Heliyon 9 (2023) e22202

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Predictive model for early death risk in pediatric hemophagocytic lymphohistiocytosis patients based on machine learning

Li Xiao [a,1], Yang Zhang [b,1], Ximing Xu [c], Ying Dou [a], Xianmin Guan [a], Yuxia Guo [a], Xianhao Wen [a], Yan Meng [a], Meiling Liao [a], Qinshi Hu [c], Jie Yu [a,*]

[a] *Department of Hematology and Oncology, Children's Hospital of Chongqing Medical University, National Clinical Research Center for Child Health and Disorders, Chongqing Key Laboratory of Pediatrics, Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing, China*

[b] *College of Medical Informatics, Chongqing Medical University, Chongqing, China*

[c] *Big Data Center for Children's Medical Care, Children's Hospital of Chongqing Medical University, Chongqing, China*

ARTICLE INFO

Keywords:
machine learning
Predictive modeling
HLH
early death
Interpretability

ABSTRACT

Background: Hemophagocytic Lymphohistiocytosis (HLH) is a rare and life-threatening disease in children, with a high early mortality rate. This study aimed to construct machine learning model to predict the risk of early death using clinical indicators at the time of HLH diagnosis.

Methods: This observational cohort study was conducted at the National Clinical Research Center for Child Health and Disease. Data was collected from pediatric HLH patients diagnosed by the HLH-2004 protocol between January 2006 and December 2022. Six machine learning models were constructed using the Least Absolute Shrinkage and Selection Operator (LASSO) to select key clinical indicators for model construction.

Results: The study included 587 pediatric HLH patients, and the early mortality rate was 28.45 %. The logistic and XGBoost model with the best performance after feature screening were selected to predict early death of HLH patients. The logistic model had an AUC of 0.915 and an accuracy of 0.863, while the XGBoost model had an AUC of 0.889 and an accuracy of 0.829. The risk factors most associated with early death were the absence of immunochemotherapy, decreased TC levels, increased BUN and total bilirubin, and prolonged TT. We developed an online calculator tool for predicting the probability of early death in children with HLH.

Conclusions: We developed the first web-based early mortality prediction tool for pediatric HLH to assist clinicians in risk stratification at diagnosis and in developing personalized treatment protocols. This study is registered on the China Clinical Trials Registry platform (ChiCTR2200061315).

1. Introduction

Hemophagocytic Lymphohistiocytosis (HLH) is a rare and serious hyperinflammatory disease that can rapidly progress and become life-threatening if left undiagnosed and untreated. HLH is usually divided into primary hemophagocytic lymphohistiocytosis (pHLH) and secondary hemophagocytic lymphohistiocytosis (sHLH) [1]。pHLH occurs primarily in the setting of an underlying genetic defect

* Corresponding author. Address: 2 Zhongshan Road, Yuzhong District, Chongqing 400014, China.
E-mail address: yujie@hospital.cqmu.edu.cn (J. Yu).

[1] Co-first authors:Li Xiao, Yang Zhang.

https://doi.org/10.1016/j.heliyon.2023.e22202
Received 12 May 2023; Received in revised form 23 October 2023; Accepted 6 November 2023
Available online 11 November 2023

Abbreviations :

Abbreviation	*Full phrase*
HLH	Hemophagocytic lymphohistiocytosis
FHL	Familial hemophagocytic lymphohistiocytosis
XLP-1	X-linked lymphoproliferative disorder type 1
XLP-2	X-linked lymphoproliferative disorder type 2
XMEN	X-linked immunodeficiency with magnesium defect,Epstein-Barr virus infection,and neoplasia
TC	Total cholesterol
HDL	High-density lipoprotein
LDL	Low density lipoprotein
EBV	EB virus
ALB	Albumin
MAS	Macrophage activation syndrome
AST	Aspartate transaminase
ALT	Alanine transaminase
LDH	Lactate dehydrogenase
GGT	Gamma-glutamyl transpeptidase
BUN	Blood urea nitrogen
Crea	Creatinine
FIB	Fibrinogen
APTT	Activated partial thromboplastin time
PT	Prothrombin time
TT	Thrombin time
INR	International normalized ratio
CNS-HLH	Central nervous system hemophagocytic lymphohistiocytosis
AUC	Area under the receiver operating characteristic curve
CI	Confidence interval
Logistic	Logistic Regression
XGBoost	Extreme Gradient Boosting
GBDT	Gradient Boosting Classifier
LightGBM	Light Gradient Boosting Machine
CatBoost	CatBoost Classifier
SVM	Support Vector Machine
LASSO	Least Absolute Shrinkage and Selection Operator
SHAP	SHapley Additive exPlanations

in immune function, whereas sHLH may be triggered by events such as infections [2,3]、malignancies [4]、rheumatic diseases [5,6]、transplantation [7]and metabolic disorders [8]. The main clinical manifestations of HLH include persistent fever, mono/-multilineage hemocytopenia, hepatosplenomegaly, hyperferritinemia, hepatic impairment, and hematophagocytosis [9]。

HLH is a clinical syndrome with multiple causes and variable changes in clinical presentation and laboratory findings. Some patients with mild cases of infection associated HLH may resolve spontaneously without specific treatment for HLH [10]. The treatment often leads to myelosuppression causing death in children with severe co-infections, and the long-term side effect of Etoposide in the regimen is secondary tumor development [11,12]. In children with severe or refractory relapses, complete remission may be difficult to achieve with the HLH-94/04 chemotherapy regimen, and more aggressive treatment is needed to save the patient's life [13–15]. HLH has a poor prognosis, with a median survival time of only 1.8–2.2 months without appropriate treatment and a long-term survival rate of less than 5 % [16].Early initiation of appropriate treatment is essential to improve survival of patients with this challenging disease [17–19].Therefore, stratifying the severity of children with HLH according to different clinical manifestations and laboratory tests and other findings, and constructing an early mortality prediction model are essential for assessing the progression and severity of HLH, taking timely interventions for high-risk patients, and improving patient survival.

Machine learning is now widely used in disease diagnosis, prognosis prediction, treatment plan design and personalized medicine [20].Due to the relative rarity of HLH, the complexity of the patient's condition, and the often accompanying multisystem damage, machine learning-based early death prediction models for pediatric HLH patients have not been reported at home and abroad. This study relies on the Clinical Science Research Big Data Platform (CSRBDP) of the National Center for Clinical Research on Child Health and Disease and proposes to analyze a large number of clinical characteristics of real-world HLH patients to construct a machine learning model to predict the risk of early death in pediatric HLH and provide a basis for decision making in the stratified and graded management of pediatric HLH patients.

2. Materials and methods

2.1. Study Population and design

This observational retrospective cohort study was conducted at the Children's Hospital of Chongqing Medical University, one of two national clinical research centers for child health and disease in China that completed the Clinical Science Research Big Data Platform (CSRBDP) in 2021, which is available to clinical researchers and includes more than 750,000 pediatric outpatients and inpatients as of January 2022. The inclusion criteria of this study were as follows: ① diagnosis between January 2006 and December 2022; ②patients with newly diagnosed HLH or who have been diagnosed out-of-hospital but have not been treated with hormone or etoposide (VP16) chemotherapy and whose diagnosis meets HLH-04 criteria; and ③ age under 18 years. Exclusion criteria: One of the following conditions was present and was not included in this study: ①clearly tumor- or rheumatology-related HLH; ②clearly diagnosed out-of-hospital with HLH and treated without immunochemotherapy. ③ Hematopoietic stem cell transplantation (HSCT) within 30 days after diagnosis; ④ Patients who were lost to follow-up within 30 days after diagnosis.

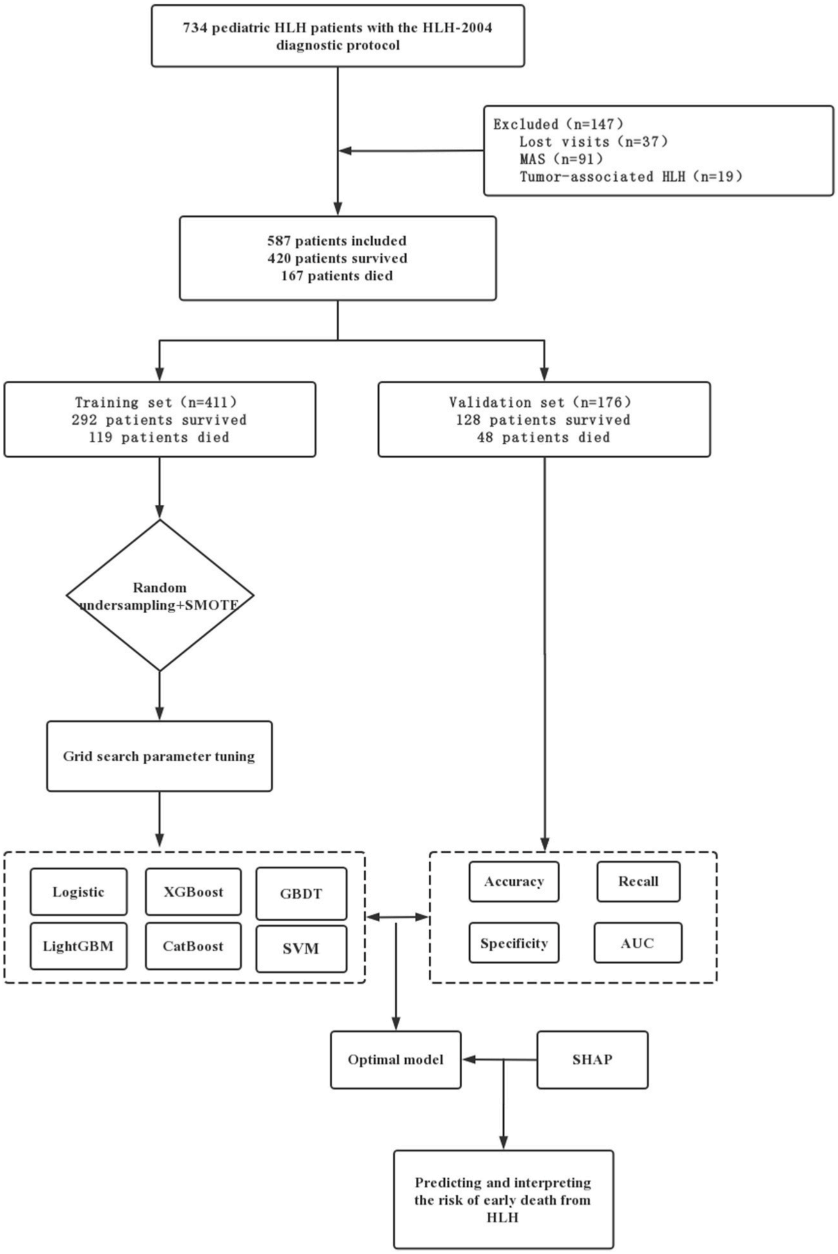

Fig. 1. Machine learning model construction Process.

The study followed the Transparent Reporting of Multivariate Predictive Models for Individual Prognosis or Diagnosis (TRIPOD) specification for model development and validation [21]。Multidimensional clinical data were collected from the CSRBDP, and the study outcome was survival status within 30 days of diagnosis. The study was registered on the Chinese clinical trial registration platform (clinicaltrials.gov, identifier: ChiCTR2200061315) and was approved by the Ethics Committee of the Children's Hospital of Chongqing Medical University with a waiver of the informed consent requirement (File No. 2022,136). Fig. 1 presents an overview of the study design.

2.2. Data collection

A total of 57 clinically relevant indicators of HLH were collected in this study, including demographic information, clinical characteristics, laboratory test information, potential etiology, and survival at 30 days after diagnosis. Demographic information included age at diagnosis and gender. Clinical and laboratory data were collected within 24 h of admission, or if this information was not available, the results of the first collection during hospitalization were used. The underlying etiology was assessed by medical history and laboratory tests. To ensure data accuracy, two clinicians separately interpreted the medical records, and the opinion of the third clinician was adopted in case of disagreement. Closing events were determined by assessing survival within 30 days of diagnosis. The main treatment plan was decided according to the judgment of the attending physician and the treatment wishes of the patient's family and included the following three types:①immunochemotherapy group: including glucocorticoids + etoposide (VP16) ± cyclosporin (CSA); ②hormone therapy group: only glucocorticoids and symptomatic treatment; and ③ symptomatic support therapy.

2.3. Model development

The candidate predictor variables in this study excluded indicators with a missing rate greater than 30 %. For indicators with a missing rate of 30 % or less, the missing values were filled using the nearest neighbor algorithm (K-Nearest Neighbor, KNN) [22]. The least absolute shrinkage and selection operator (LASSO) were applied to identify influential clinical variables (predictor variables with $p < 0.05$) to remove irrelevant and redundant information, select the best predictor variables, and improve the predictive power of machine learning models [23].

Random stratified sampling was used to divide the patients into training and validation sets in a 7:3 ratio. The former was used to train the machine learning models and the latter was used to evaluate the model performance. The training set was used to train six machine learning models, including CatBoost Classifier (CatBoost), Gradient Boosting Decision Tree (GBDT), Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LGBM), Support Vector Machine (SVM), and Logistic Regression (Logistic). The training set was preprocessed using random undersampling combined with SMOTE oversampling techniques to balance positive and negative categories. Grid Search and 5-fold cross-validation were used to find the best hyperparameters for the six models [24]. The best-performing model was chosen for risk prediction based on its combined metrics of AUC, recall, specificity, accuracy,Brier Score, net reclassification improvement (NRI) and integrated discrimination improvement (IDI).Decision curve analysis (DCA) for assessing the clinical significance of predictive models. The predictive power of the best machine learning models was explained using SHapley Additive exPlanations (SHAP) [25]. Logistic models were visualized using a nomograph, and a shinyapp was used to build calculator tools on the web for clinical use by physicians.

2.4. Statistical methods

Statistical analyses were performed using R (version 4.1.2), while predictive model construction and evaluation were conducted using Python (version 3.7). Continuous variables were assessed for normality using the Shapiro-Wilk test, and non-normal continuous variables were presented as median with interquartile range (IQR). Categorical variables were presented as frequencies and percentages (n, %). The Mann-Whitney *U* test was used for continuous variables, and the chi-square test was used for categorical variables to compare the differences in variable distribution between the training and validation cohorts. Statistical significance was defined as a two-sided P value < 0.05.

3. Results

3.1. General patient characteristics

A total of 587 pediatric patients diagnosed with HLH were included in this study, comprising 306 males and 281 females, with a median age of 46 months. Among the 57 indicators considered, those with missing values exceeding 30 % were excluded, resulting in 37 indicators that were analyzed. Table S1 presents a comparison of clinical and laboratory characteristics between patients in the training and validation cohorts at diagnosis. The results indicate that there were no significant differences between the two cohorts, except for LDL-C, Crea, and CD3 and CD19 ratios, which differed significantly.

3.2. Analysis of potential etiology

This study examined 190 (32.37 %, 190/587) patients for HLH-related genes, including PRF1, UNC13D, STX11, STXBP2, AP3B1, LYST, RAB27A, SH2D1A, XIAP, ITK, MAGT1, and CD27. Of these patients, 55 (28.95 %, 55/190) were diagnosed with pHLH, including

30 cases of FHL and 25 cases of other pHLH. No mutations in the HLH related genes were found in the remaining 135 (71.05 %, 135/190) patients.

At the time of diagnosis, 534 patients (90.97 %, 534/587) had evidence of infection, and some patients had two or more pathogenic infections. Among thepathogens, viruses were the most common, with Epstein-Barr virus (EBV) being the most prevalent (85.84 %, 491/572). Of the 491 patients with EBV infection, 197 underwent genetic testing, of whom 39 were confirmed to have pHLHtriggered by EBV infection. Table S2 shows the classification of underlying etiology in the 587 patients.

3.3. Overall survival

A total of 587 pediatric HLH patients were enrolled in this study, and within 30 days of diagnosis, 167 patients died, for a 30-day post-diagnosis mortality rate of 28.45 %. In the training cohort, 119 died in the training cohort, and 292 survived, resulting in a 30-day post-diagnosis mortality rate of 28.95 %. In the validation cohort, 48 patients died, and 128 survived, resulting in a 30-day post-diagnosis mortality rate of 27.27 %. From 2006 to 2022, the 30-day post-discharge mortality rate of pediatric HLH patients showed a decreasing trend (Fig. S1).

3.4. Feature screening

In the LASSO regression for feature screening, 37 candidate predictor variables were included. By adjusting λ, the number of variables was reduced, which resulted in a simplified model. The number of clinical characteristics of pediatric HLH patients decreased continuously with increasing log (λ) (Fig. S2A). By performing 10-fold cross-validation to find the optimal λ value, LASSO selected 15 variables with non-zero regression coefficients at $\lambda = 0.034$ and $\log(\lambda) = -3.395$ (Fig. S2B). The 15 predictor variables that were ultimately included in the follow-up study were TC, total bilirubin, ALB, LDH, BUN, FIB, APTT, TT, INR, bone marrow histiocyte ratio, treatment plan, presence of CNS-HLH, and presence of visceral hemorrhage.

3.5. Model development and screening

The optimal parameters for developing the early mortality prediction model for HLH in children through grid search and cross-validation are presented in Supplementary Table S3. Two types of models were constructed based on the training cohort: six all-variable prediction models that included all 37 variables without LASSO regression feature screening and six post-screening prediction models that included 15 variables using LASSO regression feature screening.

The ROC curves of the prediction model based on machine learning were shown in Fig. 2. Among the six prediction models with full variable inclusion, the Logistic Regression model showed the best prediction performance with the highest AUC of 0.900 (95 % CI: 0.850–0.942), followed by the GBDT model with an AUC of 0.892 (95 % CI: 0.840–0.940) (Fig. 2A). Among the six prediction models constructed after LASSO regression feature screening, it was also the logistic regression model that showed the best predictive performance with the highest AUC of 0.915 (95 % CI: 0.871–0.953), followed by the XGBoost model with an AUC of 0.889 (95 % CI: 0.836–0.941) (Fig. 2B).

In order to comprehensively assess the predictive performance of each model, this study calculated the accuracy, recall, specificity, and brier score of six models before and after LASSO screening and performed clinical decision curve analysis(DCA). Table 1 presents the performance parameters of the six models with all variables included, where the logistic regression model has an accuracy of 0.835, a recall of 0.750, a specificity of 0.867, and a brier score of 0.123. The XGBoost model has an accuracy of 0.812, a recall of 0.750, and a specificity of 0.835, and a brier score of 0.120. Table 2 shows the performance parameters of the six models constructed after LASSO feature screening. The logistic regression model has an accuracy of 0.863, a recall of 0.812, a specificity of 0.882, and a brier score of

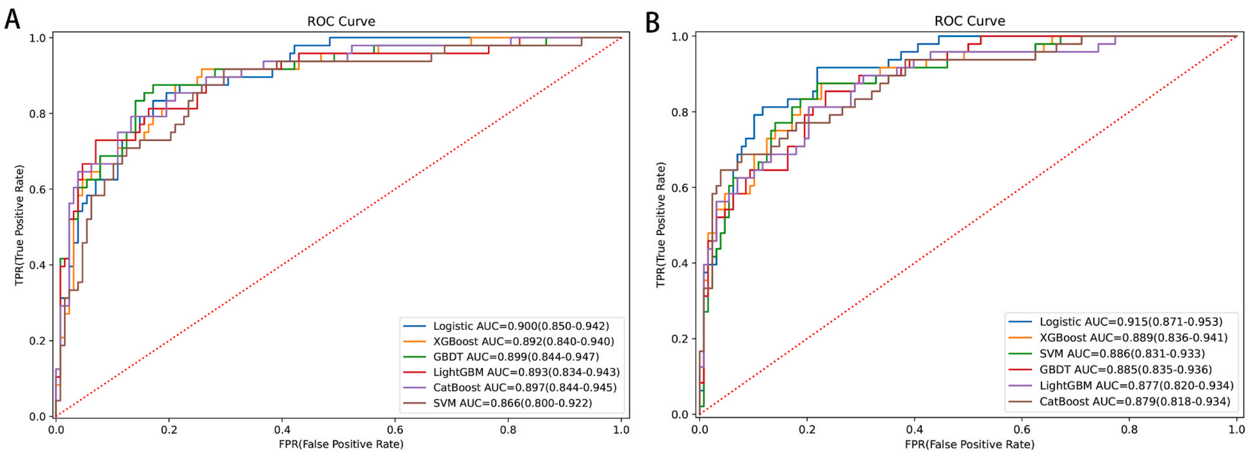

Fig. 2. Six machine learning-based predictive model ROC curves. (A) all variables incorporated; (B) variables incorporated after feature screening.

Table 1
Model performance (all features incorporated).

Model	AUC (95%CI)	Accuracy	Recall	Specificity	Brier Score
Logistic	0.900 (0.850–0.942)	0.835	0.750	0.867	0.123
XGBoost	0.892 (0.840–0.940)	0.812	0.750	0.835	0.120
GBDT	0.899 (0.844–0.947)	0.835	0.729	0.875	0.110
LGBM	0.893 (0.834–0.943)	0.841	0.729	0.882	0.121
CatBoost	0.897 (0.844–0.945)	0.840	0.750	0.875	0.109
SVM	0.866 (0.800–0.922)	0.773	0.771	0.773	0.144

0.115, while the XGBoost model has an accuracy of 0.829, a recall of 0.750, a specificity of 0.859, and a brier score of 0.121.In the clinical decision analysis, when all features were included for model construction, the SVM model had the smallest range of net gains greater than 0, the GBDT had the largest range of net gains greater than 0 (Fig. S3), and the Logistic intervened clinically in the predicted outcomes of the Logistic model at a threshold probability range of 0.051–0.835 (Table S4),At this point the Logistic model resulted in the greatest net benefit. After LASSO regression feature screening, the SVM model had the smallest range of net gains greater than 0, CatBoost had the largest range of net gains greater than 0 (Fig. S4), and Logistic intervened clinically on the predicted results of the Logistic model at a threshold probability range of 0.061–0.933, at which point the Logistic model delivers the greatest net benefit (Table S5).

The net reclassification improvement (NRI) and integrated discrimination improvement(IDI) analyses were performed for the predictive models. In the case of incorporating all the features for model construction, compared with the other five models (XGBoost, GBDT, LGBM, CatBoost, and SVM), the five NRIs of the logistic model are all greater than 0, and the P values are all less than 0.001, indicating that the logistic model prediction model is better than the other five models in local prediction effect; for IDI analysis, except for Logistic vs SVM, the IDIs of the other 4 models are all less than 0, and the P-values are all greater than 0.05, indicating that there is no difference in the overall prediction effect between the Logistic model and the other 4 models (XGBoost, GBDT, LGBM, and CatBoost), whereas for SVM, the IDI >0, and the $P < 0.05$, indicating that Logistic model has better overall prediction effect than SVM model (Table S6). In the case of LASSO regression feature screening, compared with the other five models (XGBoost, GBDT, LGBM, CatBoost, SVM), the five NRIs of the Logistic model are also greater than 0, and $P < 0.001$, indicating that the Logistic model predicts the model better than the other five models (Table S7).

Overall, after LASSO feature screening, the prediction performance of the logistic regression and SVM models improved (0.900 vs. 0.915), (0.866 vs. 0.886), while the prediction performance of the remaining four machine learning models showed a slight decrease (Fig. S5). Therefore, to avoid overfitting and to align with clinical practicality, the feature-screened logistic regression model and XGBoost model were chosen as the final HLH patient mortality risk prediction models.

3.6. Visualization of feature importance

To visually explain the selected variables, this study used SHAP to illustrate how these variables affect early mortality in the model. The feature importance ranking of the five prediction models, namely CatBoost, GBDT, XGBoost, LGBM, and SVM, is shown in Fig. S8. The model importance scores were calculated by the built-in attributes of different machine learning algorithms. Combining the feature importance rankings of the five models, the factors most associated with early death were the treatment plan, TC, BUN, ALB,TT, APTT, and INR.

For the logistic regression model, 15 variables were used to construct Nomogram to visualize the model. Each risk factor's value is assigned a score on the point scale axis. By aggregating each score and using that value on the total scale axis, a total score can be easily calculated to correspond to the probability of early death in individual patients. The Nomogram present the effect of each factor on survival 30 days after diagnosis in patients with HLH, with the most important influences being prolonged TT, increased BUN, and decreased neutrophils (Fig. 3). For easier and faster predictive application in the clinic, a calculator tool was constructed on the web using Shinyapp (https://zheshiyigeceshi.shinyapps.io/DynNomapp/). Screenshots of the calculator tool are shown in Supplement Fig. S7.

Table 2
Model performance (after feature screening).

Model	AUC (95%CI)	Accuracy	Recall	Specificity	Brier Score
Logistic	0.915 (0.871–0.953)	0.863	0.812	0.882	0.115
XGBoost	0.889 (0.836–0.941)	0.829	0.750	0.859	0.121
GBDT	0.885 (0.835–0.936)	0.795	0.708	0.828	0.129
LGBM	0.876 (0.820–0.934)	0.784	0.687	0.820	0.126
CatBoost	0.879 (0.818–0.934)	0.818	0.708	0.859	0.123
SVM	0.886 (0.831–0.933)	0.818	0.813	0.820	0.130

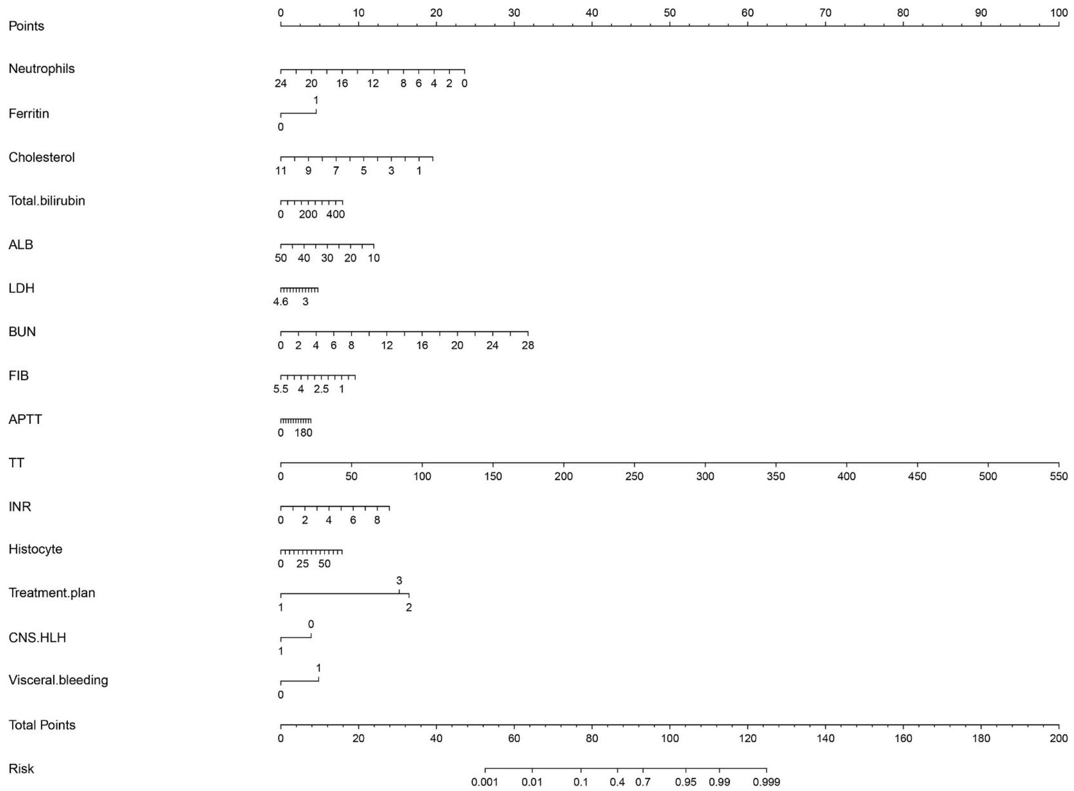

Fig. 3. Nomogram predicting the risk of death in patients with HLH.

3.7. *Interpretation of personalized predictions*

In this study, an analysis was conducted to enhance the interpretability of the XGBoost model by importing the SHAP package into the Tree-Explainer class. The SHAP explanation of the XGBoost model is shown in Fig. 4. Fig. 4A displays the importance ranking of 15 risk factors evaluated by the average absolute SHAP value, while Fig. 4B shows the distribution of the impact of each feature on the model output results. The feature ranking (y-axis) indicates the significance of the prediction model, and the SHAP value (x-axis) is a uniform index that shows the impact of specific features on model output. Each dot in each row represents a patient, and the dot's color represents the feature value: red denotes a higher value, while blue denotes a lower value. Red and blue bars represent risk factors and protective factors, respectively; the longer bars indicate more significant feature importance. The treatment regimen is the most critical influencing factor, and the risk of death is greater in pediatric HLH patients treated with hormone therapy alone and symptomatic support. Early immunochemotherapy considerably lowers the risk of death in HLH patients. Decreased TC, increased BUN, prolonged TT, increased total bilirubin, and increased bone marrow histiocyte ratio all lead to an increased risk of early death in pediatric HLH. The SHAP-dependent plot (Fig. S8) visualizes the top six clinical features that contribute to the prediction of the XGBoost model.

To illustrate the interpretability of the model, we present two typical examples: Fig. 4C portrays a pediatric HLH patient with BUN of 5.29 mmol/L, total bilirubin of 157.7 μmol/L, an immunochemotherapy treatment regimen, TC of 5.11 mmol/L, and a TT of 26 s. In this prediction model, the 30-day early mortality rate was 4.97 %, with a SHAP value of −2.95. In Fig. 4D, a pediatric HLH patient with a symptomatic treatment regimen, TC of 2.29 mmol/L, total bilirubin of 164 μmol/L, and a TT of 20s had a 30-day early mortality rate of 88.99 % in the prediction model, with a SHAP value of 2.09.

4. Discussion

This study is the first to utilize machine learning techniques to identify risk factors for early death in pediatric HLH patients and construct early death prediction models to assess the risk of early death in these patients. Two types of prediction models were constructed for the risk of death in HLH patients 30 days after diagnosis: six prediction models with full variables and six prediction models that were included after screening by LASSO features. The logistic regression model and XGBoost model, after screening for LASSO characteristics, were selected as the prediction models for the risk of death in patients with HLH through comprehensive evaluation. The importance of common features in each model indicated that the most significant factors for early mortality were the absence of immunochemotherapy, decreased TC, increased BUN and total bilirubin, and prolonged TT. To facilitate use, a web app was created to allow clinical staff to input patient-related information, quickly calculating the probability of early death for that patient and

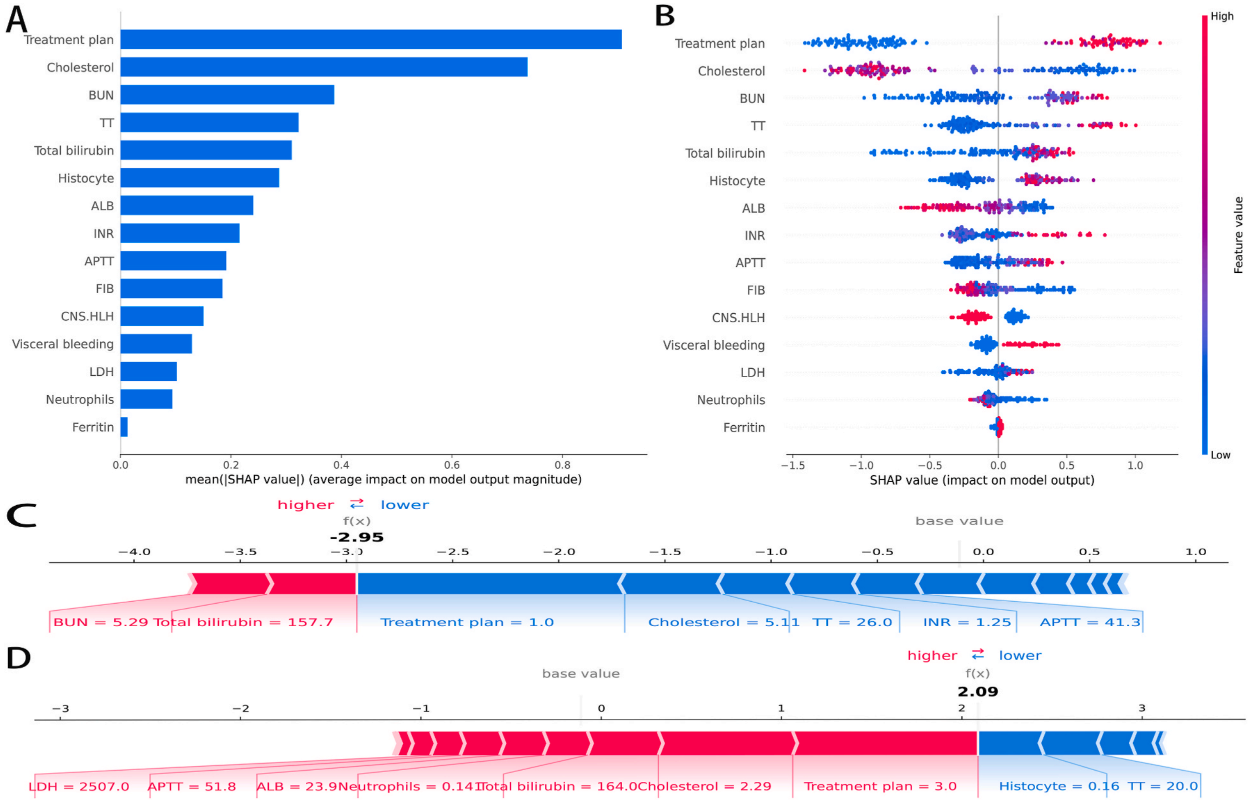

Fig. 4. SHAP interpretation of the XGBoost model. (A) Importance ranking of features that have an impact on model prediction (B) Impact of each feature on model prediction, in each row, each dot represents a patient, and the color of the dot represents the feature value: red represents a larger value, blue represents a lower value. (C, D) Individualized predictions of the model for two patients. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

enabling early intervention, optimal treatment planning, and improved efficiency.

The AUC for the logistic and XGBoost model in this study was 0.915 and 0.889. The better performance of the logistic regression model might be due to the smaller size of the training data, and the 15 clinical indicators selected after LASSO feature selection being more correlated and discriminatory, enabling the logistic regression model to better differentiate between samples from different categories. Frizzell et al. constructed prediction models for all-cause readmission rates 30 days after heart failure discharge using machine learning tree augmented plain Bayesian networks, random forests, gradient boosting algorithms, logistic regression, and LASSO. Their results showed that all five models had similar AUCs: 0.618, 0.607, 0.614, 0.624, and 0.618, with the logistic model being slightly better [26]。

Several studies have shown that machine learning-based models have better performance than prediction models using traditional logistic regression. In 2022, Duan MJ et al. developed a CatBoost model based on machine learning to predict the risk of 30-day unplanned readmission in pediatric pulmonary hypertension patients, with an AUC of 0.81 Compared to the traditional logistic regression algorithm, the machine learning model showed significantly better performance, with an AUC of 0.72 [27]. In 2021, Ke Wang et al. developed a machine learning-based XGBoost risk stratification tool to accurately assess and stratify the 3-year risk of all-cause mortality in patients with heart failure due to coronary heart disease [28]. Although several studies have explored risk factors for early mortality in children with HLH [4–6], machine learning-based risk stratification models have not been reported.

With the emergence of medical informatics, machine learning has become a important tool for intelligently assessing the conditions of patients' diseases. However, due to the opaque nature of machine learning, its application in clinical practice has been limited [29]。SHAP is a game-theoretic technique developed by Lundberg and Lee to address the black-box nature of machine learning by providing consistent interpretability. SHAP values assess the importance of the output for all feature combinations and provide consistent and locally accurate attribute values for each feature in the predictive model [30]。In this study, we performed interpretable analysis of the XGBoost model by importing the Three-Explainer class through the SHAP package. Our results showed that the factors most associated with the risk of death within 30 days after discharge from HLH in children were failure to use immunochemotherapy, reduced TC levels and neutrophils, elevated BUN and total bilirubin, and prolonged TT.

Previous studies have demonstrated that the risk of death within 30 days of diagnosis is lower in pediatric HLH patients treated with immunochemotherapy regimens. The early use of immunochemotherapy is beneficial in improving survival, with more than half of HLH patients treated with the HLH-94 regimen achieving 5-year long-term survival, timely immunochemotherapy regimens can rapidly control inflammatory factors and reduce ensuing organ damage [31]. Our research group has also previously confirmed that decreased TC and elevated BUN are independent risk factors for early death in children with HLH [32]。Lower cholesterol levels may impair the ability to regulate inflammatory immune responses and neutralize endotoxins in pathological conditions, ultimately leading to a storm of uncontrolled inflammatory factors [33]. Elevated BUN indicates that patients with HLH may have more severe circulatory perfusion deficits, infections, or acute kidney injury, which increases the risk of death [34,35]. Total bilirubin is a common indicator of liver impairment, and Honghao Ma et al. also found that high bilirubin levels are a poor prognostic factor for pHLH [36]. Previous studies have shown that coagulation disorders are common in patients with HLH and that abnormal coagulation increases the risk of bleeding and early death [37,38]. It has also been confirmed that severe neutropenia is a prognostic factor for early death in HLH, and that low neutrophil concentrations are suggestive of decreased immune function in patients, which may lead to an increased risk of death.

This study has several limitations. First, the data used in this study were from the National Children's Regional Medical Center (Southwest Region), and more cross-center studies are needed to verify the validity and generalizability of the results. Second, as a retrospective study, this may have affected the model due to some missing information, such as NK cell activity, sCD25 and cytokine assays were not included in the study variables. Finally, external validation was not performed in this study. Future studies should conduct validation in other independent datasets to increase the reliability and generalizability of the results.

5. Conclusion

This study introduces the first machine learning-based model for predicting early death in children with HLH, which exhibits excellent clinical applicability. The model can assist clinicians in making more precise prognostic predictions during early diagnosis and provide data to support early intervention for patients.

Ethics approval and consent to participate

The ethics committee of the Children's Hospital of Chongqing Medical University approved this study protocol (File No. 2022, 136) and waived the need for informed consent of the patients.

Data availability statement

The data analyzed in this study are subject to the following licenses/restrictions: due to ethics committee regulationsthe access to the dataset supporting the findings of this study is limited and not public. Deidentified patient-level data including presenting features, treatment response is available by contacting Dr. Jie Yu (yujie@hospital.cqmu.edu.cn) or Dr. Li Xiao(feeceebee@hospital.cqmu.edu.cn)

Funding

This work was supported by the Intelligence Medicine Project of Chongqing Medical University (YJSZHYX202103 and YJSZHYX202215) .

CRediT authorship contribution statement

Li Xiao: Conceptualization, Data curation, Funding acquisition, Writing – original draft, Writing – review & editing. **Yang Zhang:** Methodology, Validation, Visualization. **Ximing Xu:** Methodology, Writing – review & editing. **Ying Dou:** Data curation, Resources. **Xianmin Guan:** Data curation, Formal analysis. **Yuxia Guo:** Data curation, Formal analysis, Resources. **Xianhao Wen:** Data curation, Formal analysis. **Yan Meng:** Data curation, Formal analysis. **Meiling Liao:** Data curation. **Qinshi Hu:** Data curation. **Jie Yu:** Conceptualization, Methodology, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Not applicable.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e22202.

References

[1] G.E. Janka, Familial and acquired hemophagocytic lymphohistiocytosis, Annu. Rev. Med. 63 (2012) 233–246, https://doi.org/10.1146/annurev-med-041610-134208.
[2] B.U. Kayaaslan, D. Asilturk, F. Eser, et al., A case of Hemophagocytic lymphohistiocytosis induced by COVID-19, and review of all cases reported in the literature, J Infect Dev Ctries 15 (11) (2021) 1607–1614, https://doi.org/10.3855/jidc.14829.
[3] N.G. Rouphael, N.J. Talati, C. Vaughan, et al., Infections associated with haemophagocytic syndrome, Lancet Infect. Dis. 7 (12) (2007) 814–822, https://doi.org/10.1016/S1473-3099(07)70290-6.
[4] N. Daver, K. McClain, C.E. Allen, et al., A consensus review on malignancy-associated hemophagocytic lymphohistiocytosis in adults, Cancer 123 (17) (2017) 3229–3240, https://doi.org/10.1002/cncr.30826.
[5] S. Fukaya, S. Yasuda, T. Hashimoto, et al., Clinical features of haemophagocytic syndrome in patients with systemic autoimmune diseases: analysis of 30 cases, Rheumatology 47 (11) (2008) 1686–1691, https://doi.org/10.1093/rheumatology/ken342.
[6] J.B. Arlet, T.H. Le, A. Marinho, et al., Reactive haemophagocytic syndrome in adult-onset Still's disease: a report of six patients and a review of the literature, Ann. Rheum. Dis. 65 (12) (2006) 1596–1601, https://doi.org/10.1136/ard.2005.046904.
[7] A. Karras, E. Thervet, C. Legendre, Hemophagocytic syndrome in renal transplant recipients: report of 17 cases and review of literature, Transplantation 77 (2) (2004) 238–243, https://doi.org/10.1097/01.TP.0000107285.86939.37.
[8] H. Saneifard, B. Shamsian, M. Shakiba, Z.S. Karizi, A. Sheikhy, A rare case of glycogen storage disease type 1a presenting with hemophagocytic lymphohistiocytosis (HLH), Case Rep Pediatr 2020 (2020), 8818617, https://doi.org/10.1155/2020/8818617.
[9] G.E. Janka, K. Lehmberg, Hemophagocytic syndromes–an update, Blood Rev. 28 (4) (2014) 135–142, https://doi.org/10.1016/j.blre.2014.03.002.
[10] Y. Matsuo, K. Iwanami, E. Hiraoka, R. Oda, Spontaneous recovery of hemophagocytic lymphohistiocytosis due to primary epstein-barr virus infection in an adult patient, Am J Case Rep 22 (2021), e933272, https://doi.org/10.12659/AJCR.933272.
[11] J.I. Henter, A. Horne, M. Aricó, et al., HLH-2004: diagnostic and therapeutic guidelines for hemophagocytic lymphohistiocytosis, Pediatr. Blood Cancer 48 (2) (2007) 124–131, https://doi.org/10.1002/pbc.21039.
[12] J.I. Henter, M. Aricò, R.M. Egeler, et al., HLH-94: a treatment protocol for hemophagocytic lymphohistiocytosis. HLH study Group of the Histiocyte Society, 347, Med. Pediatr. Oncol. 28 (5) (1997) 342, https://doi.org/10.1002/(sici)1096-911x, 199705)28:5<342::aid-mpo3>3.0.co;2-h.
[13] Y. Zhao, Z. Li, L. Zhang, et al., L-DEP regimen salvage therapy for paediatric patients with refractory Epstein-Barr virus-associated haemophagocytic lymphohistiocytosis, Br. J. Haematol. 191 (3) (2020) 453–459, https://doi.org/10.1111/bjh.16861.
[14] Y. Wang, W. Huang, L. Hu, et al., Multicenter study of combination DEP regimen as a salvage therapy for adult refractory hemophagocytic lymphohistiocytosis, Blood 126 (19) (2015) 2186–2192, https://doi.org/10.1182/blood-2015-05-644914.
[15] J. Wang, R. Zhang, X. Wu, et al., Ruxolitinib-combined doxorubicin-etoposide-methylprednisolone regimen as a salvage therapy for refractory/relapsed haemophagocytic lymphohistiocytosis: a single-arm, multicentre, phase 2 trial, Br. J. Haematol. 193 (4) (2021) 761–768, https://doi.org/10.1111/bjh.17331.
[16] B. Degar, Familial hemophagocytic lymphohistiocytosis, Hematol Oncol Clin North Am 29 (5) (2015) 903–913, https://doi.org/10.1016/j.hoc.2015.06.008.
[17] H. Trottestam, A. Horne, M. Aricò, et al., Chemoimmunotherapy for hemophagocytic lymphohistiocytosis: long-term results of the HLH-94 treatment protocol, Blood 118 (17) (2011) 4577–4584, https://doi.org/10.1182/blood-2011-06-356261.
[18] E. Bergsten, A. Horne, M. Aricó, et al., Confirmed efficacy of etoposide and dexamethasone in HLH treatment: long-term results of the cooperative HLH-2004 study, Blood 130 (25) (2017) 2728–2738, https://doi.org/10.1182/blood-2017-06-788349.
[19] R. Machowicz, G. Basak, How can an internal medicine specialist save a patient with hemophagocytic lymphohistiocytosis (HLH)? Pol. Arch. Intern. Med. 130 (5) (2020) 431–437, https://doi.org/10.20452/pamw.15226.
[20] A.K. Waljee, P.D. Higgins, Machine learning in medicine: a primer for physicians, Am. J. Gastroenterol. 105 (6) (2010) 1224–1226, https://doi.org/10.1038/ajg.2010.173.
[21] G.S. Collins, J.B. Reitsma, D.G. Altman, K.G. Moons, Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement, BMJ 350 (2015) g7594, https://doi.org/10.1136/bmj.g7594.
[22] L. Beretta, A. Santaniello, Nearest neighbor imputation algorithms: a critical evaluation, BMC Med Inform Decis Mak 16 (Suppl 3) (2016) 74, https://doi.org/10.1186/s12911-016-0318-z.
[23] J. Kang, Y.J. Choi, I.K. Kim, et al., LASSO-based machine learning algorithm for prediction of lymph node metastasis in T1 colorectal cancer, Cancer Res Treat 53 (3) (2021) 773–783, https://doi.org/10.4143/crt.2020.974.
[24] A. Doğru, S. Buyrukoğlu, M. Arı, A hybrid super ensemble learning model for the early-stage prediction of diabetes risk, Med. Biol. Eng. Comput. 61 (3) (2023) 785–797, https://doi.org/10.1007/s11517-022-02749-z.
[25] S.M. Lundberg, G. Erion, H. Chen, et al., From local explanations to global understanding with explainable AI for trees, Nat. Mach. Intell. 2 (1) (2020) 56–67, https://doi.org/10.1038/s42256-019-0138-9.
[26] J.D. Frizzell, L. Liang, P.J. Schulte, et al., Prediction of 30-day all-cause readmissions in patients hospitalized for heart failure: comparison of machine learning and other statistical approaches, JAMA Cardiol 2 (2) (2017) 204–209, https://doi.org/10.1001/jamacardio.2016.3956.
[27] M. Duan, T. Shu, B. Zhao, et al., Explainable machine learning models for predicting 30-day readmission in pediatric pulmonary hypertension: a multicenter, retrospective study, Front Cardiovasc Med 9 (2022), 919224, https://doi.org/10.3389/fcvm.2022.919224.
[28] K. Wang, J. Tian, C. Zheng, et al., Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP, Comput. Biol. Med. 137 (2021), 104813, https://doi.org/10.1016/j.compbiomed.2021.104813.
[29] D. Yu, H. Wu, Variable importance evaluation with personalized odds ratio for machine learning model interpretability with applications to electronic health records-based mortality prediction, Stat. Med. (2023), https://doi.org/10.1002/sim.9642.
[30] S.M. Lundberg, B. Nair, M.S. Vavilala, et al., Explainable machine-learning predictions for the prevention of hypoxaemia during surgery, Nat. Biomed. Eng. 2 (10) (2018) 749–760, https://doi.org/10.1038/s41551-018-0304-0.
[31] J.H. Yoon, S.S. Park, Y.W. Jeon, et al., Treatment outcomes and prognostic factors in adult patients with secondary hemophagocytic lymphohistiocytosis not associated with malignancy, Haematologica 104 (2) (2019) 269–276, https://doi.org/10.3324/haematol.2018.198655.
[32] L. Xiao, X. Xu, Z. Zhang, et al., Low total cholesterol predicts early death in children with hemophagocytic lymphohistiocytosis, Front Pediatr 10 (2022), 1006817, https://doi.org/10.3389/fped.2022.1006817.
[33] B.R. Gordon, T.S. Parker, D.M. Levine, et al., Relationship of hypolipidemia to cytokine concentrations and outcomes in critically ill surgical patients, Crit. Care Med. 29 (8) (2001) 1563–1568, https://doi.org/10.1097/00003246-200108000-00011.
[34] F. Aulagnon, N. Lapidus, E. Canet, et al., Acute kidney injury in adults with hemophagocytic lymphohistiocytosis, Am. J. Kidney Dis. 65 (6) (2015) 851–859, https://doi.org/10.1053/j.ajkd.2014.10.012.
[35] S. Wang, J. Zhou, J. Yang, et al., Clinical features and prognostic factors of acute kidney injury caused by adult secondary hemophagocytic lymphohistiocytosis, J. Nephrol. 35 (4) (2022) 1223–1233, https://doi.org/10.1007/s40620-021-01147-2.
[36] H. Ma, R. Zhang, L. Zhang, et al., Treatment of pediatric primary hemophagocytic lymphohistiocytosis with the HLH-94/2004 regimens and hematopoietic stem cell transplantation in China, Ann. Hematol. 99 (10) (2020) 2255–2263, https://doi.org/10.1007/s00277-020-04209-w.
[37] Y. Zhao, D. Lu, S. Ma, et al., Risk factors of early death in adult patients with secondary hemophagocytic lymphohistiocytosis: a single-institution study of 171 Chinese patients, Hematology 24 (1) (2019) 606–612, https://doi.org/10.1080/16078454.2019.1660458.
[38] X. Li, H. Yan, X. Zhang, et al., Clinical profiles and risk factors of 7-day and 30-day mortality among 160 pediatric patients with hemophagocytic lymphohistiocytosis, Orphanet J. Rare Dis. 15 (1) (2020) 229, https://doi.org/10.1186/s13023-020-01515-4.