WILEY | Hindawi

Research Article

Development and Internal Validation of Machine Learning Algorithms for Determining Sodium Valproate Concentrations below the Standard Level Using a Risk Prediction Model of Children with Epilepsy

Yinhui Yao [ID]**,**[1] **Jingyi Zhao** [ID]**,**[2,3] **Ying Wang** [ID]**,**[1] **Wei Qiu,**[4] **Yingxue Lin,**[1] **and Yaru Zang** [ID][1]

[1]*Department of Pharmacy, The Affiliated Hospital of Chengde Medical University, Chengde 067000, China*
[2]*Hebei Province (Chengde) Chinese Herbal Medicine Industry Technology Research Institute, Chengde Medical University, Chengde 067000, China*
[3]*Hebei Province Key Laboratory of Study and Exploitation of Chinese Medicine, Chengde Medical University, Chengde 067000, China*
[4]*Department of Laboratory, The Affiliated Hospital of Chengde Medical University, Chengde 067000, China*

Correspondence should be addressed to Yaru Zang; cyfyzyr0314@163.com

Background. There is a narrow therapeutic window for sodium valproate, and the blood concentration is too low to control epilepsy, while it is easy to poison the body if the concentration is too high. It is therefore necessary to monitor drug concentration reasonably in order to control epilepsy. The purpose of this study was to establish a model for predicting concentrations of sodium valproate below 50 μg/mL in children with epilepsy. *Methods.* The clinical data and biochemical examination results of children with epilepsy treated in the pediatric outpatient department of our hospital from June 2019 to March 2022 were retrospectively collected and divided into a development group and a validation group according to a patient ratio of 8 to 2. Five machine learning algorithms were used to identify the key variable factors, and a risk prediction model for sodium valproate blood concentrations lower than the standard concentration was established. The area under the curve (AUC), calibration curve, GiViTi calibration band, and clinical influence curve were used to evaluate the diagnostic efficacy and clinical application value of the model. *Results.* A total of 525 children with epilepsy were enrolled. In the development group, the random forest algorithm performed best in predicting that the blood concentration of sodium valproate was lower than the standard concentration, showing the highest AUC (1.00). Six factors were determined as a nomogram to predict the incidence of low concentrations. In the validation group and the development group, the calibration curve, GiViTi calibration band, and clinical influence curve all performed well in the evaluation of the diagnostic efficacy and clinical application value of the model. *Conclusions.* This finding highlights the importance of examining biochemical indices in patients when data regarding the blood concentration of sodium valproate are lacking.

1. Introduction

Epilepsy is a chronic disease of the central nervous system characterized by recurrent, paroxysmal, and transient dysfunction of the central nervous system caused by the excessive discharge of neurons in the brain [1]. The prevalence rate of this disease in children is high, and most of them need to take antiepileptic drugs for a long time to control or prevent seizures [2]. Sodium valproate, as a broad-spectrum antiepileptic drug commonly used in the clinic, is the first-line therapeutic drug for the treatment of major seizures, minor seizures, and myoclonic seizures in children with epilepsy, and it has a remarkable curative effect [3, 4]. However, the blood concentration of valproate should be

maintained within the effective therapeutic range (50–100 μg/mL). Low blood concentrations of sodium valproate indicate limited ability to control the disease, and high blood concentrations can easily lead to intellectual disability or memory impairment in children. Therefore, monitoring the blood concentration of sodium valproate is an effective method to ensure curative effects and safety and realize individualized treatment [5]. Therefore, this study adopted a retrospective analysis method to collect and analyze the results of the monitoring of the serum concentration of sodium valproate in children with epilepsy in our hospital and used a machine learning method to analyze the relationship between age, sex, dosage form, experimental examination indices, and blood concentration to provide a reference for the rational clinical application of sodium valproate in children with epilepsy.

2. Methods

2.1. Study Design. A retrospective analysis of epileptic children admitted to the Department of Pediatrics of the Affiliated Hospital of Chengde Medical University from June 2019 to March 2022 was conducted. The research was approved by the Ethics Committee of the Affiliated Hospital of Chengde Medical University (No. LL2020012). The inclusion criteria were as follows: (a) children diagnosed with epilepsy according to the International Anti-Epilepsy Alliance (ILAE) criteria; (b) children with good medication compliance and stable blood concentration of sodium valproate after more than 7 days and with blood samples that were collected before the last medication to monitor the valley concentration of sodium valproate; and (c) patients with complete clinical medical records. In this study, according to the results of the monitoring of blood drug concentrations, patients with concentrations below 50 μg/mL were classified as the nonstandard concentration group, and patients with levels within the effective concentration range of 50–100 μg/mL were classified as the standard group. The exclusion criteria were as follows: (a) patients with incomplete clinical records; (b) patients not taking medicine according to the doctor's advice; (c) patients not reaching the steady-state blood concentration within 5 days after taking the medicine; (d) patients taking sodium valproate in the morning and who had blood taken for monitoring; and (e) patients with blood drug concentrations greater than 100 μg/mL.

The clinical information of the children assessed included age, sex, dosage of sodium valproate, dosage form of sodium valproate, drugs taken in combination, blood concentration of sodium valproate, and laboratory examination information (white blood cell count, red blood cell count, hemoglobin level, hematocrit, platelet count, percentage of neutrophils, percentage of lymphocytes, percentage of monocytes, percentage of eosinophils, percentage of basophils, absolute value of neutrophils, absolute value of lymphocytes, absolute value of monocytes, absolute value of eosinophils, absolute value of basophils, average volume of red blood cells, average hemoglobin content, average hemoglobin concentration, red blood cell distribution width

coefficient of variation, red blood cell distribution width standard deviation, average platelet volume, platelet distribution width, large platelet ratio, thrombocytocrit, total protein, albumin level, total bilirubin level, serum total bile acid level, glutamic-pyruvic transaminase level, glutamic-oxaloacetic transaminase level, γ-glutamyltransferase level, alkaline phosphatase level, blood urea nitrogen level, creatinine level, uric acid level, and bicarbonate level). Finally, the child was asked whether the clinical symptoms of epilepsy were controlled during the treatment with medication.

2.2. Concentration Measurement. All patients reached a steady state after taking medicine for 7 days. The sample was obtained from 3 mL of blood collected from the patient before the last medication. The sample was centrifuged at 3500 r/min for 5 min, and then 25 μL of serum was taken. Then, a sodium valproate determination kit from Siemens was used. Finally, the valley concentration of valproic acid was quantitatively determined by the Siemens ADVIA Centaur CP system.

2.3. Statistical Analysis. Through the CBCgrps package, the clinical features and experimental data of the nonstandard concentration group and the standard concentration group in this study were compared [6]. The average and standard deviation were used to describe continuous data, and classified data were recorded as quantity and percentage. Significance was defined as results with a p value of <0.05. This study was conducted by Spearman to determine whether clinical information and outcome were correlated.

Children were randomly divided into a training set (80%) and a verification set (20%) by the createDataPartition function in the caret package. If the number of seeds is fixed, createDataPartition will have random splitting proportional to the number of outcome variables. The model was developed using the training set (80% of data) and internally verified by the verification set (20% of data). Five machine learning algorithms were used to select key features and establish risk prediction models: random forest (RF), support vector machine (SVM), gradient boosting machine (GBM), generalized linear model (GLM), and neural network (NNET). The machine learning approach utilized R packages, including "randomForest" and "caret" [7, 8]. To ensure the data segmentation in modeling and the repeatability of machine learning, a fixed seed number was set when R software was run. It is completed by NNET, GBM, GBM and GLM, which are all included in the caret package. It is selected among the four machine learning methods to run the method parameter in the train function. For other parameters, defaults are used. A comparison is made between the results of the above four methods and those of RF. During this process, in the method used for the random forest model, repeatedcv was adopted by setting the train function, performing 5-fold cross-validation, using a repeated value of 3, and selecting grid in the search. Other default values were used. To determine which data analysis method is most suitable for this study, the following two results are primarily considered by DALEX package: reverse cumulative distribution of residual and boxplots of residual.

As a result of hyperparameter optimization, the following steps are taken. Firstly, the randomForest package sets the parameters ntree, mtry, and nodesize for the randomForest reference function. Then set mtry from 2 to the maximum value (the square root of the total number of dependent variables) by using the tuneGrid function of carte package. As a result, we obtain the out-of-bag (OOB) estimate of error rate. After determining the main risk factors that affected low blood drug concentration by machine learning, the variance inflation factor (VIF) was used to exclude the factors with VIF values greater than 10 [9].

Then, using the "rms" package in R software, we established a nomogram model based on the risk factors obtained by the selected machine learning candidates to predict the incidence of nonstandard concentrations in children. For the dynamic nomogram model, the "dynnom" R package was adopted [10]. The area under the receiver operator characteristic (ROC) curve (AUC), calibration curve, and GiViTi calibration band were used to evaluate the consistency between our predicted value and the actual value [11, 12]. Decision curve analysis (DCA) was performed, and a clinical influence curve was drawn to evaluate whether the decision based on this model is beneficial to patients [13]. The model was verified using 20% of the data. The variables used in the validation data were those that were previously identified in the modeling group. The method used to evaluate the validation data was the same as that used to evaluate the model data. R software version 4.0.2 was used for all statistical analyses.

3. Results

3.1. Characteristics of Patients with Childhood Epilepsy. First, 748 patients with blood concentrations of sodium valproate were identified, and 571 patients remained after removing the data of 177 patients with incomplete clinical data. Then, 46 patients with blood concentrations of sodium valproate greater than 100 μg/mL were excluded. According to the range of the blood concentration of sodium valproate, 525 children were divided into two groups: standard concentration group ($n = 377$) and low concentration group ($n = 148$) (see Table S1). Among the 525 patients in the study, 49 independent variables were taken into account. An analysis of Spearman correlation found 18 independent variables correlated with outcome groups ($p < 0.05$). There are eight independent variables (dosage, percentage of lymphocytes, percentage of monocytes, average volume of red blood cells, average hemoglobin content, glutamic pyruvic transaminase, glutamic oxaloacetic transaminase, and γ-glutamyltransferase) that are positively correlated with the outcome and ten independent variables (carbamazepine, red blood cell count, hemoglobin, hematocrit, platelet count, percentage of neutrophils, absolute value of neutrophils, thrombocytocrit, total protein, and alkaline phosphatase) that are negatively correlated with it (Figure 1).

Finally, a total of 525 cases were selected, among which 421 and 104 patients were divided into a development group and verification group, respectively. Based on the data randomly partitioned by the createDataPartition function,

the proportion of individuals with standard concentration to those with low concentration remains approximately 7 : 3 in both the development group and the verification group. Furthermore, the ratio of standard concentration to low concentration in this partitioned dataset closely aligns with the ratio observed in the undivided total grouping. A total of 29% of children with epilepsy had a blood concentration of sodium valproate lower than the standard concentration in the development group, and the value was 23% in the verification group. Table S2 lists the laboratory examination results and demographic characteristics of the patients in this study. There was no significant difference in the baseline characteristics or intraoperative variables between the development group and the verification group (except in red blood cell count). Based on these, it seems reasonable and analytical to split data randomly. There was a statistically significant difference between the two groups in the time the patients took the medicine to the time of the hospital examination, regardless of whether the epileptic symptoms of patients were controlled ($p = 0.04$). The average drug concentration in blood in the children with controlled symptoms was 60.11 (47.29, 72.59), while that in children with uncontrolled epilepsy was slightly higher, with an average drug concentration in blood of 67.61 (52.31, 78.03).

3.2. Development of Machine Learning Algorithms. We established RF, SVM, GBM, GLM, and NNET and selected potential factors from 49 variables to predict the occurrence of valproate plasma concentrations below the standard concentration. "Reverse cumulative distribution of residual" (Figure 2(a)) and "boxplots of residual" (Figure 2(b)) both reveal that the RF model exhibited the smallest residuals. The residual error of most samples in the model is relatively small, which indicates that the model is good. Therefore, the RF model is considered the best model for predicting the occurrence of sodium valproate blood concentrations lower than the standard concentration. The number of variables used was as small as possible, and the out-of-band error was as low as possible. The hyperparameter optimization was carried out to improve the prediction performance of the RF model by setting the mtry parameter from 2 to 7, ntree to 3000, and nodesize to 5. Last but not least, when ntree equals 500 and mtry equals 4, OOB estimate of error rate is 27.1%. As shown in the diagram illustrating the relationship between the model error and the number of decision trees, we selected 3000 trees as the parameters of the final model, which indicates the stability error in the model (Figure 3(a)).

After sorting these variables in RF according to their importance, we visualized 30 variables (Figure 3(b)). However, the top 10 (γ-glutamyltransferase level, red blood cell count, alkaline phosphatase level, platelet count, percentage of lymphocytes, red blood cell distribution width standard deviation, dose, blood urea nitrogen level, percentage of neutrophils, and thrombocytocrit) of the 30 variables were selected as the candidate factors. The multicollinearity test of these 10 factors showed that the VIF of each factor was less than 10 when six of the current factors were included. Finally, an ROC curve was drawn to evaluate
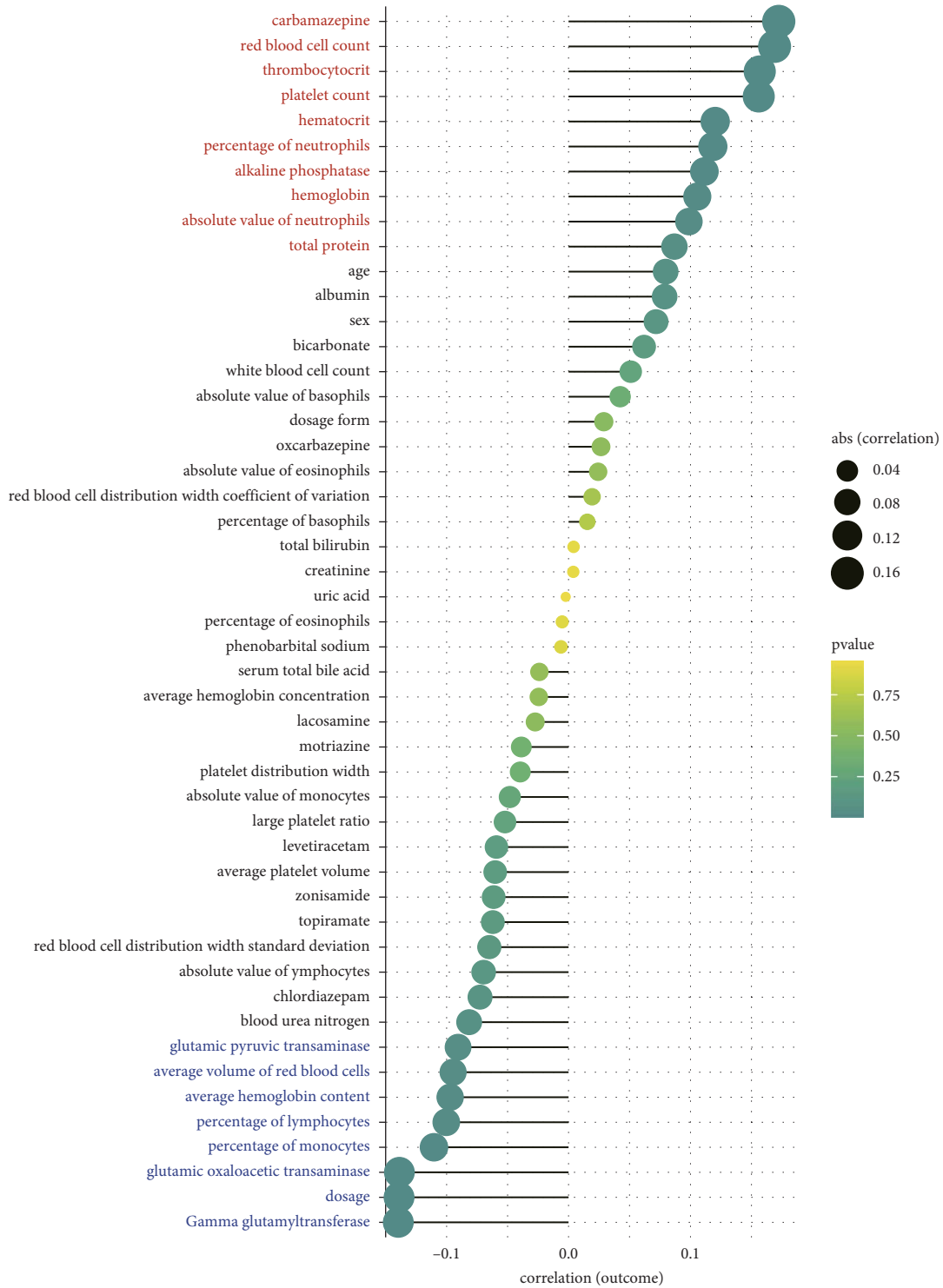
FIGURE 1: Analysis of the correlation between the independent variables and the outcome.

the model, and the AUC value of the ROC curve also showed that the RF model had higher accuracy than that of the other models (Figure 3(a)). In comparison of the AUC achieved using the five algorithms, the AUC achieved using RF and SVM ($Z = 5.17$, $p < 0.001$) further illustrated the superiority of the RF model. Table 1 shows the evaluation index of machine learning methods, indicating that RF is the best.

3.3. Establishment and Verification of the Nomogram. In the development group, a nomogram based on 6 candidate variables was generated by using the "rms" package in R to predict the incidence of valproate plasma concentrations lower than the standard concentration (Figure 4(c)). The calibration curve revealed the predictability of the nomogram model (Figure 5(a)). The GiViTi calibration band
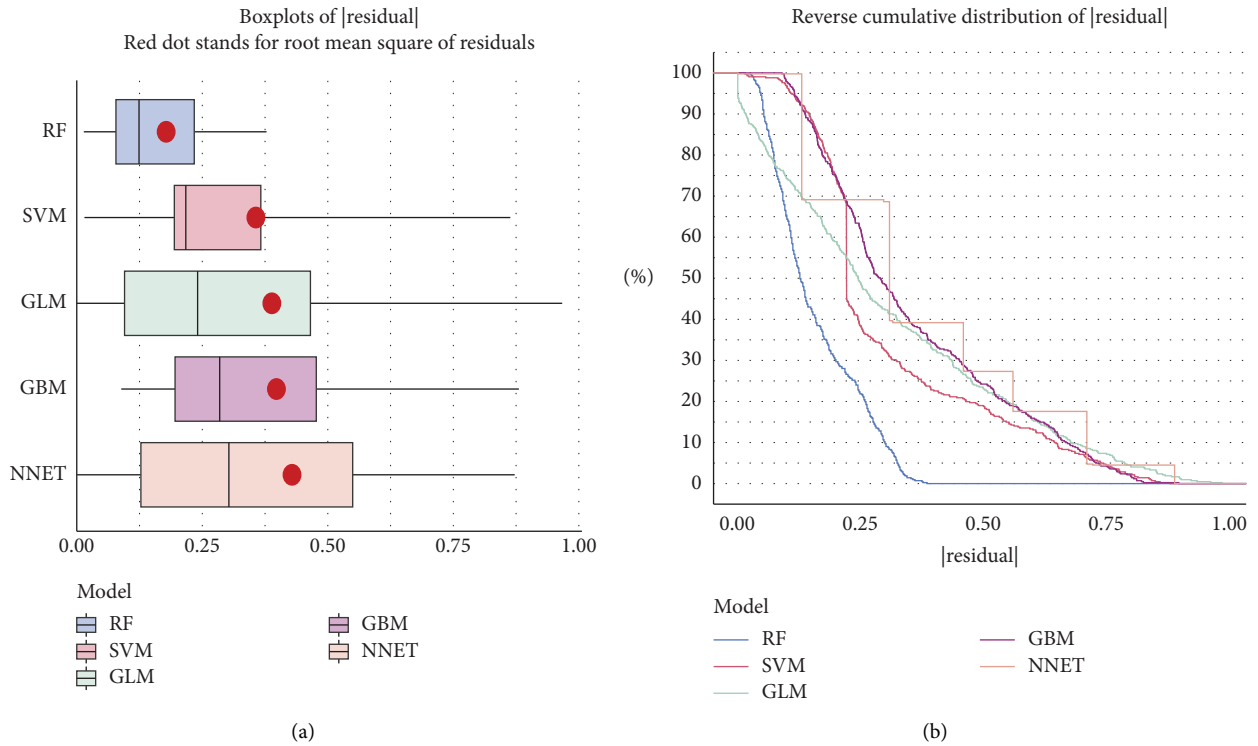
(a)

(b)

Figure 2: Development of machine learning algorithms. (a) Reverse cumulative distribution of residuals was plotted to show the residual distribution of random forest (RF), support vector machine (SVM), gradient boosting machine (GBM), generalized linear model (GLM), and neural network (NNET). (b) Boxplots of residuals were plotted to show the residual distribution of the RF, SVM, GBM, GLM, and NNET models.
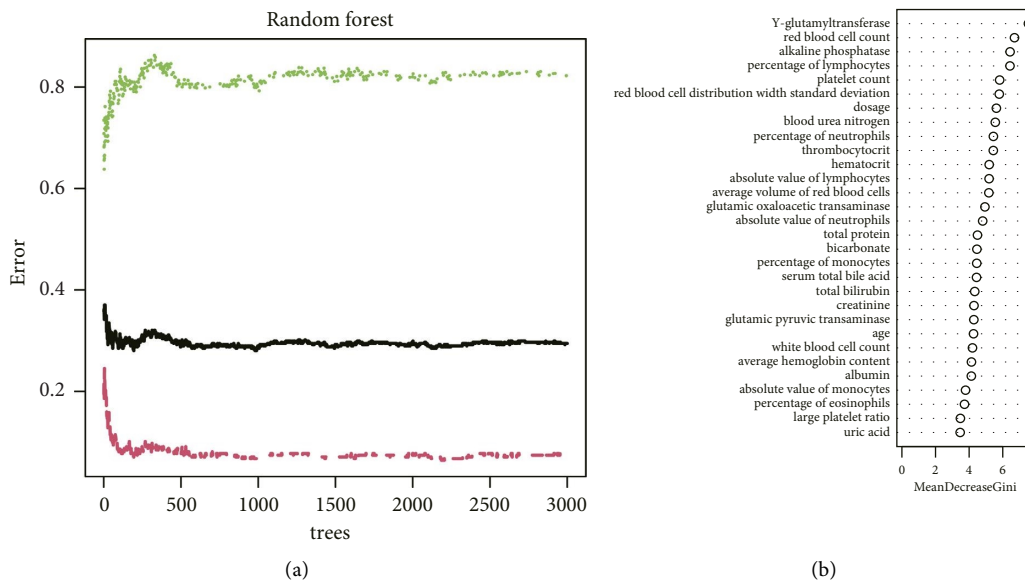

(a)

(b)

Figure 3: Random forest (RF) model establishment. (a) The relationship between the number of decision trees and the model error. The *x*-axis represents the number of decision trees, and the *y*-axis represents the error rate of the established model. (b) The importance of the 30 variables based on the RF model.

showed that the model established by the 6 included factors fit well ($p = 0.944$) (Figure 5(b)). The red line in the DCA curve remained above the gray line and black line from 0.01 to 0.6, indicating that the decision made based on the nomogram model may be beneficial to the prediction of concentrations lower than the standard level (Figure 5(c)). The clinical influence curve showed that the predictive ability of the nomogram model was very significant

TABLE 1: Forecast results for development group and validation group of machine learning algorithms.

	Sen	Spe	TP	FN	FP	TN	Accuracy	Precision	Recall	F1 score
Development group (standard/low = 297/124)										
RF	1.00	1.00	124	0	0	297	1.00	1.00	1.00	1.00
SVM	0.806	0.771	96	28	58	239	0.796	0.623	0.774	0.691
GLM	0.718	0.636	79	45	84	213	0.694	0.485	0.637	0.551
GBM	0.734	0.815	101	23	79	218	0.758	0.561	0.815	0.664
NNET	1.00	0.24	3	121	0	297	0.713	1.000	0.024	0.047
Validation group (standard/low = 80/24)										
RF	1.00	1.00	24	0	0	80	1.00	1.00	1.00	1.00
SVM	0.958	0.912	23	1	7	73	0.923	0.767	0.958	0.852
GLM	0.667	0.662	16	8	27	53	0.663	0.372	0.667	0.478
GBM	0.833	0.737	20	4	21	59	0.760	0.488	0.833	0.615
NNET	0.625	0.612	15	9	31	49	0.615	0.326	0.625	0.429

RF: random forest; SVM: support vector machine; GBM: gradient boosting machine; GLM: generalized linear model; NNET: neural network; Sen: sensitivity; Spe: specificity; TP: true positive; FN: false negative; FP: false positive; TN: true negative.

(Figure 5(d)). To provide a tool for real-time prediction, a dynamic nomogram was used in this study to demonstrate the performance of these six variables (γ-glutamyltransferase level, red blood cell count, alkaline phosphatase level, platelet count, percentage of lymphocytes, and red blood cell distribution width standard deviation) in predicting a low concentration of sodium valproate in blood (Figure S1).

We used 104 patients to further verify the above model. The AUC of RF in the verification group was more accurate than that of the other algorithms (Figure 4(b)). The calibration curve and GiViTi calibration band showed that the model was good in the validation group (Figure 4). The red line in the DCA curve remained above the gray line, and the black line was located from 0.01 to 0.31, indicating that the decision made based on the nomogram model may be beneficial to the prediction of concentrations below the standard level (Figure 5(g)). The clinical influence curve showed that the predictive ability of the nomogram model was also good in the validation group (Figure 5(h)).

4. Discussion

When using the risk prediction models used for children with epilepsy, the plasma concentration of sodium valproate is found to be lower than the standard concentration. After comparing the performance of five machine learning algorithms, we found that the RF prediction model showed the highest AUC. This result indicates that the baseline characteristics of epileptic children can be used to predict that the blood concentration of valproate is lower than the standard concentration. In addition, among the 50 variables included in this study, the nomogram of the 6 factors (γ-glutamyltransferase level, red blood cell count, alkaline phosphatase level, platelet count, percentage of lymphocytes, and red blood cell distribution width standard deviation) based on the RF algorithm was developed to predict that the blood concentration of sodium valproate was lower than the standard concentration. The first five independent variables are associated with the outcome ($p < 0.05$). Because the detection of drug concentrations in blood cannot be completed in every hospital, experimental results are difficult to

obtain. Therefore, the nomogram of these 6 indicators is a simple and practical risk calculator for clinicians. The clinical application value of the top 6 factors was proven by DCA and the GiViTi calibration band. In addition, in the verification group of 104 patients, the RF model results further reflected the above results.

In this study, we found that γ-glutamyltransferase plays an important role in diagnosing whether the blood concentration of sodium valproate in children with epilepsy is within the normal range after taking the drug. γ-Glutamyltransferase, which is an important index, has an important physiological function. It is an enzyme located on the outer surface of the cell membrane in many tissues, mainly in the liver, kidney, and pancreas, and is expressed in all cells except red blood cells [14]. Many studies have shown that long-term use of sodium valproate causes abnormal liver function [14–17]. For example, even when the drug concentration in blood can be maintained within the effective concentration range of 50–100 μg/mL, γ-glutamyltransferase levels still increase [15]. Attilakos et al. reported that the blood concentration of sodium valproate in patients taking the drug alone was maintained at 50–100 mg/L during treatment, and the γ-glutamyltransferase levels in children increased after 6, 12, and 24 months [15]. Interestingly, the average level of γ-glutamyltransferase in children with low concentrations of valproic acid was 14.4 (11.88, 18.72), which was significantly lower ($p = 0.001$) than that in children with normal concentrations of valproic acid 16 (12.9, 22). These findings also show that it is necessary to assess liver function indices when sodium valproate is administered.

Some studies of the other five biochemical indices involved in predicting that the blood concentration of sodium valproate is lower than the standard concentration have indicated that taking sodium valproate leads to changes in biochemical results. Many years ago, it was reported that the red blood cell count in patients decreased after the use of sodium valproate [18, 19]. In this study, it was also found that the red blood cell count in the low concentration group was higher than that in the normal blood concentration group ($p < 0.001$). There has been no direct report on the
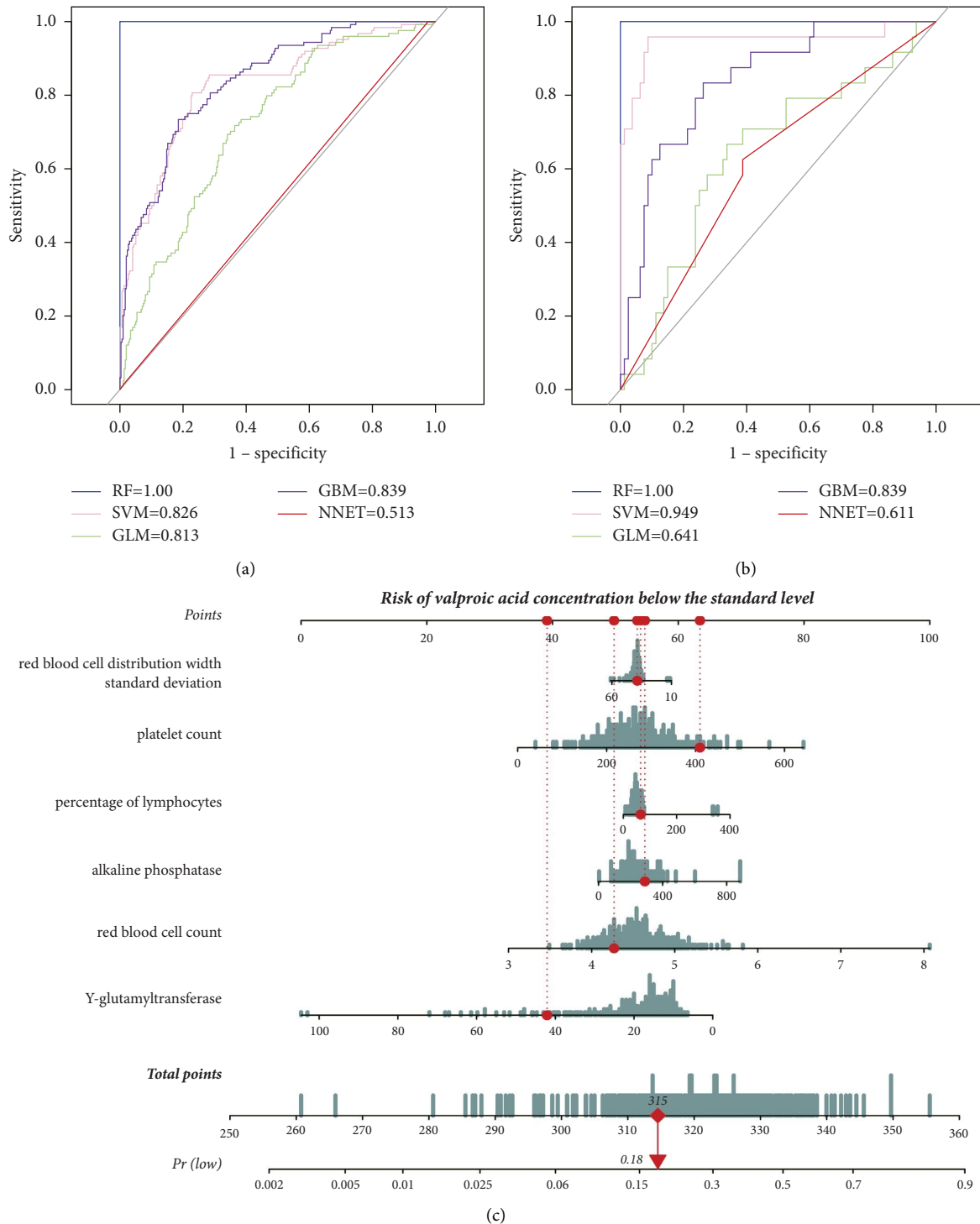
Figure 4: ROC for different machine learning algorithms and establishment of the nomogram. (a) ROC curves indicating the accuracy of the random forest (RF), support vector machine (SVM), gradient boosting machine (GBM), generalized linear model (GLM), and neural network (NNET) in the development group. (b) ROC curves indicating the accuracy of the RF, SVM, GBM, GLM, and NNET in the validation group. (c) Prediction of sodium valproate concentrations below the standard level by nomogram in the development group.

relationship between red blood cell distribution width standard deviation and the blood concentration of sodium valproate. However, the red blood cell distribution width standard deviation is another indicator that reflects the heterogeneity of red blood cells. It is often used to diagnose

anemia in children. To date, studies on anemia caused by sodium valproate include children and the elderly [20, 21]. These phenomena also suggest that we should pay attention to drug-induced anemia resulting from sodium valproate use during clinical examination. Additionally, in 3194
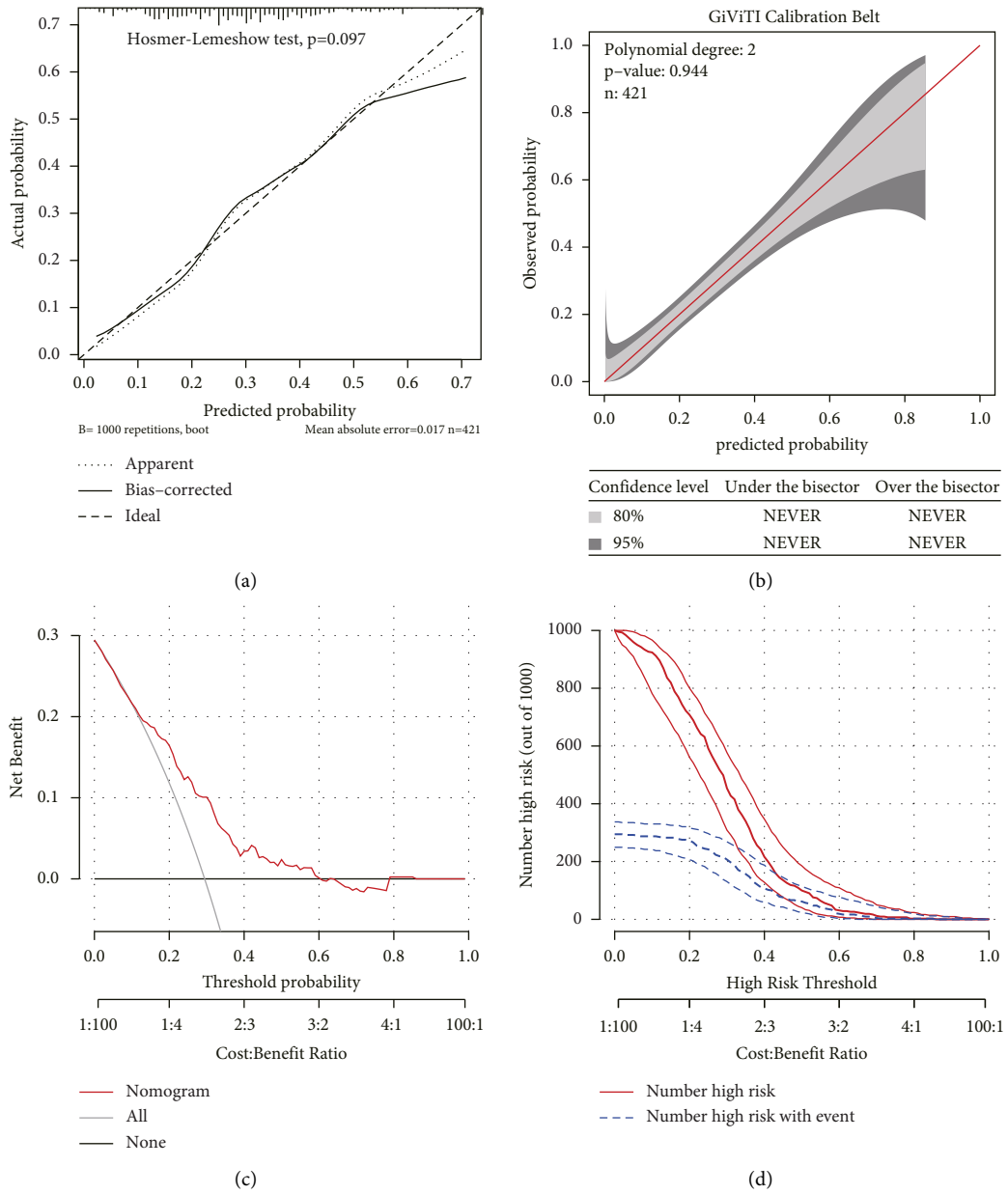
(a)

(b)

(c)

(d)

FIGURE 5: Continued.
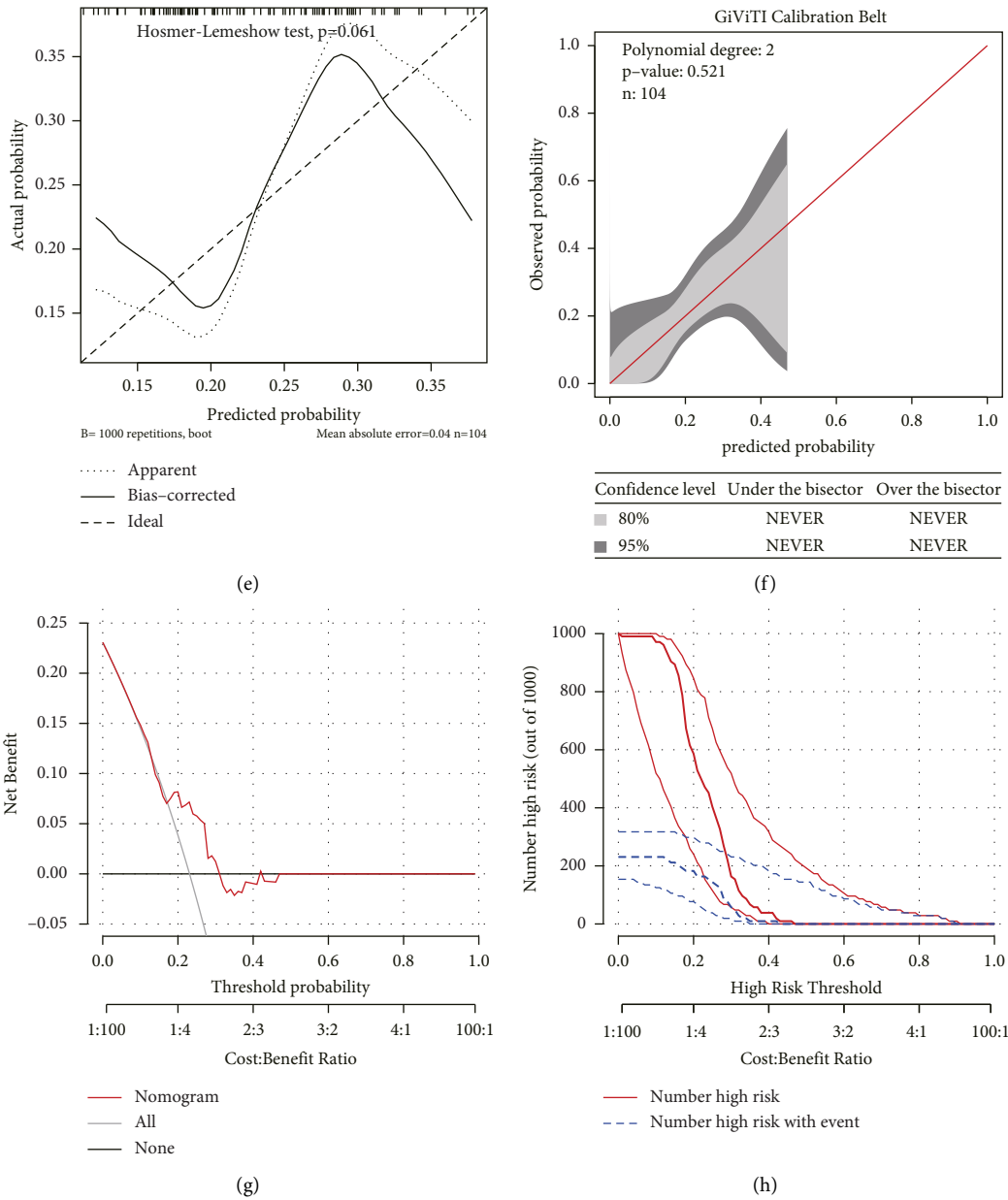
(e)

(f)

(g)

(h)

FIGURE 5: Sodium valproate concentrations below the standard level of nomogram evaluation and clinical use in children with epilepsy. (a–d) Calibration plots (a), GiViTi calibration band (b), decision curve analysis (c), and clinical impact plot (d) for predicting sodium valproate concentrations below the standard level in the development group. (e–h) Calibration plots (e), GiViTi calibration band (f), decision curve analysis (g), and clinical impact plot (h) for predicting sodium valproate concentrations below the standard level in the validation group.

epilepsy outpatients, not only was the blood concentration of valproic acid negatively correlated with platelet count ($r = -0.086$) but also the blood concentration of valproic acid combined with other antiepileptic drugs was negatively correlated with platelet count ($r = -0.079$) [22]. The above results of big data analysis are consistent with the results of this study. This can also reflect the prediction of blood drug concentrations by biochemical indices. A total of 851 record findings that included sodium valproate levels and associated platelet counts in 265 patients were analyzed. There was a significant negative correlation between sodium valproate

levels and platelet counts [23]. It has also been reported that in children treated with valproic acid, valproic acid is associated with decreased platelet count, but platelet production is not affected [24]. There is a causal relationship between the increase in plasma sodium valproate levels and the decrease in platelet counts. This finding is consistent with the result of a high platelet count in patients with low concentrations observed in this study. It has been reported that there is a negative correlation between alkaline phosphatase level and the dosage of valproate ($p < 0.01$) [25]. Although this study did not directly explain the relationship

between alkaline phosphatase level and the blood concentration of sodium valproate, an increase in drug dosage will affect the blood concentration of the drug *in vivo*. This finding also indirectly indicates that the high level of alkaline phosphatase observed in patients in this study may affect the blood concentration of sodium valproate when it is below the standard concentration. To summarize, it was found that the levels of the above three indices (red blood cell count, platelet count, and alkaline phosphatase) were low in the group that exhibited a standard concentration, which was beneficial for predicting that the blood concentration of valproate sodium was lower than the standard concentration in the clinic. However, percentage of lymphocytes in the standard concentration group was higher than that in the low concentration group ($p = 0.022$). To date, there have been few reports about the relationship between lymphocytes and epilepsy [26, 27]. Güneş and Büyükgöl found that the level of lymphocytes in epileptic patients was higher than that in the control group [26]. Recently, an increasing number of studies have reported on the neutrophil to lymphocyte ratio during the diagnosis of epilepsy [27–30]. However, the results of the above studies are inconsistent regarding its diagnostic value. The findings also imply that the level of lymphocytes is related to the dosage of valproate, which is helpful to improve the accuracy of medication.

This study shows the advantages of RF in predicting the blood concentration of sodium valproate when it is below the standard concentration used in epileptic children. However, there are limitations of this study. First, the internal verification method adopted in this study lacks external verification. Second, the sample size of this study was relatively small. Third, this study only considered the comparison between patients with the standard concentration and those with concentrations lower than the standard concentration of sodium valproate in blood and excluded those with high concentrations of sodium valproate ($>100 \, \mu g/mL$). This is because the data processing in this paper is classified into two groups.

5. Conclusions

For children who need long-term detection of sodium valproate in areas where drug concentrations in blood cannot be detected, the relatively easy-to-obtain biochemical index data can be the focus of clinical practice, which helps doctors to judge the adjustment of drug dosage in combination with the curative effect of patients, thus ensuring economic, safe, and effective treatment.

Data Availability

The data used to support the results of this study can be obtained from the corresponding author upon request.

Consent

All participants in this study have given their permission via their guardians to examine them clinically.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Supplementary Materials

Table S1: demographics and clinical characteristics of 525 children with epilepsy. Table S2: baseline characteristics of the study population. Figure S1: dynamic nomogram prediction of sodium valproate concentrations below the standard level in children with epilepsy. (*Supplementary Materials*)

References

[1] S. Neri, G. Mastroianni, E. Gardella, U. Aguglia, and G. Rubboli, "Epilepsy in neurodegenerative diseases," *Epileptic Disorders*, vol. 24, no. 2, pp. 249–273, 2022.

[2] L. L. Thio, L. Edgar, I. Ali et al., "Epilepsy milestones 2.0- an updated framework for assessing epilepsy fellowships and fellows," *Epilepsia*, vol. 63, 2022.

[3] D. A. Sturla Álvarez, E. Sánchez Marcos, C. de Lucas Collantes, V. Cantarín Extremera, V. Soto Insuga, and C. Aparicio López, "Fanconi syndrome secondary to sodium valproate therapy: experience and literature review," *Pediatric Neurology*, vol. 130, pp. 53–59, 2022.

[4] V. E. Rinaldi, G. Di Cara, E. Mencaroni, and A. Verrotti, "Therapeutic options for childhood absence epilepsy," *Pediatric Reports*, vol. 13, no. 4, pp. 658–667, 2021.

[5] X. Lan, K. Mo, L. Nong, Y. He, and Y. Sun, "Factors influencing sodium valproate serum concentrations in patients with epilepsy based on logistic regression analysis," *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, vol. 27, Article ID e934275, 2021.

[6] Z. Zhang, A. A. Gayle, J. Wang, H. Zhang, and P. Cardinal-Fernández, "Comparing baseline characteristics between groups: an introduction to the CBCgrps package," *Annals of Translational Medicine*, vol. 5, no. 24, p. 484, 2017.

[7] Y. Tian, J. Yang, M. Lan, and T. Zou, "Construction and analysis of a joint diagnosis model of random forest and artificial neural network for heart failure," *Aging*, vol. 12, no. 24, pp. 26221–26235, 2020.

[8] Y. Zou, L. Zhao, J. Zhang et al., "Development and internal validation of machine learning algorithms for end-stage renal disease risk prediction model of people with type 2 diabetes mellitus and diabetic kidney disease," *Renal Failure*, vol. 44, no. 1, pp. 562–570, 2022.

[9] K. M. Marcoulides and T. Raykov, "Evaluation of variance inflation factors in regression models using latent variable modeling methods," *Educational and Psychological Measurement*, vol. 79, no. 5, pp. 874–882, 2019.

[10] D. L. Riddle, G. J. Golladay, W. A. Jiranek, and R. A. Perera, "External validation of a prognostic model for predicting

nonresponse following knee arthroplasty," *The Journal of Arthroplasty*, vol. 32, no. 4, pp. 1153–1158.e1, 2017.

[11] G. Nattino, S. Finazzi, and G. Bertolini, "A new calibration test and a reappraisal of the calibration belt for the assessment of prediction models based on dichotomous outcomes," *Statistics in Medicine*, vol. 33, no. 14, pp. 2390–2407, 2014.

[12] H. Chen, J. Zhao, J. Hu et al., "Identification of diagnostic biomarkers, immune infiltration characteristics, and potential compounds in rheumatoid arthritis," *BioMed Research International*, vol. 2022, Article ID 1926661, 15 pages, 2022.

[13] A. J. Vickers, A. M. Cronin, E. B. Elkin, and M. Gonen, "Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers," *BMC Medical Informatics and Decision Making*, vol. 8, no. 1, p. 53, 2008.

[14] F. Hadzagic-Catibusic, E. Hasanbegovic, M. Melunovic, S. Zubcevic, and S. Uzicanin, "Effects of carbamazepine and valproate on serum aspartate aminotransferase, alanine aminotransferase and gamma- glutamyltransferase in children," *Medical Archives*, vol. 71, no. 4, pp. 239–242, 2017.

[15] A. Attilakos, K. A. Voudris, A. Garoufi et al., "Effect of sodium valproate monotherapy on serum uric acid concentrations in ambulatory epileptic children: a prospective long-term study," *European Journal of Paediatric Neurology*, vol. 10, no. 5-6, pp. 237–240, 2006.

[16] N. Ogusu, J. Saruwatari, H. Nakashima et al., "Impact of the superoxide dismutase 2 Val16Ala polymorphism on the relationship between valproic acid exposure and elevation of γ-glutamyltransferase in patients with epilepsy: a population pharmacokinetic-pharmacodynamic analysis," *PLoS One*, vol. 9, no. 11, Article ID e111066, 2014.

[17] G. Ndrepepa and A. Kastrati, "Gamma-glutamyl transferase and cardiovascular disease," *Annals of Translational Medicine*, vol. 4, no. 24, p. 481, 2016.

[18] C. Ozkara, F. E. Dreifuss, and C. A. Hansen, "Changes in red blood cells with valproate therapy," *Acta Neurologica Scandinavica*, vol. 88, no. 3, pp. 210–212, 2009.

[19] E. Hauser, R. Seidl, M. Freilinger, C. Male, and K. Herkner, "Hematologic manifestations and impaired liver synthetic function during valproate monotherapy," *Brain & Development*, vol. 18, no. 2, pp. 105–109, 1996.

[20] B. Kaczorowska-Hac, A. Matheisel, L. Maciejka-Kapuscinska et al., "Anemia secondary to valproic acid therapy in a 13-year-old boy: a case report," *Journal of Medical Case Reports*, vol. 6, no. 1, p. 239, 2012.

[21] C. Li, L. Su, M. Lao, S. Zhu, and M. Ding, "Anemia secondary to the use of sodium valproate for preventing postoperative seizures in a 79-year-old man: a case report," *Medicine*, vol. 97, no. 50, Article ID e13626, 2018.

[22] X. Yang, P. Wang, J. Pan, C. Wang, J. Luo, and H. Ma, "Valproic acid concentration and biochemical indexes in epilepsy outpatients during monotherapy or combination therapy," *Zhong nan da xue xue bao Yi xue ban= Journal of Central South University Medical sciences*, vol. 45, no. 7, pp. 782–789, 2020.

[23] W. Nasreddine and A. Beydoun, "Valproate-induced thrombocytopenia: a prospective monotherapy study," *Epilepsia*, vol. 49, no. 3, pp. 438–445, 2008.

[24] H. Kurahashi, A. Takami, K. Murotani, S. Numoto, and A. Okumura, "Decreased platelet count in children with epilepsy treated with valproate and its relationship to the immature platelet fraction," *International Journal of Hematology*, vol. 107, no. 1, pp. 105–111, 2018.

[25] C. Rawat, S. Kukal, S. Kushwaha et al., "Elevated serum alkaline phosphatase in epilepsy: effect of age and treatment," *British Journal of Biomedical Science*, vol. 77, no. 1, pp. 44–47, 2020.

[26] M. Güneş and H. Büyükgöl, "Relationship between generalized epileptic seizure and neutrophil/lymphocyte ratio, platelet/lymphocyte ratio, and neutrophil mediated inflammation," *International Journal of Neuroscience*, vol. 130, no. 11, pp. 1095–1100, 2020.

[27] G. Morkavuk, G. Koc, and A. Leventoglu, "Is the differential diagnosis of epilepsy and psychogenic nonepileptic seizures possible by assessing the neutrophil/lymphocyte ratio?" *Epilepsy and Behavior*, vol. 116, Article ID 107736, 2021.

[28] S. Hosseini, A. M. E. Mofrad, P. Mokarian et al., "Neutrophil to lymphocyte ratio in epilepsy: a systematic review," *Mediators of Inflammation*, vol. 2022, Article ID 4973996, 7 pages, 2022.

[29] A. Faruk Ozdemir, R. Kemerdere, B. Orhan et al., "Serum endocan and preoperative systemic inflammatory markers in patients with epilepsy," *Neurochirurgie*, vol. 66, no. 1, pp. 29–35, 2020.

[30] Ü. Mete Ural, İ. Şehitoğlu, Y. Bayoğlu Tekin, and F. Kir Şahin, "Neutrophil-to-lymphocyte and platelet-to-lymphocyte ratios in patients with endometrial hyperplasia and endometrial cancer," *Journal of Obstetrics and Gynaecology Research*, vol. 41, no. 3, pp. 445–448, 2015.