Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Explainable machine-learning model for Real-Time Prediction of Intradialytic Hypotension (IDH)

ling tan
Beijing University of Posts and Telecommunications
zemin Kuang
Beijing Anzhen Hospital
haihong E (ehaihong@bupt.edu.cn)
Beijing University of Posts and Telecommunications
meina Song
Beijing University of Posts and Telecommunications
gengxian Zhou
Beijing University of Posts and Telecommunications
haoran Luo
Beijing University of Posts and Telecommunications
ruizhi Zhang
Beijing University of Posts and Telecommunications
zichen Tang
Beijing University of Posts and Telecommunications

Research Article

Keywords: explainable, machine learning, Real-Time, Intradialytic Hypotension

Posted Date: May 17th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2879979/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Background: IDH is a common and serious complication of chronic hemodialysis that can lead to adverse long-term outcomes, including increased cardiovascular and all-cause mortality. Predicting IDH in advance may aid in establishing an intervention plan.

Method: This study used 1173783 hemodialysis sessions from 837 patients and 211 features (including time-span varying and basic information features) as datasets. We constructed 21 features (including ion changes before and after dialysis) by clinical physicians from the raw datasets. We selected 89 features by grid search recursive feature elimination (RFE) based on extreme gradient boosting (XGBoost) (Gs-RFE-XGBoost) and XGBoost model to predict IDH. We used the Shapley value to achieve an explainable result for the 20 features. Besides, we used the quartile 2.5 to quartile 97.5 to get the reference range of the important continuous features.

Result: We found the most important 20 features that are blood pressure related (including the datasets of 4 features and 4 construct features), platelet count, pathway type, pre-dialysis dry weight, dialyzer type, hematocrit, basophils, platelet distribution width, pre-dialysis uric acid, KT/V, parathyroid hormone (PTH), changes in calcium before and after dialysis, dialysis age. We get 0.8446 of avg recall, 0.9149 of avg F1, 0.9410 of avg PRAUC, 0.8823 of avg precision, and 0.8944 of avg MCC to predict the IDH in real-time.

Conclusion: The Gs-RFE-XGBoost-Shapley methods can be applied as a practical, explained, and accurate method to predict IDH. The feature's reference range can provide hints for the clinical physician to prevent the IDH of maintenance hemodialysis (MHD).

1 Background

MHD is the primary treatment for individuals with end-stage renal disease (ESRD), and currently, there are approximately 600,000 people receiving this treatment. Patients with ESRD undergoing MHD face unique hemodynamic challenges, typically on a thrice-weekly basis. IDH is a common and serious MHD complication associated with adverse long-term outcomes, including increased cardiovascular and all-cause mortality [1]. The prevalence of IDH ranges from 8–40% in different studies [2,3,4,5]. However, defining IDH is challenging as there is no generally accepted "safe" blood pressure (BP) range for dialysis patients, resulting in the use of various definitions. The absence of a consensus medical definition limits our understanding of IDH. Besides, predicting IDH in advance may help establish an intervention plan. Cardiovascular and uremic mechanisms underlie IDH [7,8], but other mechanisms may also be associated with it. The research studied different features of ESRD to anticipate IDH. Yang X, Zhao D, Yu F, et al. [9] utilizing CKD-MBD biomarkers, blood pressure, demographic characteristics, and parameters during ultrafiltration. Nafisi et al. [10] identified IDH episodes during HD by investigating highlights of photoplethysmography (PPG) signals. Another examination from a comparable research group developed another PPG-based marker and got high precision [11]. Lin et al. [12] investigated demographic characteristics, blood pressure, and parameters during ultrafiltration. The prediction models had a decent

performance, a sensitivity of 86%, and a specificity of 81%. Lee H, Yun D, Yoo J, et al. [13] categorized the features into six sets: A (age and sex), B (hemodialysis-related features), C (vital signs), D (comorbidities), E (laboratory findings), and F (medications). The recurrent neural network model for predicting intradialytic hypotension achieved an area under the receiver operating characteristic curve of 0.94 (95% confidence intervals, 0.94 to 0.94). CCJLAB, ACYC, APCW, et al. [14] collected BP, pulse rate, age, sex, diabetes mellitus (DM), body temperature, and dry weight. The results are a sensitivity of 86% and a specificity of 81%. Thus, the first challenge of predicting the IDH is that casual diversity makes predicting IDH difficult, although many resources and features on IDH are available.

As datasets continue to increase, selecting the optimal feature subset from the original dataset becomes increasingly important to obtain the best performance in machine learning tasks. Highly dimensional datasets that contain an excessive number of features can lead to low performance in such tasks. Hyperparameter optimization is the process of determining the correct combination of hyperparameters to maximize the model's performance. It runs multiple experiments in one training process. Each test is the complete execution of the training program, and the hyperparameter setting value is selected within the specified range. Once this process is completed, it will be given a set of hyperparameter values that are most suitable for the model to obtain the best results. Shapley's values are from game theory literature and provide a theoretically justified method for the allocation of a coalition's output among the members of the coalition [15]. The Shapley value is known to be the unique method that satisfies certain properties [16]. Feature importance based on XGBoost can be measured by the number of features split or the gain after using the feature split. XGBoost-Bayesian optimizer RFE(BO-RFE) is used Python package Bayes Search CV to find the best parameters based on the RFE methods. The scoring is accuracy, and the random_state is 42. Thus, the second challenge of predicting the IDH is finding the best match between the model and hyperparameter search methods.

In the medical area, explainability is important. However, many studies are only getting the features related to the disease. But the reference range is the doctors' real demand in clinical. However, there haven't been papers that studied the reference range. Thus, the third challenge in predicting the IDH is the strong explanatory power related to clinical practice.

We present an IDH prediction model based on time-span varying of medical examinations(monthly), dialysis records (by day), dialysis machine monitoring information(hourly), and patient basic information, in total 211 features, and construct the 21 features about the ion changes before and after dialysis by clinical physicians, which have not been explored previously. We proposed a model Gs-RFE-XGBoost-shapely that performs the best. Finally, our prediction model enables real-time IDH prediction and explanation of the risk factors by Shapley value associated with IDH during hemodialysis. And we got a 95% confidence medical reference range for the features. Due to the skewed IDH data, we selected the quantile 2.5 to quantile 97.5 to denote the 95% threshold of normal people for the first time.

2 Results

At first, we get the result with construct features and without. Besides, to compare the Gs-RFE-XGBoost' accuracy, we compared the four features selection. Meanwhile, we obtain Shapley's value for interpretability and the figure. At last, we get the 20 features and the medical reference range.

2.1 construct features

We construct 21 features. In order to test the efficiency of features, we compared all features, Gs-RFE-XGBoost_89, and Gs-RFE-XGBoost_108. Table 1 shows that the construct features greatly improve the accuracy of predicting IDH models. At last, we used the Gs-RFE-XGBoost method to obtain the 20 important features.

Table 1
The result of the constructed feature

	avg recall	avg F1	avg PRAUC	avg precision	avg MCC
All_features	0.4631	0.6263	0.7853	0.5862	0.6117
Gs-RFE-XGBoost – 89	0.4939	0.6570	0.8017	0.6129	0.6397
Gs-RFE-XGBoost – 108	**0.8446**	**0.9149**	**0.9410**	**0.8823**	**0.8944**
Gs-RFE-XGBoost _20	**0.8049**	**0.8911**	**0.9268**	**0.8538**	**0.8683**

2.2 different features selection methods

We compared the performance of four feature selection methods in predicting IDH. Table 2 shows that the Gs-RFE-XGBoost had the best performance.

Table 2
The result of different feature selection methods

	avg recall	avg F1	avg PRAUC	avg precision	avg MCC
All-features	0.4631	0.6263	0.7853	0.5862	0.6117
Importance-89	0.4476	0.6136	0.7815	0.5766	0.6021
Shap-89	0.4822	0.6455	0.7962	0.6036	0.6297
Gs-RFE-XGBoost-89	**0.4939**	**0.6570**	**0.8017**	**0.6129**	**0.6397**
BO-RFE-89	0.4738	0.6382	0.7929	0.5971	0.6234

2.3 utilizing shape values for interpretability

We use the Gs-RFE-XGBoost method to select 20 features. And we utilize shapley values for interpretability. We identified 20 important features for predicting IDH using the Shapley value method. The mean Shapley value is shown in Table 3. It is possible to calculate the Shapley values for each feature during each observation, and from a global perspective, the Shapley values can be used for global interpretation. In Fig. 1, all values on the left indicate the cause of the predicted value moving in the negative direction, while the values on the right indicate the cause of the predicted value moving in the positive direction. These features are reflected on the left y-axis. The baseline value of the model y_base = 0.2535 is the mean fitting value of the target variable in the training set.

Table 3
The mean Shapley value of features

Features	Mean Shapley value
Pre-systolic blood pressure	-0.00021
Pre-diastolic blood pressure	0.000121
systolic pressure	-6.01E-05
diastolic pressure	-9.83E-05
platelet count	-5.40E-06
Channel type	6.24E-06
Net weight before penetration	-2.10E-05
Dialysis machine model	-3.30E-06
Hematocrit	7.01E-07
Basophil	-1.30E-06
Platelet distribution width	-6.75E-08
Uric acid before dialysis	-1.17E-06
KT/V	-1.36E-06
parathyroid hormone	-1.25E-06
Changes in systolic blood pressure before dialysis	-0.03853
Changes in diastolic blood pressure before dialysis	0.038826
Changes in diastolic blood pressure before and after dialysis	-2.22E-06
Changes in systolic blood pressure before and after dialysis	-1.48E-05
Changes in calcium before and after dialysis	7.19E-07
Dialysis age	-3.77E-06

2.4 Gs-RFE-XGBoost method selected 20 features and the reference range

We use non-IDH patients' P2.5 – P97.5 to represent 95% of normal people. The specific result is shown in Table 4.

Table 4
The reference range of the most important 20 features

Feature	Reference range
pre_SBP	[84.45, 215.55]
Pre_DBP	[43.2, 126.8]
SBP	[84.45, 215.55]
DBP	[43.2, 126.8]
plc	[23.425, 609.575]
pre_nw	[37.4975, 124.8025]
Hem	[0.15295, 0.56905]
bas	[0.013625, 0.341375]
pla_width	[9.2, 20.6]
pre_dia	[86.45, 863.55]
KT/V	[0.05575, 1.94625]
par_hor	[44.37825, 1667.29175]
Cha_SBP before dia	[-132.65, 132.65]
Cha_DBP before dia	[-84.675, 79.675]
Cha_DBP bef_aft dia	[-84.55, 84.55]
Cha_SBP bef_aft dia	[-132.05, 132.05]
Cha_cal bef_aft dia	[-1.13075, 0.74075]
dia_age	[192.475, 7468.525]

3 Discussion

In the paper, we got 4 results. They are the construct features, four feature selection methods, 20 important features related to the IDH, and the medical reference range of the important feature. They prove that the construct features may be the mechanism related to the IDH. We proposed that the Gs-RFE-

XGBoost model has the best performance. And the 20 important features give the conduct to the clinical physician for the treatment and prevention of IDH. Last but not least, the medical reference range of the important feature may provide new clues or the factor of the disease's research. The following will provide detailed explanations one by one:

3.1 Construct features

We construct 21 features. They are details as follows.

Table 5
The construct of the 21 features

number	features
1	changes in uric acid before and after dialysis
2	changes in urea nitrogen before and after dialysis
3	changes in sodium before and after dialysis
4	changes in magnesium before and after dialysis
5	changes in chlorine before and after dialysis
6	changes in phosphorus before and after dialysis
7	changes in potassium before and after dialysis
8	changes in creatinine before and after dialysis
9	changes in calcium before and after dialysis
10	changes in blood β2 changes in microglobulin before and after dialysis
11	changes in carbon dioxide before and after dialysis
12	changes in heart rate before and after dialysis
13	changes in diastolic blood pressure before and after dialysis
14	changes in systolic blood pressure before and after dialysis
15	changes in systolic blood pressure before and during dialysis
16	changes in systolic blood pressure before and during dialysis
17	changes in systolic blood pressure during and after dialysis
18	changes in systolic blood pressure during and after dialysis
19	weight gain during dialysis
20	dialysis age
21	average arterial pressure

We used the Gs-RFE-XGBoost method to predict the IDH with or without features. we get the 0.4939 of avg recall, the 0.6570 of avg F1, the 0.8017 of avg PRAUC, the 0.6129 of avg precision, and the 0.6397 of avg MCC to predict the IDH in real-time without the constructed feature. And we get the 0.8446 of avg recall, the 0.9149 of avg F1, the 0.9410 of avg PRAUC, the 0.8823 of avg precision, and the 0.8944 of avg MCC to predict the IDH in real-time with the constructed feature. It improved accuracy by 40%.

Besides, when the Gs-RFE-XGBoost picked the most important 20 features, 6 features are from the construct features. It indicated the change of the features before and after dialysis is probably more important than the independent features. It hints the clinical physician to focus more on the IDH patient's changes in features may obtain the new finding.

3.2 Different feature selection methods

Shapley value method was put forward by Shapley L.S. in 1953. It belongs to the field of cooperative games in order to solve the problem of conflicts caused by the distribution of interests among multiple players in the process of cooperation. One advantage of applying the Shapley value is to distribute benefits according to the marginal contribution rate of members to the alliance, that is, the benefits shared by member I am equal to the average value of the marginal benefits created by the member for the alliance he participates in. Only the shapley value and permutation method are consistent, and only the shapley value is personalized, so the shapley value is the only consistent personalized feature attribution method.

The Gs-RFE-XGBoost consisting of the grid algorithm of hyperparameters search method, RFE of feature selected method, and XGBoost predicted model. The grid algorithm is one of the representative algorithms of heuristic algorithms, whose core idea is to select the local optimal solution every time, but this algorithm cannot guarantee that the final result is the global optimal solution. The order to be eliminated in this process is the sorting of features. RFE method is to repeatedly build the model, then select the best (or worst) feature (based on the coefficient), put the selected feature aside, and then repeat the process on the remaining features until all the features have been traversed. Gradient Boosting Decision Tree (GBDT) is an additive model based on boosting ensemble ideas. XGBoost is the ultimate GBDT. XGBoost uses second-order Taylor formula expansion on the basis of GBDT to optimize the loss function and improve the calculation accuracy. The regularization term is used to simplify the model and avoid overfitting. Finally, the Blocks storage structure is used for parallel computing. XGBoost calculates which feature to select as the segmentation point according to the gain of the structure score, and the importance of a feature is the sum of its occurrence times in all trees. That is to say, the more an attribute is used to build a decision tree in the model, the more important it is.

The BO-RFE-XGBoost consisting of the bayesian optimizer of hyperparameters search method, RFE of feature selected method, and XGBoost predicted model. The Bayesian optimizer is a Bayesian optimization based on Gaussian processes, with many continuous parameters in the algorithm's parameter space and a relatively short running time. The input of the objective function of the Bayesian

optimizer must be a specific hyperparameter, not the entire hyperparameter space, nor the elements other than the hyperparameter such as data and algorithm. Bayesian Optimization is a black box optimizer used to find the optimal parameters.

The feature importance based XGBoost calculates which feature to choose as the segmentation point based on the gain of the structural score, and the importance of a feature is the sum of its occurrences in all trees. That is to say, the more an attribute is used to construct a decision tree in the model, the higher its importance is relatively.

We used the 4 feature-selected methods to select the features. The Gs-RFE-XGBoost selected 20 features for IDH is the best. And we get the 0.4939 of avg recall, the 0.6570 of avg F1, the 0.8017 of avg PRAUC, the 0.6129 of avg precision, and the 0.6397 of avg MCC to predict the IDH in real-time.

3.3 The important features related to the IDH

Episodes of IDH cause end-organ damage due to unstable hemodynamics. Predicting episodes of IDH pre-dialysis is the focal point of treatment because advanced clinical interventions can forestall some adverse events [9].

Blood platelets, abbreviated as PLTs, are a type of cell in our blood. They are called "platelets" because they are slightly concave, elliptical, or disc-shaped on both sides. The number of platelets in normal human blood will be maintained at a certain level. Some diseases can cause a decrease or increase in platelet count. Platelet count is helpful in the diagnosis and differential diagnosis of hemostasis and thrombotic diseases in clinical practice.

Platelet count refers to the number of platelets contained in a unit volume of blood. The number of platelets in normal human blood will be maintained at a certain level. Some diseases can cause a decrease or increase in platelet count. Platelet count is helpful in the diagnosis and differential diagnosis of hemostasis and thrombotic diseases in clinical practice.

Hematocrit refers to the volume ratio of sinking red blood cells to whole blood measured after a certain amount of anticoagulant whole blood is centrifuged and precipitated, which is a simple method to indirectly reflect the size and volume of red blood cells. By combining red blood cell count and hemoglobin content, the average value of red blood cells can be calculated, which is helpful for the morphological classification of anemia.

Basophils contain histamine, heparin, and allergenic slow-reactive substances. Heparin has anticoagulant effects, while histamine can alter the permeability of capillaries. Allergic slow-response substances are lipid molecules that can cause smooth muscle contraction. The occurrence of allergic reactions in the body is related to these substances. When basophils are in the connective tissue and mucous epithelium, they are called a mast cells. Their structure and function are similar to basophils.

Platelet distribution width is seen in thrombocytopenic disorders (such as aplastic anemia) and increased platelet destruction (such as thrombocytopenic purpura); An increase can be seen in acute infections, acute hemolysis, and myeloproliferative diseases.

Blood platelets, Hematocrit, Basophils, and platelet distribution width are from the blood routine. It hints to the doctors that the IDH's pathology may relate to the immune system. Further research work can be carried out at this point.

Pathway types are temporary intubation, graft vascular fistula, long-term intubation, direct puncture, and autologous arteriovenous fistula. An autologous arteriovenous fistula is a true surgical connection between arteries and veins, typically chosen with a non-dominant arm. This type of dialysis can increase the blood flow velocity through the veins, helping to expand and strengthen the veins. An arteriovenous fistula can accelerate blood flow from the vein to the dialysis machine. An arteriovenous fistula is the preferred choice for all types of hemodialysis access and is commonly called the "gold standard".

The dialyzer types are 4008B, 4008s, 5008S, AK200, ak96, AK98, DBB-26, and DBB-27. The dialysate concentrate and dialysate water are prepared into qualified dialysate through the dialysate supply system. Through a blood dialyzer, solute dispersion, permeation, and ultrafiltration are carried out with the patient's blood drawn out by the blood monitoring alarm system; The purified patient's blood is returned to the patient's body through the blood monitoring alarm system, and the dialyzed liquid is discharged from the dialysate supply system as a waste liquid. This process is constantly cyclical, achieving the purpose of treatment and completing the entire dialysis process. Different types selectively use all or part of the above functions based on their market positioning. Generally speaking, blood pumps, heparin pumps, arteriovenous pressure monitoring, air detectors, and venous clamps are the basic configurations of extracorporeal circulation. Any equipment used for blood purification treatment through extracorporeal circulation must have the above five basic components.

The pathway types and dialyzer types are about the hemodialysis equipment and setting. It is indicated that the hemodialysis machine also can influence the occurrence the IDH. Thus the MHD's daily nursing work is still important.

Excessive production of uric acid results in an increase in blood uric acid levels. Reduced uric acid excretion, causing an increase in blood uric acid levels. Primary low uric acid excretion is often related to genetics. Increased reabsorption of uric acid, causing an increase in blood uric acid levels.

PTH is an alkaline single-chain polypeptide hormone secreted by the main cell of the parathyroid gland. Its main function is to regulate the metabolism of calcium and phosphorus in vertebrates, promote the increase of blood calcium levels, and decrease blood phosphorus levels. Yang X, Zhao D, Yu F, et al found the most critical features that related to the IDH are PTH, ALP, dialysis vintage, and MAP [9]. The recommendation range of PTH is 150-300pg/ml according to K/DOQI guidelines, PTH shows a U-shaped curve relationship with mortality [17] and fracture risk [18].

Pre-dialysis dry weight is the weight obtained by hemodialysis patients after removing their clothing and weight before dialysis. The influence factor of weight is age, dialysis age, and DTI score of maintenance hemodialysis patients.

The KT/V value refers to the ratio of the clearance of urea by the dialyzer to the volume during a certain dialysis time. A multicenter, large-scale collaborative study in the United States suggests that for every 0.1 increase in KT/V, the mortality rate decreases by 7%.

Changes in systolic blood pressure before and during dialysis, Changes in diastolic blood pressure before and during dialysis, Changes in diastolic blood pressure before and after dialysis, Changes in systolic blood pressure before and after dialysis, Changes in calcium before and after dialysis are we construct features. Therefore blood pressure change is essential for the IDH patient. The IDH patient must accept blood pressure monitoring.

Pre-dialytic systolic blood pressure, systolic blood pressure, pre-dialysis diastolic blood pressure, and diastolic blood pressure are closely related to the IDH in our study. F.M. van der Sande and A.W. Mulder pointed out that regardless of systolic or diastolic cardiac dysfunction, they are likely to have decreased cardiac output, thus, contributing to the risk for IDH [19].

Dialysis age refers to the time interval between the patient's current dialysis time and the first dialysis. Dialysis age is a key index to evaluate hemodialysis patients. Yang X, Zhao D, Yu F, et al. found that dialysis vintage is the key indicator to predict the IDH [20].

3.4 The medical reference range of the important feature

It is the first attempt to obtain the important feature's medical reference range. It can also experiment on a large scale for research later.

Table 6
The important feature's medical reference range

Feature	Reference range
pre_SBP	[84.45, 215.55]
Pre_DBP	[43.2, 126.8]
SBP	[84.45, 215.55]
DBP	[43.2, 126.8]
Plc	[23.425, 609.575]
pre_nw	[37.4975, 124.8025]
Hem	[0.15295, 0.56905]
Bas	[0.013625, 0.341375]
pla_width	[9.2, 20.6]
pre_dia	[86.45, 863.55]
KT/V	[0.05575, 1.94625]
par_hor	[44.37825, 1667.29175]
Cha_SBP before dia	[-132.65, 132.65]
Cha_DBP before dia	[-84.675, 79.675]
Cha_DBP bef_aft dia	[-84.55, 84.55]
Cha_SBP bef_aft dia	[-132.05, 132.05]
Cha_cal bef_aft dia	[-1.13075, 0.74075]
dia_age	[192.475, 7468.525]

4 Conclusion

In our study, we find the Gs-RFE-XGBoost methods selected features to predict the IDH have the best effect on medical data. Besides, we get the Shapley value explaining the result and features reference range. The Gs-RFE-XGBoost-Shapley methods can be applied as a practical, explained, and accurate method to predict IDH.

5 Methods

5.1 Research object and dialysis settings

The study retrospectively reviewed 1173783 hemodialysis sessions from 837 patients who received treatment at Beijing Luhe Hospital affiliated with Capital Medical University between June 2012 and July 2022. Hemodialysis treatments were conducted two to three times a week, with each lasting up to 240 minutes. For most patients, the initial dialysis settings included a dialyzate sodium concentration of 138 mmol/L, a dialyzate calcium concentration of 1.5 mmol/L, a dialyzate flow rate of 500 mL/min, and a dialyzate temperature of 36.5 C.

5.2 The step of predicting the IDH

The step of predicting the IDH included data aggregation, converting categorical values to numeric values, labeling IDH standards, constructing features, training, testing, and validating sets, model prediction without construct features, feature selection with the 21 construct features, explainable results, and 95% feature reference range in total 9 steps as shown in Fig. 1. The following will provide a detailed introduction to these 9 steps:

1) Data Aggregation: We collected a total of 57069 medical examinations, 254191 dialysis records, 1246779 dialysis machine monitoring information, and 837 patient basic information. To aggregate all the features for real-time prediction of intradialytic hypotension, we used the dialysis machine monitoring information's timestamp as a target, matching medical examinations with the same month and dialysis records on the same day. The statistical description of all features is in Table 1. We used the pandas' Python package merge function to aggregate all the data, resulting in 1173783 dialysis sessions and 211 features as shown in Fig. 3. The dataset contained information at initial and any time points during hemodialysis. These features mainly include:

a. Medical examination (95 features) – include a routine blood test, liver and kidney function, electrolyte, blood fat, blood sugar, PTH, iron metabolism, TKV, prealbumin, inflammation index, 16β2-MG, natriuretic peptide, and LDH and so on. The medical examination is recorded once a month.

b. Dialysis record (99 features) - dialysis records (record of dialysis by day) and dialysis machine monitoring information (hourly recording of each dialysis).

c. Patient basic information (17 features) - sex, weight, first dialysis time, and other relevant details.

2) Convert categorical values to numeric values

We used a mapping function to convert categorical values to numeric values. For instance, the pathway type category contains temporary intubation, graft vascular fistula, long-term intubation, direct puncture, and autologous arteriovenous fistula, which we converted to 1, 2, 3, 4, and 5, respectively. We converted a total of 29 categorical features to numeric attributes.

3) Label 8 IDH standards

We labeled 8 standards for IDH. These standards include Fall20, Fall30, Nadir90, Nadir100, Fall20Nadir90, Fall30Nadir90, Flythe, J.E. et al., and KDOQI, with their corresponding calculation methods provided in Table 7. Figure 4 shows the frequency of IDH occurrence in our dataset. Due to the low frequency of standards, we focused on standard 8 KDOQI.

Table 7
The calculation methods of standards of intradialytic hypotension

standards	calculation methods
Fall20[3]	Set Pre-SBP-SBP > = 20 as 1, others as 0
Fall30[3]	Set Pre-SBP-SBP > = 30 as 1, others as 0
Nadir90[3]	Set SBP < 90 as 1, others as 0
Nadir100[3]	Set SBP < 100 as 1, others as 0
Fall20Nadir90[3]	Set Pre-SBP > = 20 and SBP < 90 as 1, others as 0
Fall30Nadir90[4]	Set Pre-SBP > = 30 and SBP < 90 as 1, others as 0
Flythe, J.E.et al[5]	Set 120 < = Pre-SBP < = 159 and SBP < 90; PreSBP > = 160 and SBP < 100 as 1, others as 0
KDOQI [1]	Set average mean arterial pressure > = 10 or Pre-SBP-SBP > = 20 as 1, others as 0

4) Construct features

As the dataset does not include the dialysis age feature, we calculate it by subtracting the patient's first dialysis time from the current dialysis date. The weight gain during dialysis is obtained by subtracting the weight after the previous dialysis from the weight before the current dialysis. The average arterial pressure is computed using the following formula: (pre_SBP + 2 * pre_DBP) / 3 - (SBP + 2 * DBP) / 3. Besides, we construct 18 features. They are ion changes between before dialysis and after dialysis. For example, changes in uric acid before and after dialysis and changes in urea nitrogen before and after dialysis.

5) Train, test, and validate sets

To begin, the data was divided into 10 sets for subsequent analysis. Eight sets were used for training, while the remaining two sets were used for testing. Next, the K-Fold function in the scikit-learn package was utilized to perform 5-fold cross-validation by dividing the training dataset into five sets of train and validate sets. Given the imbalance between IDH and non-IDH patients, the SMOTE algorithm was employed for cross-validation sampling.

6) model prediction without construct features

In order to evaluate the importance of construct features, we compare with and without features that predict the IDH. The prediction model that we used is XGBoost. And the number of features is 211.

7) Feature selected 89 features with 21 construct features

We used four approaches to compare different feature selection methods: Gs-RFE-XGBoost, BO-RFE, Shapley value, and feature importance based on XGBoost.

Gs-RFE-XGBoost algorithm flow:

1. The grid search to find the best parameters. The best parameters are that the max-depth = 10, min-child-weight = 4, n-jobs=-1, random-state = 0, reg-alpha = 0.01, reg-lambda = 0.01, scale-pos-weight = 1, seed = None, subsample = 0.8, verbosity = 1, num-class = 2.
2. XGBoost is trained from the training set *T* containing all samples, and the five-fold cross-validation method is used to evaluate the prediction accuracy of XGBoost based on the new feature subset after each round of feature elimination based on the mean absolute error (MAE);
3. Calculate and sort the importance IM(∝) of each feature ∝ in the feature set based on gain;
4. According to the backward selection of the sequence, delete the feature with the lowest feature importance, and repeat the remaining feature subset for 1–2 until the feature subset is empty. According to the cross-validation results of each feature subset, the feature subset with the highest prediction accuracy is determined.

And BO_RFE algorithm with the use of the RFE-CV function. We make use of Bayes Search CV to find the best parameters. We make the learning rate's threshold 0.01–0.3, the max-depth 3–10, the n-estimators integer 50–500, the gamma 0–1, the subsample 1.5-1, the colsample-bytree − 1, the min-child-weight integer 1–10, the reg-alpha 0–1, and the reg-lambda 0–1. And Bayes Search CV finds the optimal hyperparameter among these parameters. At last, we get the parameters that the colsample-bytree is 1.0, the gamma is 0.71, the learning rate is 0.3, the max depth is 4, the min child weight is 10, the n estimators is 431, the reg-alpha is 0.93, the reg-lambda is 0.93, and subsample is 0.72.

8) Gs-RFE-XGBoost Feature selected 20 features

We used the best performing method Gs-RFE-XGBoost to get the most important 20 features and the Shapley method to get the explainable results.

9) Explainable results of the Shapley value

We used the Shapley value to get the explainable results.

10) 95% feature reference range

We make the most important continuous feature split into 100 equal parts, and P2.5(2.5*1 equal parts + the minimum of the continuous feature) to P97.5(97.5*1 equal parts + the minimum of the continuous feature) is the reference range.

6 Evaluation criteria

In the feature selection experiments, several evaluation metrics are used to assess the performance of the models, including Recall, F1-measure, precision, and Matthews Correlation Coefficient (MCC). The equations for these metrics are shown in Eqs. (1)-(4).

Recall = TP/ (TP + FN) (1)

F1 = (2 × Precision × Recall) / (Precision + Recall) (2)

Precision = (TP) / (TP + FP) (3)

MCC = $\frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP) \times (TP+FN) \times (TN+FP) \times (TN+FN)}}$ (4)

In these equations, TP represents the number of true positive instances, TN represents the number of true negative instances, FN represents the number of false negative instances, and FP represents the number of false positive instances. Recall measures the proportion of true positive cases that are correctly identified by the classifier. A higher recall score indicates better identification of true positive cases. F1-measure provides a balanced assessment of the classifier's precision and recall. Higher F1-measure values indicate better classification results. Precision measures the proportion of positive samples that are correctly identified by the classifier. A higher precision score indicates more accurate positive case prediction. MCC is a measure of the correlation between the predicted and true values, with higher values indicating better model reliability and prediction accuracy.

7 List of abbreviations

Feature	abbreviations
Intradialytic Hypotension	IDH
recursive feature elimination	RFE
extreme gradient boosting	XGBoost
grid search recursive feature elimination (RFE) based on extreme gradient boosting (XGBoost)	Gs-RFE-XGBoost
parathyroid hormone	PTH
maintenance hemodialysis	MHD
end-stage renal disease	ESRD
blood pressure	BP
The National Kidney Foundation's Kidney Disease Outcomes Quality Initiative	KDOQI
photoplethysmography	PPG
body mass index	BMI
interdialytic weight gain	IDWG
XGBoost-Bayesian optimizer RFE	BO-RFE

Declarations

Ethical approval and consent to participate

The study was conducted in accordance to relevant guidelines and regulations. Informed consent was obtained from all the participants and/or their LAR. The study was approved by [echics committee of Beijing anzhen hospital capital medical university 2018016]

Consent for publication

Not applicable

Availability of data and materials

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

Not applicable

Author's contributions

E haihong, Kuang zemin and Song meina designed research. Zhou gengxian, Tan ling collected and analyzed data. E haihong, Tan ling and Kuang zemin designed method. Tan ling performed the experiments. Tan ling wrote the manuscript. E haihong, Tan ling and Song meina reviewed and edited the manuscript. All authors have read and approved the final manuscript.

Acknowledgements

We acknowledge some limitations to this study. Data were collected from a single center, which could raise issues regarding the generalization of the findings.

References

1. Mehmet K , Ertuglu L A , Baris A , et al. An update review of intradialytic hypotension: concept, risk factors, clinical implications and management[J]. Clinical Kidney Journal, 2020(6):6.
2. Yu J, Liu Z, Shen B et al. Intradialytic hypotension as an in dependent risk factor for long-term mortality in maintaining hemodialysis patients: a 5-year follow-up cohort study. Blood Purif 2018; 45: 320–326
3. Okoye OC, Slater HE, Rajora N. Prevalence and risk factors of intra-dialytic hypotension: a 5 year retrospective report from a single Nigerian Centre. Pan Afr Med J 2017; 28: 62
4. Rocha A, Sousa C, Teles P et al. Effect of dialysis day on intradialytic hypotension risk. Kidney Blood Press Res 2016; 41: 168–174
5. Kuipers J, Verboom LM, Ipema KJR et al. The prevalence of intradialytic hypotension in patients on conventional hemodialysis: a systematic review with meta-analysis. Am J Nephrol 2019; 49: 497–506
6. Gul A, Miskulin D, Harford A, et al. Intradialytic hypotension [J]. Current opinion in nephrology and hypertension, 2016, 25(6):545-550.
7. Reeves PB, Mc Causland FR: Mechanisms, clinical implications, and treatment of intradialytic hypotension.Clin J Am Soc Nephrol 13: 1297–1303, 2018
8. K/DOQI Workgroup. K/DOQI clinical practice guidelines for cardiovascular disease in dialysis patients. Am J Kidney Dis 2005; 45(4 Suppl 3): S1–S153
9. Yang X, Zhao D, Yu F, et al. An optimized machine learning framework for predicting intradialytic hypotension using indexes of chronic kidney disease-mineral and bone disorders[J]. Computers in Biology and Medicine, 2022, 145: 105510.
10. V.R. Nafisi, M. Shahabi, Intradialytic hypotension related episodes identification based on the most effective features of photoplethysmography signal, Comput Methods Programs Biomed, 157 (2018) 1-9.

11. F. Sandberg, R. Bailón, D. Hernando, P. Laguna, J.P. Martínez, K. Solem, L. Sörnmo, Prediction of hypotension in hemodialysis patients, Physiol Meas, 35 (2014) 1885-1898.
12. C.J. Lin, C.Y. Chen, P.C. Wu, C.F. Pan, H.M. Shih, M.Y. Huang, L.H. Chou, J.S. Tang, C.J. Wu, Intelligent system to predict intradialytic hypotension in chronic hemodialysis, J Formos Med Assoc, 117 (2018) 888- 893.
13. Lee H, Yun D, Yoo J, et al. Deep learning model for real-time prediction of intradialytic hypotension[J]. Clinical Journal of the American Society of Nephrology, 2021, 16(3): 396-406.
14. CCJLAB, ACYC, APCW, et al. Intelligent system to predict intradialytic hypotension in chronic hemodialysis[J]. Journal of the Formosan Medical Association, 2018, 117(10):888-893.
15. Saria S, Rajani AK, Gould J, Koller D, Penn AA, Integration of Early Physiological Responses Predicts Later Illness Severity in Preterm Infants, Sci. Transl. Med 2, 48ra65–48ra65 (2010).
16. Sundararajan M, Najmi A. The many Shapley values for model explanation[C]//International conference on machine learning. PMLR, 2020: 9269-9278.
17. J.L. Fernández-Martín, P. Martínez-Camblor, M.P. Dionisi, J. Floege, M. Ketteler, G. London, F. Locatelli, J.L. Gorriz, B. Rutkowski, A. Ferreira, W.J. Bos, A. Covic, M. Rodríguez-García, J.E. Sánchez, D. Rodríguez-Puyol, J.B. Cannata-Andia, Improvement of mineral and bone metabolism markers is associated with better survival in haemodialysis patients: the COSMOS study, Nephrol Dial Transplant, 30 (2015) 1542- 1551.
18. S. Iimori, Y. Mori, W. Akita, T. Kuyama, S. Takada, T. Asai, M. Kuwahara, S. Sasaki, Y. Tsukamoto, Diagnostic usefulness of bone mineral density and biochemical markers of bone turnover in predicting fracture in CKD stage 5D patients--a single-center cohort study, Nephrol Dial Transplant, 27 (2012) 345- 351.
19. F.M. van der Sande, A.W. Mulder, S.J. Hoorntje, K.H. Peels, W.H. van Kuijk, J.P. Kooman, K.M. Leunissen, The hemodynamic effect of different ultrafiltration rates in patients with cardiac failure and patients without cardiac failure: comparison between isolated ultrafiltration and ultrafiltration with dialysis, Clin Nephrol, 50 (1998) 301-308.
20. Yang X, Zhao D, Yu F, et al. Boosted machine learning model for predicting intradialytic hypotension using serum biomarkers of nutrition[J]. Computers in Biology and Medicine, 2022, 147: 105752.

Figures

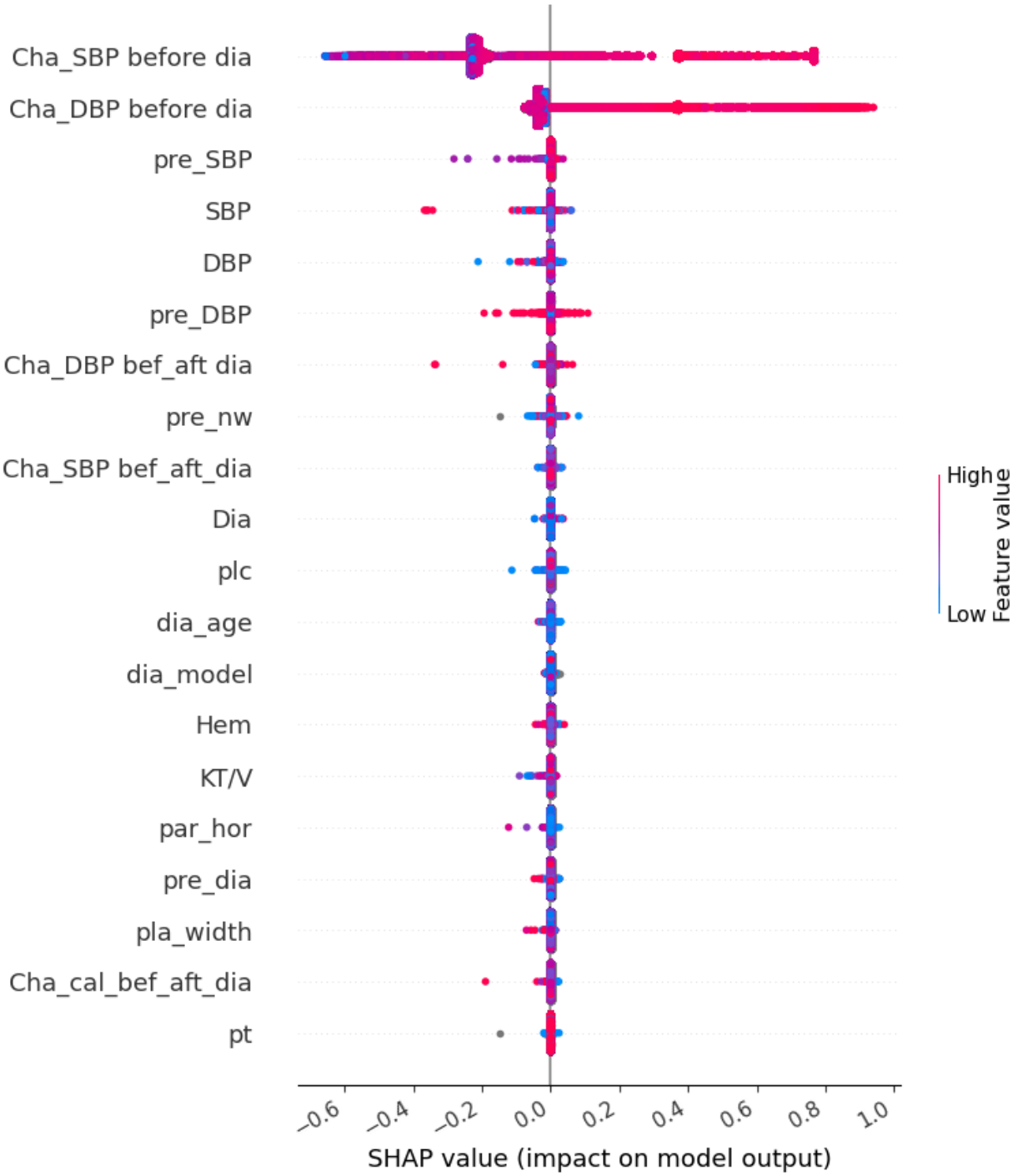


Figure 1

Shapley's value of the important 20 feature

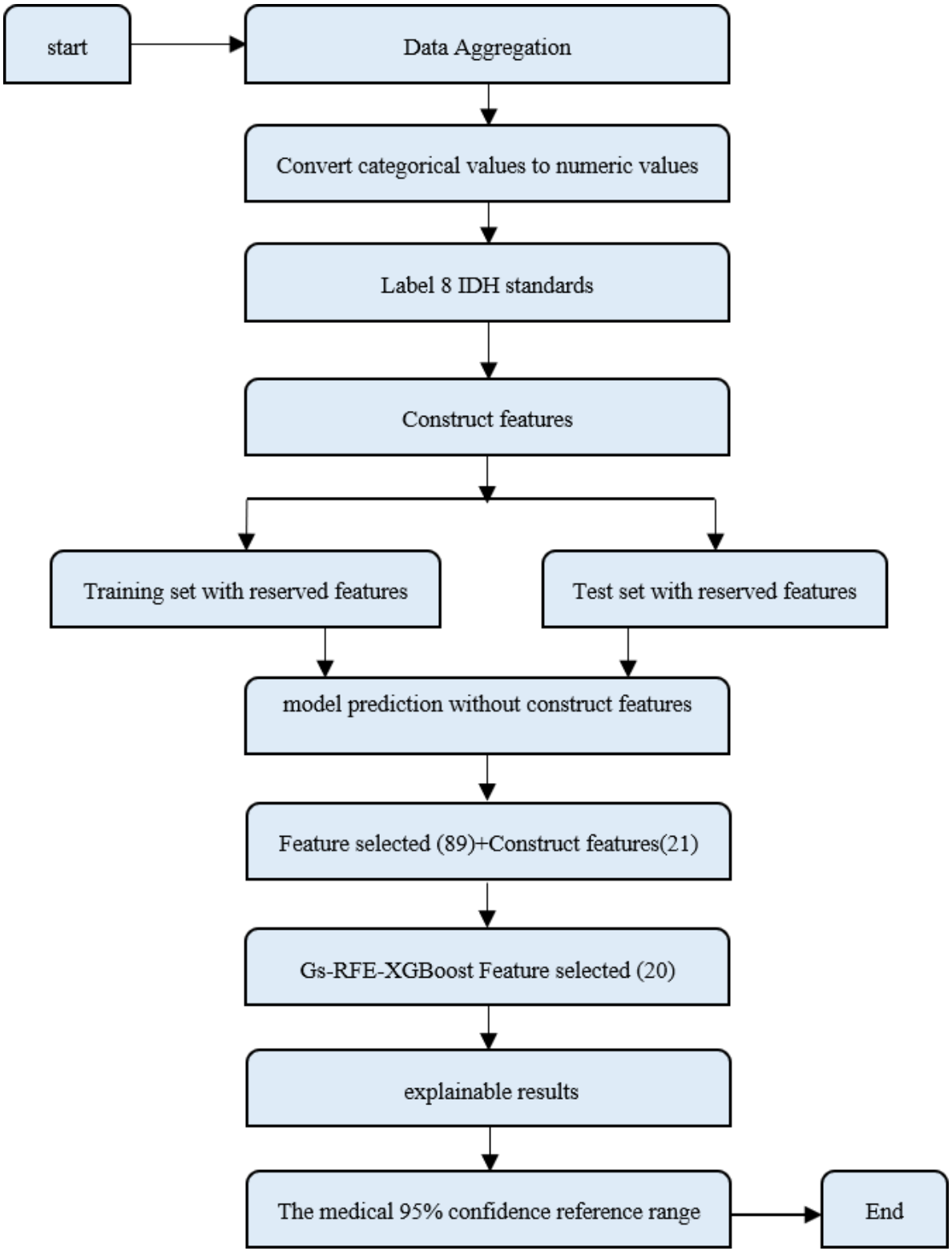


Figure 2

the flowchart of predicting the IDH

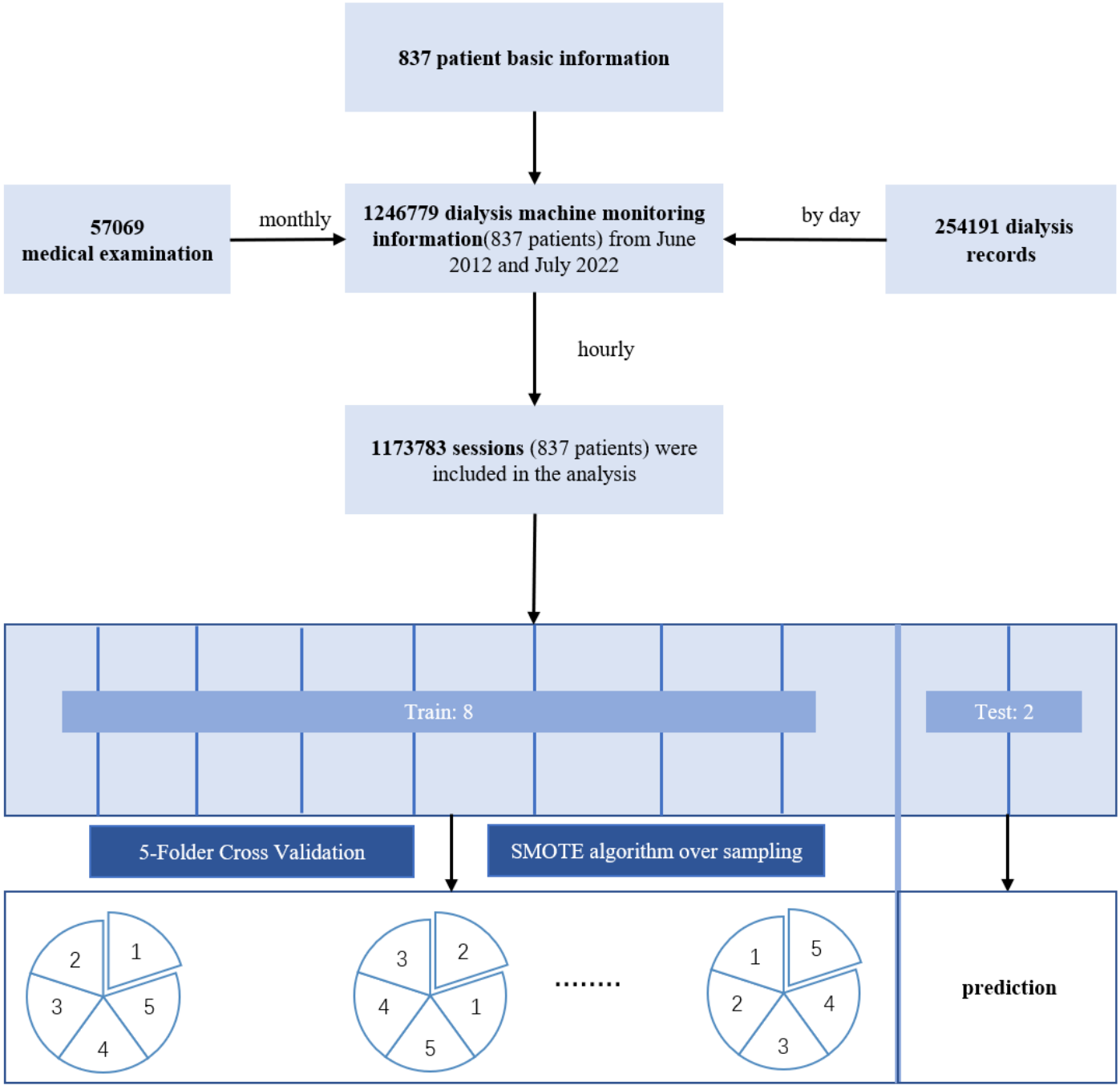


Figure 3

the pre-process of data

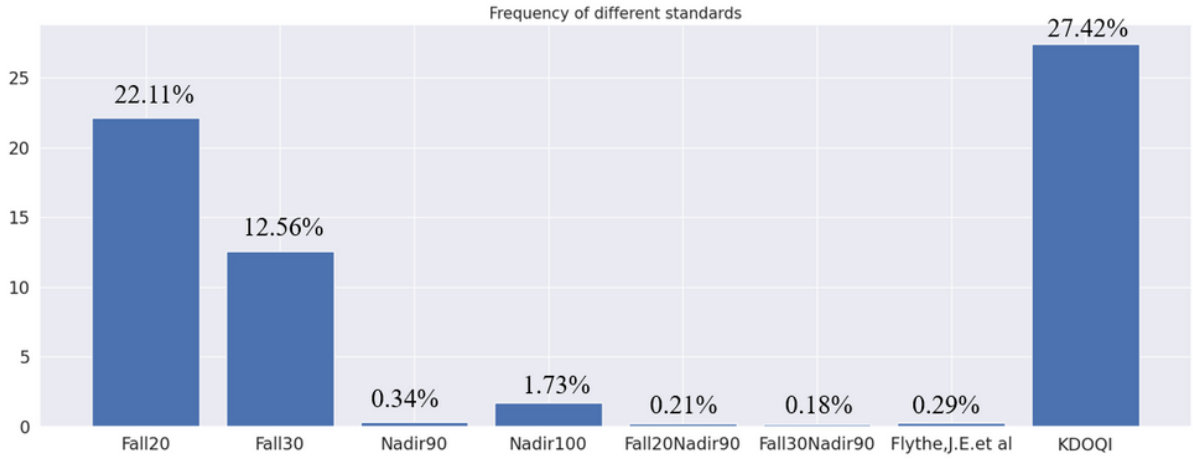


Figure 4

the frequency of different standards of the IDH