, Children's Hospital of Nanjing Medical University, Nanjing, People's Republic of China; [2]Department of Cardiology, Children's Hospital of Nanjing Medical University, Nanjing, People's Republic of China; [3]Department of Emergency/Critical Medicine, Children's Hospital of Nanjing Medical University, Nanjing, People's Republic of China; [4]Department of Clinical Research, Children's Hospital of Nanjing Medical University, Nanjing, People's Republic of China

*These authors contributed equally to this work

Correspondence: Wei Li, Department of Clinical Research, Children's Hospital of Nanjing Medical University, Guangzhou Road #72, Nanjing, Jiangsu, 210008, People's Republic of China, Email weili126@126.com; Hongbing Chen, Department of Clinical Laboratory, Children's Hospital of Nanjing Medical University, Guangzhou Road #72, Nanjing, Jiangsu, 210008, People's Republic of China, Email Chenghb1999@126.com

Objective: Severe Kawasaki disease (SKD) poses numerous risks. Early identification of SKD is crucial for precise pharmacological intervention, which can reduce the incidence of complications. This study introduces a novel machine learning approach for the early prediction of SKD in pediatric populations, utilizing routinely collected laboratory parameters.

Methods: We extracted patients' age, sex, and 67 standard laboratory markers from the clinical records of 1,466 patients diagnosed with KD at the Children's Hospital of Nanjing Medical University. Using Lasso regression, we identified 15 critical predictors from the initial 69 candidates, demonstrating a significant impact on the accuracy of predictive outcomes. We forecasted the binary diagnosis of SKD or ordinary Kawasaki Disease (OKD) using 16 machine learning models, with model performance assessed through AUC-ROC, accuracy, F1 score, DCA, and calibration analysis.

Results: Our study included 1,466 patients with KD, categorized into 1,286 cases of OKD and 180 cases of SKD. Both groups predominantly consisted of male. A significantly lower median age was observed in SKD patients (4.29 years) compared to the OKD group. In our comparative analysis of predictive models, the Gradient Boosting model (AUC 0.952) emerged as the most accurate, followed closely by Ada Boost (AUC 0.945), Random Forest (AUC 0.944), CatBoost (AUC 0.957), and Naive Bayes (AUC 0.951). The GBC model achieved a high accuracy of 0.925, with a sensitivity of 0.628, specificity of 0.967, precision of 0.740, and an F1 score of 0.666, underscoring its robustness in distinguishing between the two KD subgroups. Our analysis identified 15 independent predictors, including absolute basophil count and conjugated bilirubin, that significantly enhanced the diagnostic accuracy of SKD.

Conclusion: Our most effective model demonstrates commendable performance in differentiating OKD from SKD. This advancement empowers pediatric clinicians to make swift clinical decisions, facilitating prompt therapeutic intervention and preventing the onset of severe complications.

Keywords: biomarkers, machine learning, Kawasaki disease, diagnosis

Introduction

Kawasaki disease (KD), a systemic vasculitis primarily affecting children under five, remains a significant pediatric concern.[1] Incomplete KD manifestations in infants under six months frequently result in delayed diagnosis and treatment initiation. This demographic is particularly prone to coronary artery lesions (CALs) and may exhibit resistance to initial treatments with intravenous immunoglobulin (IVIG).[2] Additionally, KD has been identified as a significant pediatric precursor to both childhood-acquired cardiovascular disorders in industrialized nations,[3] and is closely related to adult coronary arteriosclerosis and myocardial fibrosis.[4] Despite significant advances in KD management over the past half-century, its root causes remain obscure. Recent years have seen an increase in its incidence.[5] Currently, KD diagnosis

Received: 11 April 2025
Accepted: 12 July 2025
Published: 4 August 2025

depends on recognizing characteristic clinical features that develop sequentially over time, while validated biomarkers for early detection remain scarce in published studies.

In the acute phase, KD can lead to profound systemic involvement,[6] characterized by the emergence of Kawasaki disease shock syndrome (KDSS) and impairments across the digestive, nervous, renal, and cardiac systems, thereby significantly endangering the child's life. IVIG non-responsiveness occurs more frequently in this population, resulting in both heightened coronary artery injury risk and marked deterioration in quality of life measures.[7] The concept of severe Kawasaki disease (SKD) has been posited by scholars worldwide, who have conducted extensive research to facilitate the early identification of such cases.[8] SKD can lead to coronary artery aneurysms, which may result in heart attacks, heart failure, and even sudden death. However, a universally accepted definition and predictive framework for SKD applicable in clinical settings is still lacking.[9] Therefore, through systematic examination of the risk factors linked to the onset and progression of SKD and establishing a systematic approach with early predictive indicators, it is possible to promptly recognize these cases, intervene medically, and thus ameliorate disease progression and improve patient outcomes.

Recent advancements in machine learning and statistical methodologies have significantly propelled the development of sophisticated clinical predictive models. These models not only enhance diagnostic accuracy but also facilitate physicians in reaching swift diagnostic conclusions. Scholarly advancements have effectively utilised machine learning algorithms to devise predictive models of notable precision. These models are particularly adept at distinguishing Kawasaki disease from other clinically similar pathologies, such as sepsis[10] and multisystem inflammatory syndrome.[11] At present, no machine learning model has been specifically tailored to distinguish ordinary Kawasaki disease (OKD) from SKD based on Laboratory test index.

This study introduces a classification framework for KD, differentiating between OKD, which presents without significant organ damage, and SKD, which is marked by extensive system impairment. In our study, we have incorporated a set of standard laboratory tests as the metrics for evaluation. These tests are cost-effective, ensuring that our approach is accessible and widely applicable in clinical settings. By retrospectively examining the laboratory indices of these two groups, we propose to uncover previously unrecognized risk determinants for SKD advancement while developing a biomarker panel for pre-symptomatic diagnosis. These tests are widely available in both inpatient and outpatient settings and are cost-effective, making them accessible across healthcare facilities without requiring advanced or specialized laboratories. By utilizing these readily accessible and standardized tests, the model enhances its potential to be broadly implemented across diverse healthcare environments, ensuring that early and accurate identification of SKD can be achieved nationwide.

Methods

Experimental Design and Patient Cohort

We conducted a retrospective cohort study analyzing clinical records of pediatric KD patients at the Children's Hospital of Nanjing Medical University. The study period spans from January 2017 to July 2023. In our study, ordinary Kawasaki Disease (OKD) is characterized by hemodynamic stability and the absence of significant organ dysfunction. SKD refers to cases that fulfill the diagnostic criteria for Kawasaki disease and exhibit any of the following complications necessitating intensive care: (1). Hemodynamic compromise, including hypotension, poor perfusion, or shock; (2). Severe dysfunction of single or multiple organs, such as the cardiovascular, gastrointestinal, renal, or neurological systems, which is attributable to Kawasaki Disease; (3). Acute exacerbation of pre-existing comorbidities that results in instability of vital signs. The study cohort comprised 180 SKD patients (119 males, 61 females; median age: 4.29 years). For comparison, 1286 contemporaneously hospitalized OKD patients (795 males, 491 females; median age: 4.83 years) were enrolled as controls (Figure 1). The study included patients who met the following criteria: (1) fulfilment of the diagnostic criteria for KD;[3] (2) availability of complete clinical data; (3) completion of a comprehensive laboratory examination. Participants were excluded based on the following criteria: (1) the presence of chronic underlying diseases and immune system abnormalities in patients; (2) prior administration of IVIG or glucocorticoids before hospitalisation; (3) co-occurrence of severe infectious diseases; (4) other confusing diseases, such as Epstein-Barr virus infection, sepsis, juvenile idiopathic arthritis, etc; (5) incompleteness of clinical data.

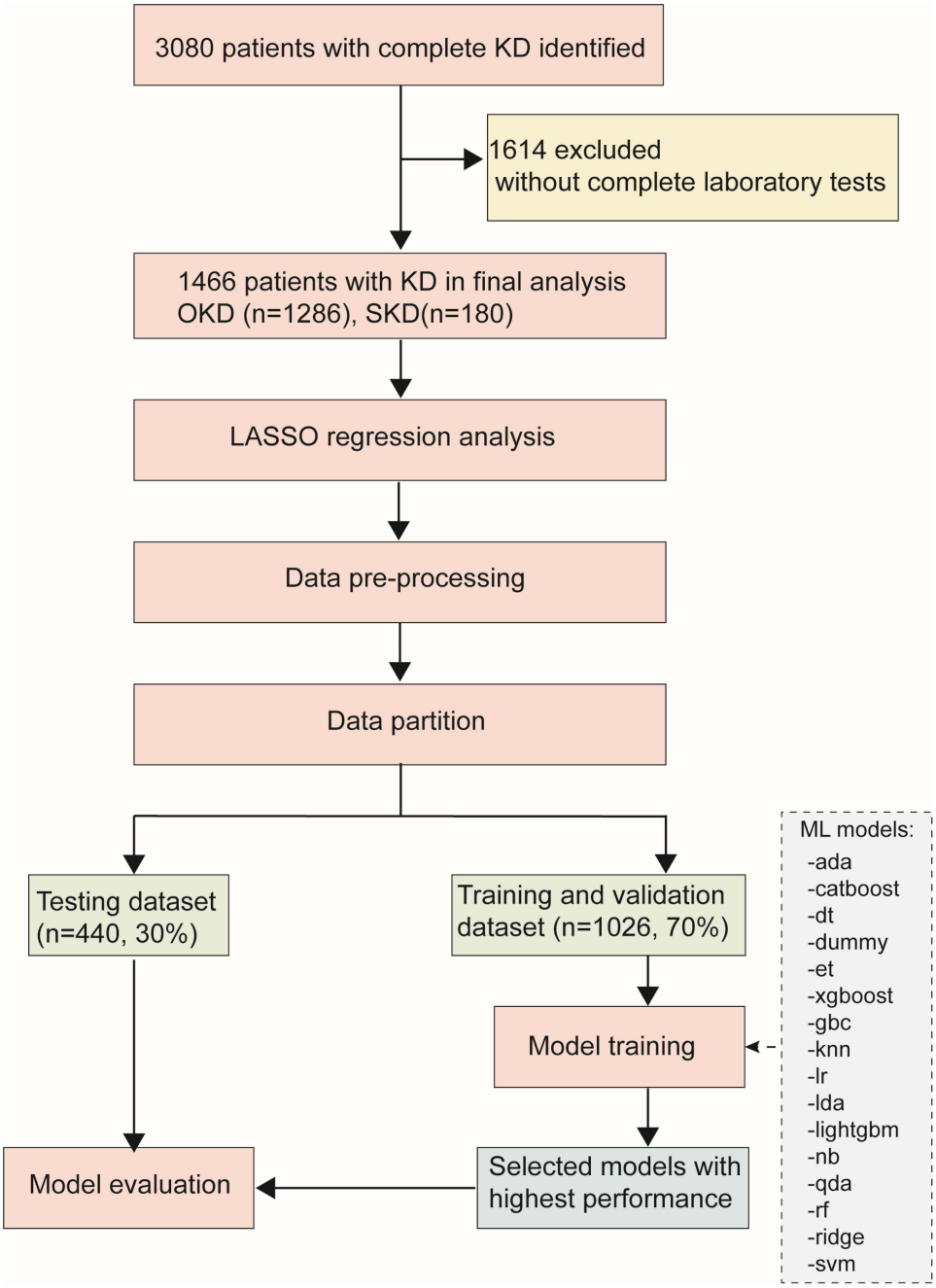

Figure 1 Flowchart illustrating the method overview.
Abbreviations: ML, Machine Learning; ada, AdaBoost; catboost, CatBoost; dt, Decision Tree; dummy, Dummy; et, Extra Trees; xgboost, XGBoost; gbc, Gradient Boosting; knn, K-Neighbors; lr, Logistic Regression; lda, Linear Discriminant Analysis; lightgbm, LightGBM; nb, Naïve Bayes; qda, Quadratic Discriminant Analysis; rf, Random Forest; ridge, Ridge; svm, Support Vector Machine—Linear Kernel.

Data Collection

We systematically collected data across 69 variables from a spectrum of tests, encompassing age, sex, routine blood, biochemical, coagulation, and immunotyping analyses. Routine blood analyses included a panel of standard metrics: erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), white blood cell (WBC) count, differential counts for neutrophils, lymphocytes, monocytes, eosinophils, and basophils, as well as red blood cell (RBC) parameters like haemoglobin concentration, haematocrit, mean corpuscular volume (MCV), mean corpuscular haemoglobin (MCH), mean corpuscular haemoglobin concentration (MCHC), red cell distribution width (RDW) standard deviation (SD),

https://doi.org/10.2147/JIR.S528341

RDW coefficient of variation (CV), platelet count, platelet distribution width (PDW), mean platelet volume (MPV), plateletcrit (PCT), and platelet large cell ratio (P-LCR). Biochemical tests further expanded this dataset with measurements of alanine transaminase (ALT), aspartate transaminase (AST), alkaline phosphatase (ALP), gamma-glutamyltransferase (GGT), cholinesterase, lactate dehydrogenase (LDH), creatine kinase (CK), CK-MB, alpha-hydroxybutyrate dehydrogenase (α-HBDH), total protein (TP), albumin (ALB), globulin (GLB), albumin/globulin ratio (A/G), glucose (GLU), total bilirubin (TBIL), direct bilirubin (DBIL), indirect bilirubin (IBIL), prothrombin activity (PA), urea, uric acid (UA), creatinine (Cr), cystatin C (Cys-C), retinol-binding protein (RBP), and electrolytes including potassium (K), sodium (Na), chloride (Cl), calcium (Ca), magnesium (Mg), and iron (Fe). Lipid profiles were also assessed through triacylglycerol (TRIG), total cholesterol (TC), and high-density lipoprotein cholesterol (HDL-C). Coagulation profiles included prothrombin time (PT), international normalised ratio (INR), activated partial thromboplastin time (APTT), fibrinogen (Fbg), thrombin time (TT), D-dimer (DD), and fibrinogen degradation products (FDP). Lymphocyte typing analyses rounded out the dataset with T lymphocyte (TLBX), $CD3^+CD4^+$, $CD3^+CD8^+$, $CD4^+CD8^+$, CD3-$CD16^+CD56^+$, $CD3^-CD19^+$, and the $CD4^+/CD8^+$ ratio. All data for KD patients were gathered at the initial presentation prior to the administration of IVIG.

Data Preprocessing

Data preprocessing involved several steps to ensure the quality and usability of the dataset for ML modeling. Missing values were handled using imputation methods suitable for each type of data. Laboratory test results were standardized to a common scale to facilitate model training. Anomalous data points were detected and processed using established statistical protocols to maintain model integrity.

Predictor Selection Using LASSO

With its inherent capacity for dimension reduction, Least Absolute Shrinkage and Selection Operator (LASSO) simultaneously performs variable selection by driving non-essential predictors' coefficients to exact zero. By leveraging LASSO for feature selection, this study aims to identify the most influential laboratory parameters that predict the progression of OKD to SKD. The identified biomarker panel may elucidate pathogenic pathways while facilitating clinically actionable predictive algorithms for precision medicine approaches. The implementation of LASSO for feature selection using the glmnet package in R program.

Machine Learning Models

The machine learning framework was implemented with PyCaret 3.0 on Python 3.11. Data were partitioned (7:3 train: test ratio) with stratified 10-fold cross-validation to ensure robust performance evaluation. Sixteen machine learning classifiers were systematically evaluated to determine the optimal prognostic model for SKD patients. Primary endpoints were AUC-ROC for discriminative ability, alongside standard classification metrics (accuracy/precision/recall/F1). Clinical utility was further assessed through decision curve analysis(DCA), complemented by probability calibration evaluation.

Feature Interpretation

We used the Shapley Additive Explanations (SHAP) algorithm to explain the major features that made significant contributions to the prediction after the ML model was trained, as appropriate. SHAP quantifies feature importance through Shapley values - the average marginal prediction impact when a feature is added to any possible feature coalition. These Shapley values are aggregated through weighted averaging to determine each feature's total predictive influence.

Statistical Analysis

In addition to ML model evaluation, statistical analyses were conducted to compare the baseline characteristics and laboratory parameters between patients with SKD and those with OKD. Continuous and categorical variables were analyzed using parametric/nonparametric tests (*t*-test/Mann–Whitney U) and χ^2/Fisher's exact tests, respectively (significance threshold α=0.05).

Ethical Considerations

The study protocol was approved by Nanjing Medical University Children's Hospital Institutional Review Board approval (202306013–1) following Helsinki Declaration guidelines. Parental/guardian informed consent was obtained for all participants. Parental/guardian consent documentation was obtained for every enrolled subject. By employing advanced ML techniques on routinely collected laboratory data, this study aims to enhance the early prediction of SKD, ultimately contributing to improved management and outcomes for children affected by Kawasaki Disease.

Results

Intergroup Demographic and Laboratory Characteristics Comparison

This study included a total of 1,466 Kawasaki disease patients, of which 180 were severe cases and 1,286 were non-severe cases. Patients in the severe group were younger (median 4.29 vs 4.83 years, P=0.002). No significant gender disparity was observed (χ^2=1.24, P=0.266), with males comprising the majority in both cohorts. Among the 21 routine blood test indicators compared between the two groups, except for Absolute Monocyte Count and MCHC, all other indicators showed statistically significant differences. The comparison of 32 blood biochemical indicators between the two groups revealed that, except for CK, CK-MB, α-HBDH, and Urea, the other 28 blood biochemical indicators showed statistically significant differences. Coagulation profiles and cellular immunity markers differed significantly between groups. Detailed comparative metrics are tabulated in Table 1.

Table 1 Basic Characteristics and Laboratory Test Values Between Children with OKD and SKD

Variables	OKD	SKD	Z/t/χ^2	*P* value
Basic clinical information				
Patients, n	1286	180		
Age	4.83(3.67, 6.00)	4.29(2.77, 6.00)	−3.050	0.002**
Sex, n(%)	Male, 795(61.8%)	Male, 119(66.1%)	1.24	0.266
	Female, 491(38.2%)	Female,61(33.9%)		
Blood routine parameters				
ESR (mm/h)	68.00(44.75, 87.00)	65.59(42.25, 79.75)	−2.252	0.024*
CRP (mg/L)	48.63(20.00, 80.00)	84.00(50.17, 119.87)	−7.046	0.000***
WBC (10^9 cells per L)	11.46(8.01, 15.71)	15.64(10.68, 20.59)	−7.062	0.000***
Lymph# (10^9 cells per L)	3.39(2.42, 4.62)	2.11(1.29, 3.43)	−8.789	0.000***
Mono# (10^9 cells per L)	0.68(0.51, 0.92)	0.70(0.42, 1.16)	−0.869	0.385
Neut# (10^9 cells per L)	6.45(3.28, 10.43)	11.60(7.51, 16.44)	−10.271	0.000***
Eo# (10^9 cells per L)	0.32(0.14, 0.59)	0.10(0.03, 0.26)	−9.350	0.000***
Baso# (10^9 cells per L)	0.21(0.07, 0.39)	0.02(0.01, 0.04)	−13.931	0.000***
RBC (10^{12} cells per L)	3.93(3.71, 4.15)	4.14(3.78, 4.44)	−5.572	0.000***
HGB (g/L)	105.00(98.00, 111.00)	112.00(102.00, 119.00)	−6.826	0.000***
HCT (%)	31.80(29.70, 33.70)	33.60(31.13, 36.58)	−6.833	0.000***
MCV (fL)	80.80(78.29, 83.33)	82.20(78.90, 85.05)	−3.776	0.000***
MCH (pg)	26.80(25.81, 27.71)	27.20(26.50, 28.00)	−4.287	0.000***
MCHC (g/L)	330.34(322.98, 337.46)	330.00(320.00, 338.00)	−0.227	0.820
RDW-SD (fL)	38.80(37.40, 40.60)	39.80(38.13, 41.90)	−4.027	0.000***
RDW-CV (%)	13.20(12.70, 13.90)	13.40(12.90, 14.20)	−2.822	0.005**
PLT (10^9 cells per L)	384.00(316.00, 475.00)	272.50(205.25, 348.50)	−11.417	0.000***
PCT (%)	0.37(0.31, 0.45)	0.29(0.22, 0.35)	−9.522	0.000***
MPV (fL)	9.60(9.10, 10.20)	10.10(9.44, 10.78)	−6.173	0.000***
PDW (fL)	10.40(9.50, 11.40)	10.90(9.83, 12.65)	−4.799	0.000***
P-LCR (%)	21.10(17.00, 25.70)	25.05(20.13, 30.38)	−6.655	0.000***

(*Continued*)

Table 1 (Continued).

Variables	OKD	SKD	Z/t/χ^2	*P* value
Biochemistry parameters				
ALT (U/L)	18.00(11.00, 36.00)	57.50(17.00, 141.75)	−8.178	0.000***
AST (U/L)	29.00(21.00, 38.00)	37.00(25.00, 99.50)	−6.176	0.000***
ALP (U/L)	151.00(124.00, 184.00)	206.00(140.00, 283.75)	−7.615	0.000***
GGT (U/L)	22.00(12.00, 53.00)	111.00(31.50, 182.50)	−11.217	0.000***
CHE (U/L)	5567.00(4307.75, 6968.00)	4772.00(3735.00, 99.50)	−5.155	0.000***
LDH (U/L)	273.00(238.00, 314.00)	295.50(251.25, 351.25)	−3.879	0.000***
CK (U/L)	36.00(25.00, 53.00)	35.00(25.00, 62.75)	−0.569	0.569
CK-MB (U/L)	20.00(16.00, 27.00)	19.50(15.00, 25.00)	−1.793	0.073
α-HBDH (U/L)	207.00(182.0, 235.00)	214.50(188.00, 248.75)	−1.896	0.058
TP (g/L)	73.30(63.38, 81.00)	59.95(55.65, 65.68)	−12.064	0.000***
ALB (g/L)	35.70(32.50, 38.50)	34.70(29.93, 38.58)	−2.354	0.019*
GLO (g/L)	37.70(26.08, 46.93)	23.75(29.93, 28.38)	−11.102	0.000***
A/G	0.99(0.72, 1.45)	1.50(1.12, 1.79)	−8.555	0.000***
GLU (mmol/L)	4.19(3.80, 4.75)	5.01(4.30, 6.03)	−8.615	0.000***
TBIL (umol/L)	4.90(3.50, 7.21)	12.25(6.25, 31.49)	−12.965	0.000***
DBIL (umol/L)	2.00(1.40, 3.10)	7.96(2.90, 24.70)	−14.246	0.000***
IBIL (umol/L)	2.73(1.88, 4.10)	3.95(2.60, 7.20)	−7.048	0.000***
PA (g/L)	0.09(0.06, 0.16)	0.06(0.04, 0.08)	−9.525	0.000***
Urea (mmol/L)	1.09±0.95	1.04±0.37	0.700	0.484
UA (umol/L)	248.00(198.75, 305.00)	229.00(188.00, 285.75)	−2.474	0.013*
Cr (umol/L)	25.00(20.58, 30.00)	26.95(22.00, 33.75)	−3.372	0.001**
Cys-C (mg/L)	1.02(0.88, 1.23)	0.97(0.79, 1.17)	−3.155	0.022*
RBP (mg/L)	16.70(10.92, 27.10)	10.69(7.16, 14.40)	−9.354	0.000***
K (mmol/L)	4.47(4.09, 5.07)	4.01(3.64, 4.46)	−9.278	0.000***
Na (mmol/L)	136.00(134.10, 138.00)	133.60(131.20, 136.15)	−8.710	0.000***
Cl (mmol/L)	100.20(97.90, 102.30)	97.80(95.00, 100.28)	−7.965	0.000***
Ca (mmol/L)	2.28(2.18, 2.38)	2.21(2.09, 2.35)	−4.580	0.000***
Mg (mmol/L)	0.96(0.89, 1.02)	0.90(0.83, 0.95)	−7.177	0.000***
P (mmol/L)	1.49(1.27, 1.67)	1.10(0.87, 1.29)	−13.303	0.000***
TRIG (mmol/L)	1.31(1.02, 1.73)	1.22(0.98, 1.58)	−2.610	0.009**
TC (mmol/L)	3.38(2.87, 3.37)	2.94(2.57, 3.41)	−6.929	0.000***
HDL-C (mmol/L)	0.66(0.50, 0.83)	0.57(0.41, 0.88)	−2.267	0.023*
Coagulation parameters				
PT (s)	12.60(11.70, 13.50)	13.45(12.70, 14.70)	−7.982	0.000***
INR	1.12(1.04, 1.20)	1.19(1.11, 1.31)	−7.537	0.000***
APTT (s)	33.60(30.98, 36.20)	34.05(31.50, 37.63)	−2.619	0.009**
Fbg (g/L)	4.72(3.73, 4.91)	4.77(3.73, 4.91)	−5.515	0.000***
TT (s)	14.70(13.80, 15.80)	14.25(13.30, 15.10)	−4.490	0.000***
DD (ng/mL)	542.00(315.75, 874.25)	516.28(567.00, 1037.75)	−7.543	0.000***
FDP (ug/mL)	4.91(3.24, 7.90)	8.00(5.16, 9.02)	−7.256	0.000***
Cellular immunity parameters				
TLBX (%)	63.89(56.48, 70.57)	52.32(43.30, 61.31)	−10.654	0.000***
CD3+CD4+ (%)	39.20(32.39, 46.12)	31.74(23.96, 27.92)	−9.663	0.000***
CD3+CD8+ (%)	19.72(15.68, 23.85)	16.62(12.37, 22.86)	−4.770	0.000***
CD4+CD8+ (%)	0.48(0.26, 0.92)	0.37(0.20, 0.68)	−2.798	0.005**

(*Continued*)

Table 1 (Continued).

Variables	OKD	SKD	Z/t/χ^2	*P* value
CD3-CD16+CD56+ (%)	7.71(4.88, 11.81)	9.16(5.41, 12.65)	−2.027	0.043*
CD3-CD19+ (%)	22.96(17.39, 29.62)	33.28(25.56, 41.88)	−10.194	0.000***
CD4+/CD8+ (%)	1.97(1.50, 2.65)	1.89(1.24, 2.56)	−1.930	0.054

Notes: All continuous variables underwent normality testing. Data conforming to a normal distribution are expressed as mean ± standard deviation, while non-normally distributed data are expressed as median (P25, P75). The enumeration data were expressed as frequency and rate (%).*P<0.05, **P<0.01, ***P<0.001.

Abbreviations: OKD, ordinary Kawasaki disease; SKD, severe Kawasaki disease; ESR, erythrocyte Sedimentation Rate; CRP, C-reactive protein; WBC, White blood cell count; Lymph#, absolute lymphocyte count; Mono#, absolute monocyte count; Neut#, absolute neutrophil count; Eo#, absolute eosinophil count; Baso#, absolute basophil count; RBC, red blood cell count; HGB, haemoglobin concentration; HCT, hematocrit; MCV, mean erythrocyte volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RDW-SD, red cell distribution width-SD; RDW-CV, red cell distribution width-CV; PLT, platelet; PCT, plateletcrit; MPV, mean platelet volume; PDW, platelet distribution width; P-LCR, platelet-large contrast ratio; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase; GGT, gamma-glutamyl transferase; CHE, cholinesterase; LDH, lactic dehydrogenase; CK, creatine kinase; CK-MB, creatine kinase isoenzymes; α-HBDH, alpha-hydroxybutyric dehydrogenase; TP, total protein; ALB, albumin; GLO, globulin; A/G, albumin/globulin; GLU, glucose; TBIL, total bilirubin; DBIL, direct bilirubin; IBIL, indirect bilirubin; PA, prealbumin; UA, Uric acid; Cr, creatinine; Cys-C, cystatin C; RBP, retinol-binding protein; K, potassium; Na, sodium; Cl, chlorine; Ca, calcium; Mg, magnesium; P, phosphorus; TRIG, triacylglycerol; TC, total cholesterol; HDL-C, high-density lipoprotein cholesterol; PT, prothrombin time; INR, international normalized ratio; APTT, activated partial thromboplastin time; Fbg, fibrinogen; TT, thrombin time; DD, D-dimer; FDP, fibrinogen degradation products; TLBX, T lymphocyte.

LASSO Regression for the Feature Selection

LASSO regression with 10-fold CV identified optimal predictors while controlling overfitting (Figure 2A). The optimal regularization strength was determined by analyzing the cross-validation error across log of lambda values.

The left dashed vertical line indicated that the log of the optimal value of lambda was approximately-3.5, which minimized the prediction error (Figure 2A). The LASSO regression analysis selected 15 variables out of the initial 69 candidates (Figure 2B). These variables were found to significantly contribute to the model's predictive power, as indicated by their non-zero coefficients (Table S1). These coefficients reflect the relative importance of each variable in predicting SKD.

After standardizing each variable, we used LASSO regression again to select characteristic variables. The results showed that the 15 characteristic variables were consistent with the unstandardized results (Figure S1 and Table S2).

Machine Learning Model Selection

Table 2 summarizes the test set performance of all 16 candidate models. Tree-based ensembles (GBC, ADA, RF, CatBoost) outperformed linear models in SKD/OKD classification. Among all models, GBC demonstrated superior performance in training set SKD prediction (AUC=0.952, accuracy=0.925), with balanced sensitivity (0.628) and specificity (0.967), alongside strong precision (0.740) and F1-score (0.666).

Evaluation of the GBC Model Performance for SKD Diagnosis in the Training Set

We use the testing set to evaluate the performance of the GBC model for SKD diagnosis, GBC model maintained robust performance on the test set (AUC=0.952, accuracy=0.918), with specificity exceeding sensitivity (0.95 vs 0.72). The GBC model demonstrated discriminative performance as visualized by its ROC curve (Figure 3A). The GBC model demonstrated 72% sensitivity for SKD detection and 95% specificity for OKD identification (Figure 3B). Decision curve analysis confirmed the GBC model's clinical utility throughout the threshold probability spectrum. DCA confirmed the GBC model's significant advantage versus "treat-none", demonstrating tangible clinical utility (Figure 4A). Figure 4B's calibration plot confirms the GBC model's well-calibrated predictions across risk strata.

Feature Importance Analysis

The GBC model's decision logic was deconstructed via SHAP, quantifying each variable's contribution to SKD prediction. Figure 5A ranks the top 15 predictors by mean |SHAP| value, quantifying their global importance. SHAP

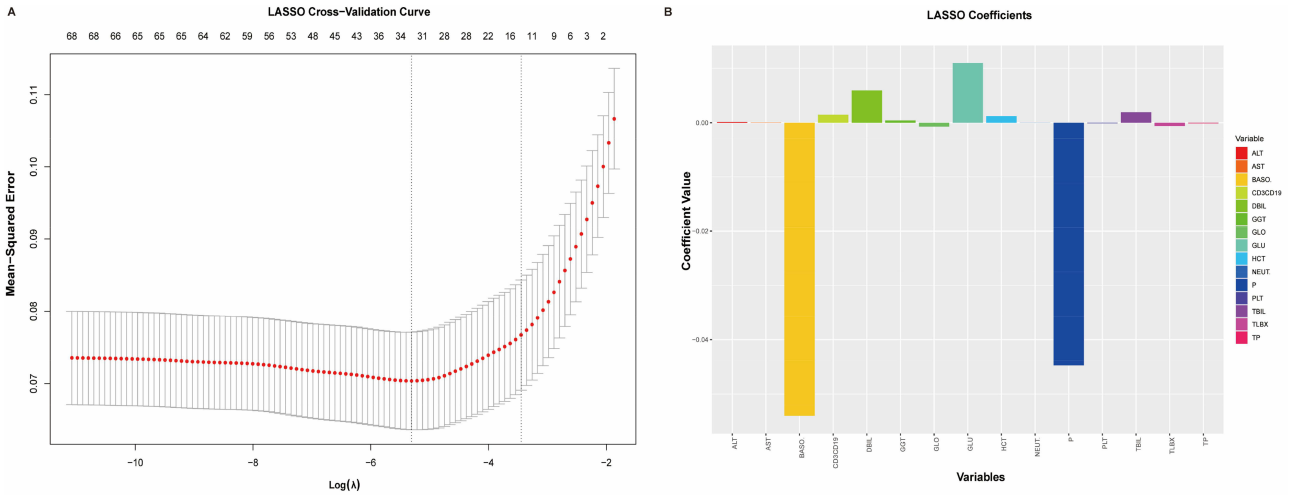

Figure 2 Feature selection via Lasso-Cox regression (**A**) The selection process of the optimum value of the parameter λ in the Lasso regression model; (**B**) Variable coefficient dynamics across λ values.

values quantified each feature's marginal contribution to model predictions. The impact of each predictor on the GBC-based prediction model were visualized through SHAP value distributions in Figure 5B. The horizontal axis displays SHAP values, while the vertical axis lists predictors ordered by importance. Feature values are color-mapped (red=high, blue=low) relative to their population distribution. Elevated SHAP values identified BASO#, DBIL, GLU, PLT and GLO as key SKD determinants.

Table 2 The Comparison of Model Performance Metrics

Model	Accuracy	AUC	Sensitivity	Specificity	Precision	F1-Score	Kappa	MCC	TT (Sec)
gbc	0.925	0.952	0.628	0.967	0.740	0.666	0.6251	0.6353	0.289
ada	0.919	0.945	0.611	0.962	0.702	0.640	0.596	0.605	0.130
rf	0.917	0.944	0.485	0.977	0.784	0.584	0.542	0.569	0.204
catboost	0.916	0.957	0.547	0.968	0.717	0.607	0.563	0.576	3.536
nb	0.915	0.951	0.626	0.955	0.674	0.642	0.595	0.599	0.028
lr	0.914	0.937	0.501	0.972	0.725	0.580	0.535	0.553	1.790
et	0.914	0.934	0.428	0.982	0.796	0.544	0.502	0.537	0.163
lightgbm	0.913	0.951	0.572	0.961	0.672	0.609	0.561	0.569	0.232
ridge	0.909	0.000	0.349	0.987	0.816	0.478	0.438	0.490	0.028
lda	0.908	0.921	0.444	0.973	0.740	0.537	0.491	0.519	0.026
qda	0.907	0.947	0.564	0.955	0.646	0.598	0.546	0.550	0.033
xgboost	0.907	0.945	0.565	0.955	0.652	0.593	0.541	0.550	0.100
dt	0.906	0.758	0.561	0.954	0.638	0.589	0.538	0.543	0.034
knn	0.900	0.799	0.348	0.977	0.681	0.452	0.405	0.435	0.050
svm	0.885	0.000	0.389	0.954	0.643	0.428	0.374	0.415	0.027
dummy	0.877	0.500	0.000	1.000	0.000	0.000	0.000	0.000	0.028

Abbreviations: gbc, Gradient Boosting Classifier; ada, Ada Boost Classifier; rf, Random Forest Classifier; catboost, CatBoost Classifier; nb, Naive Bayes; lr, Logistic Regression; et, Extra Trees Classifier; lightgbm, Light Gradient Boosting Machine; ridge, Ridge Classifier; lda, Linear Discriminant Analysis; qda, Quadratic Discriminant Analysis; xgboost, Extreme Gradient Boosting; dt, Decision Tree Classifier; knn, K Neighbors Classifier; svm, SVM - Linear Kernel; dummy, Dummy Classifier.

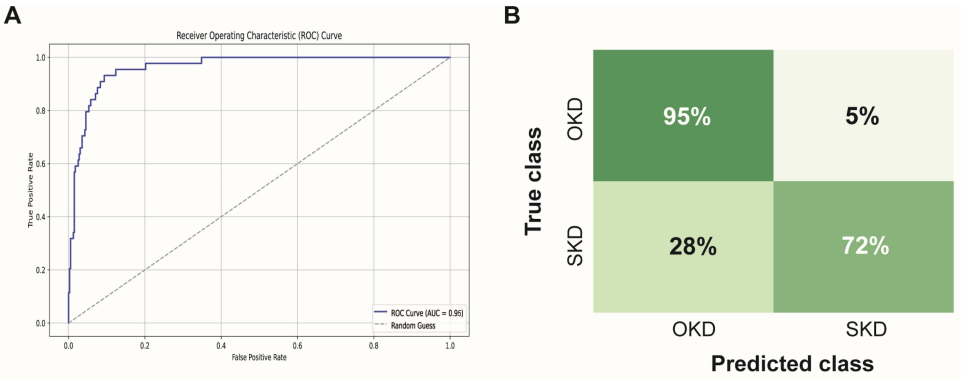

Figure 3 ROC curve and confusion matrix of the Gradient Boosting Classifier (GBC) for SKD diagnosis.(**A**) Receiver Operating Characteristic (ROC) curve of the Gradient Boosting Classifier (GBC) model on the testing set. The ROC curve illustrates the discriminative performance of the GBC model in distinguishing SKD from OKD. The model achieved a high area under the curve (AUC) of 0.952, indicating excellent overall classification capability; (**B**) Confusion matrix of the GBC model on the testing set. The confusion matrix summarizes the classification results of the GBC model. The model demonstrated a sensitivity of 72% for SKD detection and a specificity of 95% for correctly identifying OKD cases, supporting its potential application in clinical screening and decision-making.

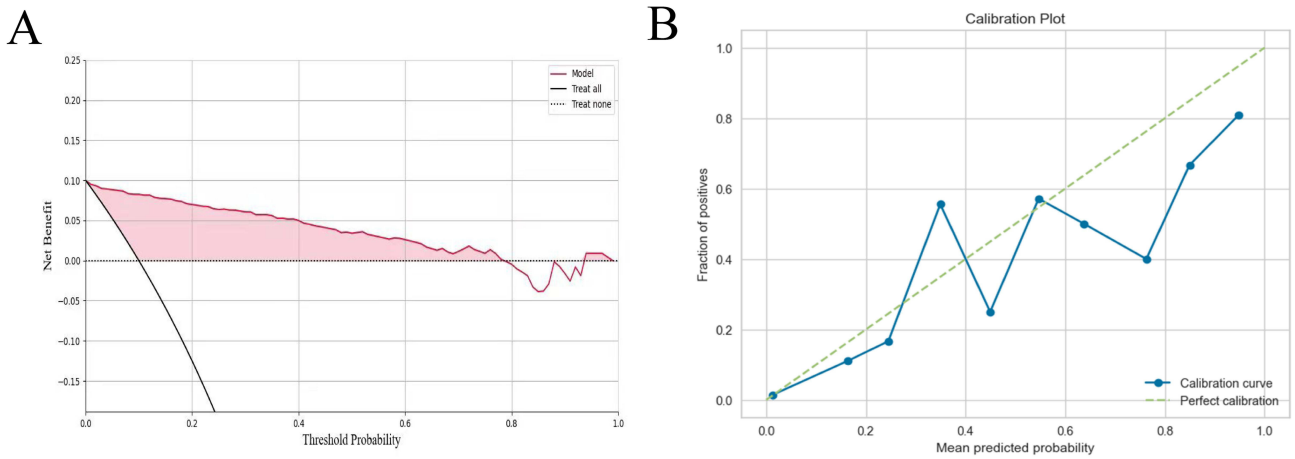

Figure 4 GBC model's clinical utility throughout the threshold probability spectrum. (**A**) DCA of model in the modeling group; (**B**) Calibration curve of model in the modeling group.

Discussion

Currently, clinicians diagnose KD and SKD primarily through clinical signs and echocardiography, as there are no specific laboratory diagnostic methods available. The early clinical features of KD can be subtle, increasing the risk of misdiagnosis with other febrile conditions. In the acute phase, KD can cause significant damage to various systems throughout the body, indicative of SKD.[6,12] The risks associated with misdiagnosis, delayed diagnosis, and missed diagnosis are heightened by the increased likelihood of IVIG resistance and CAL in KD patients, potentially leading to life-threatening situations. Machine learning represents a paradigm of data analysis that is inherently unbiased, utilizing a large number of variables to identify patterns and make predictions. Our approach utilizes machine learning to develop a predictive model based on routine laboratory tests data, specifically designed to identify SKD. This represents the pioneering application of artificial intelligence in differentiating between OKD and SKD. The model exhibits superior discriminatory power, predictive accuracy, and clinical applicability, thereby assisting pediatricians in promptly diagnosing both OKD and SKD during a child's initial presentation. The model incorporates routine laboratory data, including routine blood tests, biochemical profiles, coagulation assays, and lymphocyte typing tests. These tests are widely available in both inpatient and outpatient settings and are cost-effective, making them accessible across healthcare facilities without requiring advanced or specialized laboratories. By utilizing these readily accessible and standardized

https://doi.org/10.2147/JIR.S528341

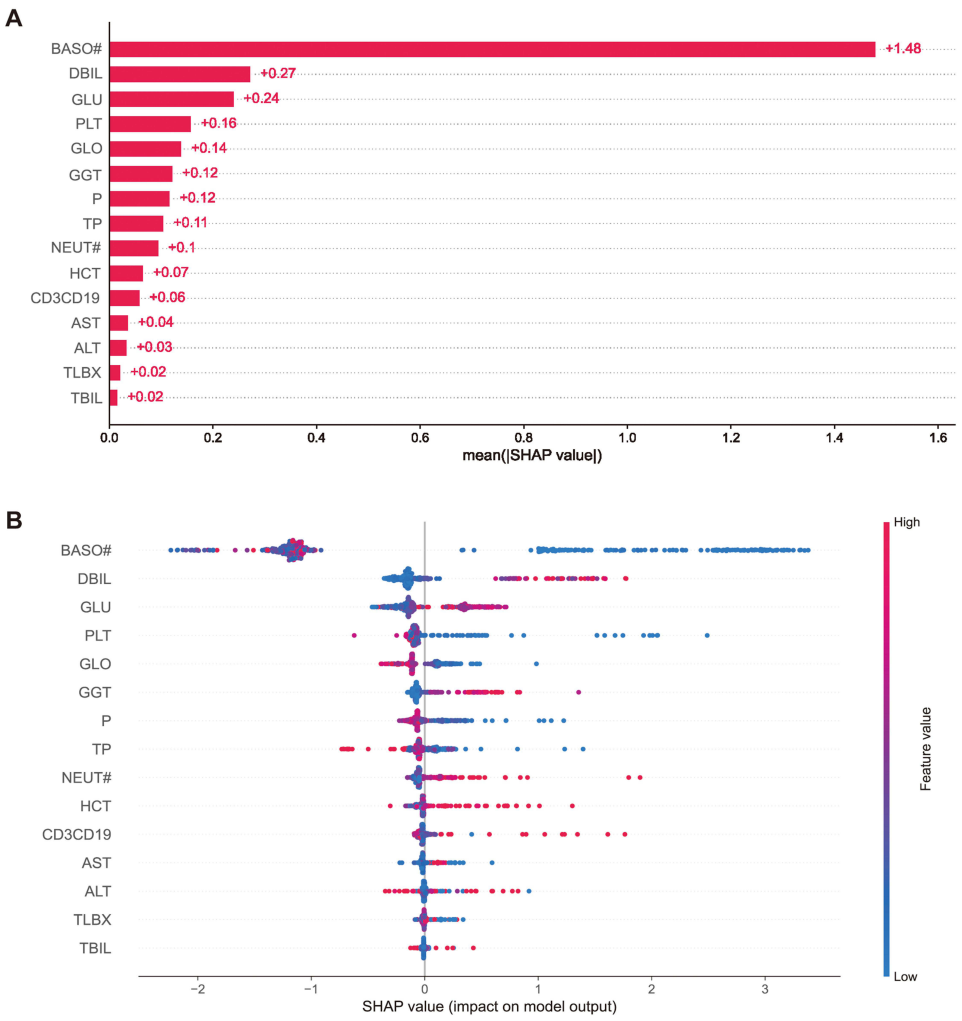

Figure 5 Visualization of the features importance in the Gradient Boosting Classifier model. (**A**) Top 15 predictors by mean |SHAP| value, quantifying their global importance; (**B**) Visualization of the features importance in the Gradient Boosting Classifier model.
Abbreviations: Neut#, absolute neutrophil count; Baso#, absolute basophil count.

tests, the model enhances its potential to be broadly implemented across diverse healthcare environments, ensuring that early and accurate identification of SKD can be achieved nationwide.

Our study presents the following key findings: (1) Of the 69 standard laboratory parameters assessed, only gender, absolute monocyte count, mean corpuscular hemoglobin concentration (MCHC), creatine kinase (CK), creatine kinase-MB (CK-MB), alpha-hydroxybutyrate dehydrogenase (α-HBDH), urea, and the $CD4^+/CD8^+$ ratio showed no significant differences between the SKD and OKD groups, while the remaining 61 parameters exhibited statistically significant differences. (2) Absolute basophil count (BASO#), conjugated bilirubin (DBIL), blood glucose levels (GLU), platelet count(PLT), globulin(GLO), gamma-glutamyltransferase (GGT), phosphorus (P), total protein (TP), absolute neutrophil count (NEUT#), hematocrit (HCT), B-lymphocyte count (CD3CD19), aspartate transaminase (AST), alanine amino-transferase (ALT), T lymphocyte (TLBX), and total bilirubin (TBIL) were independent predictors of predicting SKD. (3) Our study demonstrates that a machine learning model utilizing the Gradient Boosting algorithm can effectively differentiate between OKD and SKD. The confusion matrix shows that the Gradient Boosting model accurately predicted 72% of SKD and 95% of OKD. We employed SHAP to identify the key variables influencing our predictive model. Kawasaki disease is most prevalent in children under five years of age, with SKD patients typically presenting at a younger age than those with OKD. Accurate diagnosis traditionally relies on the expertise of pediatricians in specialized children's hospitals. This newly developed model offers significant advantages to general hospitals lacking

pediatric expertise, as well as to adult hospitals and comprehensive hospitals without access to KD specialists, by enabling rapid and accurate identification of SKD in children. The implementation of this model could substantially reduce the incidence of missed diagnoses, mitigate the risk of KD-associated complications, facilitate timely medical intervention, and enhance overall prognosis.

In recent years, the pursuit of early detection for KD has gained considerable momentum within the scientific community.[11,13,14] Researchers have explored a variety of advanced technologies, including wireless optical monitoring systems,[15] HAMP promoter hypomethylation,[16] and proteome microarrays,[17] to enhance early diagnostic capabilities. Despite the promise of these innovations, they have yet to undergo the necessary clinical validation, thereby limiting their practical application. The investigation of machine learning models for differentiating KD from other febrile conditions has also been a significant focus. These models, based on the clinical presentations and serological data of pediatric patients, have demonstrated impressive levels of sensitivity and specificity. However, they are not free from error, as evidenced by a notable incidence of false-negative results, highlighting a critical area for refinement.[18] Furthermore, efforts have been made to apply machine learning algorithms to distinguish between IVIG-resistant and IVIG-responsive KD.[19,20] While the potential of these models is substantial, their broader clinical implementation has been constrained by factors such as limited sample sizes and geographical variations. In the field of KD diagnosis, our research presents a novel approach that utilizes machine learning models for the first time to differentiate between OKD and SKD. The model we developed exhibits remarkable proficiency, achieving an accuracy rate of 95% for diagnosing OKD and 72% for identifying SKD. Our methodology is notable for its dependence on routine laboratory markers, which are both cost-effective and supported by a comprehensive dataset. This dual advantage significantly enhances the robustness and reliability of our diagnostic model.

In our machine learning model, the five features with the highest SHAP values are the absolute count of basophils, conjugated bilirubin, blood glucose levels, platelet count and globulin at diagnosis. The most distinctive feature is that our results indicate the absolute value of basophils plays a significant role in its performance. Imbalances in the Th17/Treg cell ratio are increasingly recognised as pivotal contributors to the dysregulation of immune function, potentially leading to the development of IVIG-resistant KD.[21] Circulating basophils, which have the capacity for lymphatic trafficking, play a significant role in actively modulating the Treg/Th17 balance mechanism.[22,23] Additionally, the interaction with some soluble factors via membrane receptors positions basophils as potential candidates for gating in cellular diagnostic assays.[24] The cohort study conducted by Chang et al was the first to demonstrate the role of basophils in contributing to mite sensitization in patients with KD.[25] In addition to primarily affecting the heart and coronary arteries, SKD can also impact multiple organ systems throughout the body, including the liver. Bilirubin has been studied for its potential oxidative stress alleviation and inflammatory cascade suppression.[26] Lena et al's large-scale study demonstrated that infection-mediated pathways underlie most hyperbilirubinemia presentations in neonates.[27] Serum bilirubin and hemoglobin serve as sensitive biomarkers for monitoring systemic inflammation in KD.[28] Hyperbilirubinemia occurred in 30.8% of pediatric KD cases, accounting for 120 cases. Among these, increased conjugated bilirubin was recorded in 70.83% (85 cases), while increased unbound bilirubin was noted in 29.17% (35 cases). KD patients with hyperbilirubinemia showed more severe inflammation, potentially attributable to liver damage or biliary obstruction.[28] Diabetes is intricately linked to compromised immune function, often acting as a catalyst for a spectrum of infectious diseases.[29] Elevated blood glucose levels, particularly when they reach or exceed 300 mg/dL, have been correlated with an increased mortality rate in patients suffering from severe sepsis, irrespective of their diabetic status.[30] In a pioneering study by Ji et al, SKD is characterised by the presence of cardiac or cerebrovascular complications and necessitates the administration of intravenous immunoglobulin on at least two occasions. This study also introduces a significant correlation between a family history of cardiovascular diseases and the risk of KD, encompassing conditions such as hypertension, dyslipidaemia, myocardial infarction, stroke, and diabetes.[31] This correlation hints at the potential influence of blood glucose levels in the aetiology of KD. Our predictive model underscores blood glucose levels as a pivotal independent risk factor for SKD. Platelets, along with their secretions, including transcription factors, microparticles, and granules, can exacerbate inflammation and facilitate vascular remodeling.[32] An iconic feature of KD is the increase in platelet count and activation, which functionally contributes to the heightened risk of resistance to IVIG therapy and the formation of CALs. Interestingly, studies have indicated that

dynamic platelet count variations correlating with Kawasaki disease progression timelines. Furthermore, patients with KD who present with abnormally high platelet counts at admission or abnormally low counts after hospitalization are at an increased risk of developing CALs.[33] Our large-scale analysis has demonstrated that platelet counts in patients with SKD are significantly lower than those in patients with OKD. Furthermore, within our predictive model, platelet counts emerges as a vital biomarker for identifying cases of SKD. Higher baseline globulin concentrations predicted greater all-cause mortality in incident hemodialysis patients. This association remained significant even after adjusting for other indicators of malnutrition and inflammation, including albumin levels.[34] Serum globulins, which constitute a significant fraction of total plasma proteins, are elevated in inflammatory conditions. These proteins, primarily comprising immunoglobulins and acute-phase reactants, serve as indicators of the body's inflammatory response. Therefore, increased levels of serum globulins not only reflect this response but also function as valuable biomarkers for the clinical assessment of inflammation.[35] In our study, serum globulin level was an independent prognostic for SKD in the multivariate analysis. Additionally, several factors are associated with an increased risk of developing SKD, including GGT, P, TP, NEUT#, HCT, CD3CD19, AST, ALT, TLBX and TBIL in our machine learning model. Recent research has illuminated several significant risk factors associated with the development of CAL in children with KD. Key indicators include NUET,[36] Hb,[36] the AST/ALT ratio,[37] and RDW.[38] CD3CD19[39] and GGT[40] exhibit a pronounced correlation with the acute phase of KD. The identification of these factors offers valuable insights into disease progression and potential treatment strategies. The decrease in T-lymphocytes indicates an inhibition of cellular immune function, which correlates with the severity of the patients' conditions. Furthermore, HCT and TBIL[41] have emerged as predictive biomarkers for cases that may not respond to the standard treatment of IVIG, highlighting the necessity for alternative therapeutic approaches in IVIG-resistant Kawasaki Disease. It is worth mentioning that serum P were independently related to KDSS in Zhao's research.[42]

Limitations

Our study has some limitations. First, the data for our model was primarily derived from the initial laboratory test data of patients diagnosed with untreated OKD and SKD in the Nanjing area and its surrounding regions. Consequently, it remains uncertain whether this model can yield equally effective predictions for children in other regions or countries. Second, the majority of children with KD require follow-up consultations and are subject to serial laboratory assessments. Our predictive model is based exclusively on the initial, pre-treatment laboratory findings from patients diagnosed with OKD and SKD. It is important to emphasize that the utility of this model is limited to the initial diagnosis and does not extend to forecasting outcomes or predicting relapses. Optimizing the model's sensitivity, however, comes at the cost of reduced specificity in its predictions, an effect that becomes evident during threshold adjustment.

Conclusions

This diagnostic study demonstrates that objective laboratory test results have strong predictive power for distinguishing between OKD and SKD. The development of a predictive model utilizing the GBC machine learning method can aid frontline physicians, especially those with limited pediatric experience, in diagnosing Kawasaki disease in children and in the early identification of severe cases, thereby decreasing the incidence of missed and delayed diagnoses.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This work was supported by Jiangsu Provincial Health Commission (F202155), and China Postdoctoral Science Foundation (2022M721682).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Kawasaki T, Kosaki F, Okawa S, Shigematsu I, Yanagawa H. A new infantile acute febrile mucocutaneous lymph node syndrome (MLNS) prevailing in Japan. *Pediatrics*. 1974;54:271–276. doi:10.1542/peds.54.3.271
2. Singh S, Agarwal S, Bhattad S, et al. Kawasaki disease in infants below 6 months: a clinical conundrum? *Int J Rheum Dis*. 2016;19:924–928. doi:10.1111/1756-185X.12854
3. McCrindle BW, Rowley AH, Newburger JW, et al. Diagnosis, treatment, and long-term management of Kawasaki disease: a scientific statement for health professionals from the American Heart Association. *Circulation*. 2017;135:e927–e999. doi:10.1161/CIR.0000000000000484
4. Fukazawa R. Long-term prognosis of Kawasaki disease: increased cardiovascular risk? *Curr Opin Pediatr*. 2010;22:587–592. doi:10.1097/MOP.0b013e32833e12f7
5. DeHaan LL, Copeland CD, Burney JA, et al. Age-dependent variations in Kawasaki disease incidence in Japan. *JAMA Netw Open*. 2024;7:e2355001.
6. Iwanczak B, Krynicka-Scaringella O. Kawasaki disease in children: epidemiology, clinical symptoms, diagnostics and treatment. *Pol Merkur Lekarski*. 2013;35:375–378.
7. Moffett BS, Syblik D, Denfield S, Altman C, Tejtel-Sexson K. Epidemiology of immunoglobulin resistant Kawasaki disease: results from a large, national database. *Pediatr Cardiol*. 2015;36:374–378. doi:10.1007/s00246-014-1016-1
8. Dionne A, Kokta V, Chami R, Morissette G, Dahdah N. Fatal Kawasaki disease with incomplete criteria: correlation between optical coherence tomography and pathology. *Pediatr Int*. 2015;57:1174–1178. doi:10.1111/ped.12719
9. Kobayashi T, Saji T, Otani T, et al. Efficacy of immunoglobulin plus prednisolone for prevention of coronary artery abnormalities in severe Kawasaki disease (RAISE study): a randomised, open-label, blinded-endpoints trial. *Lancet*. 2012;379:1613–1620. doi:10.1016/S0140-6736(11)61930-2
10. Li C, Liu YC, Zhang DR, et al. A machine learning model for distinguishing Kawasaki disease from sepsis. *Sci Rep*. 2023;13:12553. doi:10.1038/s41598-023-39745-8
11. Lam JY, Shimizu C, Tremoulet AH, et al. A machine-learning algorithm for diagnosis of multisystem inflammatory syndrome in children and Kawasaki disease in the USA: a retrospective model development and validation study. *Lancet Digit Health*. 2022;4:e717–e726. doi:10.1016/S2589-7500(22)00149-2
12. Saha A, Khalil S, Kapoor K, Gupta A, Dubey NK. Kawasaki shock syndrome presenting as toxic shock syndrome. *Int J Rheum Dis*. 2013;16:480–482. doi:10.1111/1756-185X.12119
13. Tsai CM, Lin CR, Kuo HC, et al. Use of machine learning to differentiate children with Kawasaki disease from other febrile children in a pediatric emergency department. *JAMA Netw Open*. 2023;6:e237489. doi:10.1001/jamanetworkopen.2023.7489
14. Zhang M, Ke B, Zhuo H, Guo B. Diagnostic model based on bioinformatics and machine learning to distinguish Kawasaki disease using multiple datasets. *BMC Pediatr*. 2022;22:512. doi:10.1186/s12887-022-03557-y
15. Kuo HC, Lo CC, Lin PX, Kao CC, Huang YH, Lin BS. Wireless optical monitoring system identifies limb induration characteristics in patients with Kawasaki disease. *J Allergy Clin Immunol*. 2018;142:710–711. doi:10.1016/j.jaci.2018.04.018
16. Huang YH, Kuo HC, Li SC, Cai XY, Liu SF, Kuo HC. HAMP promoter hypomethylation and increased hepcidin levels as biomarkers for Kawasaki disease. *J Mol Cell Cardiol*. 2018;117:82–87. doi:10.1016/j.yjmcc.2018.02.017
17. Kuo HC, Huang YH, Chung FH, et al. Antibody profiling of Kawasaki disease using Escherichia coli proteome microarrays. *Mol Cell Proteomics*. 2018;17:472–481. doi:10.1074/mcp.RA117.000198
18. Liu XP, Huang YS, Xia HB, et al. A nomogram model identifies eosinophilic frequencies to powerfully discriminate Kawasaki disease from febrile infections. *Front Pediatr*. 2020;8:559389. doi:10.3389/fped.2020.559389
19. Sunaga Y, Watanabe A, Katsumata N, et al. A simple scoring model based on machine learning predicts intravenous immunoglobulin resistance in Kawasaki disease. *Clin Rheumatol*. 2023;42:1351–1361. doi:10.1007/s10067-023-06502-1
20. Liu J, Zhang J, Huang H, et al. A machine learning model to predict intravenous immunoglobulin-resistant Kawasaki disease patients: a retrospective study based on the Chongqing population. *Front Pediatr*. 2021;9:756095. doi:10.3389/fped.2021.756095
21. Jia S, Li C, Wang G, Yang J, Zu Y. The T helper type 17/regulatory T cell imbalance in patients with acute Kawasaki disease. *Clin Exp Immunol*. 2010;162:131–137. doi:10.1111/j.1365-2249.2010.04236.x
22. Martin-Orozco E, Norte-Munoz M, Martinez-Garcia J. Regulatory T cells in allergy and asthma. *Front Pediatr*. 2017;5:117. doi:10.3389/fped.2017.00117
23. Stelmaszczyk-Emmel A. Regulatory T cells in children with allergy and asthma: it is time to act. *Respir Physiol Neurobiol*. 2015;209:59–63. doi:10.1016/j.resp.2014.11.010
24. Chirumbolo S, Bjorklund G, Sboarina A, Vella A. The role of basophils as innate immune regulatory cells in allergy and immunotherapy. *Hum Vaccin Immunother*. 2018;14:815–831. doi:10.1080/21645515.2017.1417711
25. Chang LS, Huang YH, Chang HY, Lee ZM, Feng WL, Kuo HC. Basophils predict mite sensitization in patients with Kawasaki Disease. *Children*. 2023;11:10. doi:10.3390/children11010010
26. Huang ZW, Shi Y, Zhai YY, et al. Hyaluronic acid coated bilirubin nanoparticles attenuate ischemia reperfusion-induced acute kidney injury. *J Control Release*. 2021;334:275–289. doi:10.1016/j.jconrel.2021.04.033
27. Gottesman LE, Del Vecchio MT, Aronoff SC. Etiologies of conjugated hyperbilirubinemia in infancy: a systematic review of 1692 subjects. *BMC Pediatr*. 2015;15:192. doi:10.1186/s12887-015-0506-5
28. Cheng F, Kang L, Zhang F, et al. Analysis of hyperbilirubinemia in patients with Kawasaki disease. *Medicine*. 2020;99:e21974. doi:10.1097/MD.0000000000021974
29. S RKS, Kaptoge S, Thompson A, et al. Diabetes mellitus, fasting glucose, and risk of cause-specific death. *N Engl J Med*. 2011;364:829–841.

30. Kushimoto S, Gando S, Saitoh D, et al. Impact of serum glucose levels on disease severity and outcome in patients with severe sepsis: an analysis from a multicenter, prospective survey of severe sepsis. *Acute Med Surg*. 2015;2:21–28. doi:10.1002/ams2.59
31. Kwak JH, Ha EK, Kim JH, Cha HR, Lee SW, Han MY. Association of familial history of diabetes, hypertension, dyslipidemia, stroke, or myocardial infarction with risk of Kawasaki Disease. *J Am Heart Assoc*. 2022;11:e023840. doi:10.1161/JAHA.121.023840
32. Noval Rivas M, Kocaturk B, Franklin BS, Arditi M. Platelets in Kawasaki disease: mediators of vascular inflammation. *Nat Rev Rheumatol*. 2024;20:459–472. doi:10.1038/s41584-024-01119-3
33. Ae R, Abrams JY, Maddox RA, et al. Platelet count variation and risk for coronary artery abnormalities in Kawasaki Disease. *Pediatr Infect Dis J*. 2020;39:197–203. doi:10.1097/INF.0000000000002563
34. Pai AY, Sy J, Kim J, et al. Association of serum globulin with all-cause mortality in incident hemodialysis patients. *Nephrol Dial Transplant*. 2022;37:1993–2003. doi:10.1093/ndt/gfab292
35. Hsieh YP, Tsai SM, Kor CT, Chiu PF. Serum globulin is a novel predictor of mortality in patients undergoing peritoneal dialysis. *Sci Rep*. 2023;13:1139. doi:10.1038/s41598-023-27688-z
36. Beiser AS, Takahashi M, Baker AL, Sundel RP, Newburger JW. A Predictive Instrument for Coronary Artery Aneurysms in Kawasaki Disease 11This study was supported, in part, by grants HL34545 and HL48606 from the National Institutes of Health, Bethesda, Maryland, and by the Kobren Fund, Boston, Massachusetts. *Am J Cardiol*. 1998;81(9):1116–1120. doi:10.1016/S0002-9149(98)00116-7
37. Cao L, Tang YJ, Gang M, et al. AST-to-ALT ratio and coronary artery lesions among patients with Kawasaki disease. *World J Pediatr*. 2021;17:659–668. doi:10.1007/s12519-021-00479-0
38. Ming L, Cao HL, Li Q, Yu G. red blood cell distribution width as a predictive marker for coronary artery lesions in patients with Kawasaki disease. *Pediatr Cardiol*. 2021;42:1496–1503. doi:10.1007/s00246-021-02633-x
39. Wang Z, Xie L, Ding G, et al. Single-cell RNA sequencing of peripheral blood mononuclear cells from acute Kawasaki disease patients. *Nat Commun*. 2021;12:5444. doi:10.1038/s41467-021-25771-5
40. Ting EC, Capparelli EV, Billman GF, Lavine JE, Matsubara T, Burns JC. Elevated gamma-glutamyltransferase concentrations in patients with acute Kawasaki disease. *Pediatr Infect Dis J*. 1998;17:431–432. doi:10.1097/00006454-199805000-00020
41. Wang J, Huang X, Guo D. Predictors and a novel predictive model for intravascular immunoglobulin resistance in Kawasaki disease. *Ital J Pediatr*. 2023;49:126. doi:10.1186/s13052-023-01531-7
42. Zhao Z, Yuan Y, Gao L, Li Q, Wang Y, Zhao S. Predicting Kawasaki disease shock syndrome in children. *Front Immunol*. 2024;15:1400046. doi:10.3389/fimmu.2024.1400046