DOMAIN 1: Participants			
A. Risk of Bias			
Describe the sources of data and criteria for participant selection: HD-patients' demographic data and hemodialysis records from Hemodialysis Quality Control Center of Shenzhen City between Jan, 2019 and Sep, 2021 were collected. In our study, the target population was 17-to-100-year-old individuals who are indicated for a 3-to-4-hour HD due to other than acute kidney injury.			
		Dev	Val
1.1 Were appropriate data sources used, e.g. cohort, RCT or nested case-control study data?			Yes
1.2 Were all inclusions and exclusions of participants appropriate?			Yes
Risk of bias introduced by selection of participants	**RISK:** *(low/ high/ unclear)*		Low
Rationale of bias rating: All the missing data in HD sessions was missing at random and no handled by data processing. As sensitivity analysis for missing data handling, the continuous numerical variables were imputed by median and categorical variables by the most common value. The results shown the LightGBM method both had good performance in the two types of dataset which indicated slight affect of the selection of participants and HD sessions to LightGBM predictor.			
B. Applicability			
Describe included participants, setting and dates: All the missing data in HD sessions was missing at random and no handled by data processing. As sensitivity analysis for missing data handling, the continuous numerical variables were imputed by median and categorical variables by the most common value. The results shown the LightGBM method both had good performance in the two types of dataset which indicated slight affect of the selection of participants and HD sessions to LightGBM predictor.			
Concern that the included participants and setting do not match the review question	**CONCERN:** *(low/ high/ unclear)*		Low
Rationale of applicability rating: Uniform inclusion and exclusion criteria of participants and HD sessions were used in the training and test dataset in IDH-A and IDH-B model, respectively.			

DOMAIN 2: Predictors			
A. Risk of Bias			
List and describe predictors included in the final model, e.g. definition and timing of assessment: This study developed and internally validated two models (IDH-A and IDH-B) at two different times for predicting IDH incident. The prediction times were those: (1) after a HD session for the outcome at the next HD session (i.e., IDH-A). The time interval between IDH-A model and outcome assessment was two to four days varied in HD patients with twice-weekly or tri-weekly hemodialysis sessions. (2) immediately before an HD session for the outcome at that session (i.e., IDH-B). The time interval between IDH-B model and outcome assessment was relatively short, ranged three to four hours.			
		Dev	Val
2.1 Were predictors defined and assessed in a similar way for all participants?			Yes
2.2 Were predictor assessments made without knowledge of outcome data?			Yes
2.3 Are all predictors available at the time the model is intended to be used?			Yes
Risk of bias introduced by predictors or their assessment	**RISK:** *(low/ high/ unclear)*		Low
Rationale of bias rating: Two types of IDH predictors were conducted for two different times for predicting IDH incident. To excluded the bias came from different definitions of IDH, this study perform IDH predictors based on five other IDH definitions that reviewed before to prove the stability of predictor.			
B. Applicability			
Concern that the definition, assessment or timing of predictors in the model do not match the review question	**CONCERN:** *(low/ high/ unclear)*		Low
Rationale of applicability rating: Uniform criteria of definition and assessment of IDH were used in this study.			

DOMAIN 3: Outcome

A. Risk of Bias

Describe the outcome, how it was defined and determined, and the time interval between predictor assessment and outcome determination: The outcome was defined clearly in this study (ie, IDH of Fall30Nadir90 means pre-HD SBP - nadir SBP ≥30mmHg and SBP nadir<90mmHg. HD-interval SBP recorded per half-hour during HD session and pre-SBP collected 10 to 15 minutes before that HD session were used to assess the IDH outcome. When the HD-interval SBP meet the criteria of IDH at least one time, the HD session was considered to happen IDH event.

		Dev	Val
3.1 Was the outcome determined appropriately?			Yes
3.2 Was a pre-specified or standard outcome definition used?			Yes
3.3 Were predictors excluded from the outcome definition?			Yes
3.4 Was the outcome defined and determined in a similar way for all participants?			Yes
3.5 Was the outcome determined without knowledge of predictor information?			Yes
3.6 Was the time interval between predictor assessment and outcome determination appropriate?			Yes
Risk of bias introduced by the outcome or its determination	**RISK:** *(low/ high/ unclear)*		Low

Rationale of bias rating: The outcome was defined clearly in this study (ie, IDH of Fall30Nadir90 means pre-HD SBP - nadir SBP ≥30mmHg and SBP nadir<90mmHg. And the uniform criteria of definition of IDH accident was used in this study

B. Applicability

At what time point was the outcome determined: For HD patients, IDH-A model can use to predict risk after a HD session for the outcome at the next HD session (with prediction time in two to four days depending on the frequency of HD treatment); IDH-B model can use to predict risk immediately before an HD session for the outcome at that session (with prediction time in the next three to four hours).

If a composite outcome was used, describe the relative frequency/distribution of each contributing outcome:

Concern that the outcome, its definition, timing or determination do not match the review question	**CONCERN:** *(low/ high/ unclear)*		Low

Rationale of applicability rating: The uniform criteria of definition, timing of IDH accident was used in this study. The risk of bias came from the outcome was Low.

DOMAIN 4: Analysis

Risk of Bias

Describe numbers of participants, number of candidate predictors, outcome events and events per candidate predictor: , In the IDH definition of Fall30Nadir90, IDH-A model was conducted based on 27 variables from 62,227 HD sessions, in which contained 2365 IDH events. IDH-B model was conducted based on 10 variables from 64,870 HD sessions, in which contained 2507 IDH events.

Describe how the model was developed (for example in regards to modelling technique (e.g. survival or logistic modelling), predictor selection, and risk group definition): LightGBM, support vector machine, linear discriminant analysis, eXtreme Gradient Boosting , TabNet, and multilayer perceptron were utilized for prediction models. Variables who meet including criteria were collected from demographic characteristics and dialysis sessions . 27 variables for IDH-A model and 10 variables for IDH-B model were used for model development at last. The outcome was defined clearly in this study (ie, IDH of Fall30Nadir90 means pre-HD SBP - nadir SBP ≥30mmHg and SBP nadir<90mmHg.

Describe whether and how the model was validated, either internally (e.g. bootstrapping, cross validation, random split sample) or externally (e.g. temporal validation, geographical validation, different setting, different type of participants): This study developed and internally validated two models (IDH-A and IDH-B) by random split HD sessions.

Describe the performance measures of the model, e.g. (re)calibration, discrimination, (re)classification, net benefit, and whether they were adjusted for optimism: Area Under the receiver operating characteristic curve (ROC AUC) was utilized to assess model performance by calculating the C-statistic. Delong test was utilized to compared the area under the curves in different predictors. And the parameters of model was adjusted to achieve the optimal.

Describe any participants who were excluded from the analysis: The excluded participants and HD sessions were clearly described in Figure S1 and Figure S2.

Describe missing data on predictors and outcomes as well as methods used for missing data:
The excluded participants and HD sessions were clearly described in Figure S1 and Figure S2. All the missing data was described in Table S2. To exclude the risk of bias brought by the approach for missing data handling, we imputed the continuous variables by median and categorical variables by the most common value and conducted sensitivity analysis with LightGBM method.

		Dev	Val
4.1 Were there a reasonable number of participants with the outcome?			Yes
4.2 Were continuous and categorical predictors handled appropriately?			Yes
4.3 Were all enrolled participants included in the analysis?			Yes
4.4 Were participants with missing data handled appropriately?			Yes
4.5 Was selection of predictors based on univariable analysis avoided?			No involved
4.6 Were complexities in the data (e.g. censoring, competing risks, sampling of controls) accounted for appropriately?			Yes
4.7 Were relevant model performance measures evaluated appropriately?			Yes
4.8 Were model overfitting and optimism in model performance accounted for?			Yes
4.9 Do predictors and their assigned weights in the final model correspond to the results from multivariable analysis?			No involved
Risk of bias introduced by the analysis	**RISK:** *(low/ high/ unclear)*		Low

Rationale of bias rating:

Step 4: Overall assessment

Use the following tables to reach overall judgements about risk of bias and concerns regarding applicability of the prediction model evaluation (development and/or validation) across all assessed domains.
Complete for each evaluation of a distinct model.

Reaching an overall judgement about risk of bias of the prediction model evaluation	
Low risk of bias	If all domains were rated low risk of bias. If a prediction model was developed without any external validation, and it was rated as low risk of bias for all domains, consider downgrading to **high risk of bias**. Such a model can only be considered as low risk of bias, if the development was based on a very large data set and included some form of internal validation.
High risk of bias	If at least one domain is judged to be at **high risk of bias**.
Unclear risk of bias	If an unclear risk of bias was noted in at least one domain and it was low risk for all other domains.

Reaching an overall judgement about applicability of the prediction model evaluation	
Low concerns regarding applicability	If low concerns regarding applicability for all domains, the prediction model evaluation is judged to have **low concerns regarding applicability**.
High concerns regarding applicability	If high concerns regarding applicability for at least one domain, the prediction model evaluation is judged to have **high concerns regarding applicability**.
Unclear concerns regarding applicability	If unclear concerns (but no "high concern") regarding applicability for at least one domain, the prediction model evaluation is judged to have **unclear concerns regarding applicability** overall.

Overall judgement about risk of bias and applicability of the prediction model evaluation		
Overall judgement of risk of bias	**RISK:** *(low/ high/ unclear)*	Low
Summary of sources of potential bias: The excluded participants and HD sessions were clearly described in Figure S1 and Figure S2. All the missing data was described in Table S2. The subgroup analysis and sensitivity analysis were explored in this study. And the risk of bias brought by the approach for missing data handling and clusters heterogeneity were explored in this study.		
Overall judgement of applicability	**CONCERN:** *(low/ high/ unclear)*	Low
Summary of applicability concerns: For HD patients, IDH-A model can use to predict risk after a HD session for the outcome at the next HD session (with prediction time in two to four days depending on the frequency of HD treatment); IDH-B model can use to predict risk immediately before an HD session for the outcome at that session (with prediction time in the next three to four hours).		