Supplemental Material

Deep Learning Model for Real-Time Prediction of Intradialytic Dypotension

Hojun Lee, Donghwan Yun, Jayeon Yoo, Kiyoon Yoo, Yong Chul Kim, Dong Ki Kim, Kook-Hwan Oh, Kwon Wook Joo, Yon Su Kim, Nojun Kwak and Seung Seok Han

Supplemental Table 10. Performance of the recurrent neural network model in sessions with and without echocardiographic information

Supplemental Table 11. Area under the curves for predicting the differently defined intradialytic hypotension

Supplemental Figure 1. Exploratory data analysis of the rate of intradialytic hypotension (IDH).

Supplemental Figure 2. Confusion matrix plot for (A) IDH-1, (B) IDH-2, and (C) IDH-3. The case numbers are given in each cell.

Supplemental Figure 3. Precision-recall graph according to the hemodialysis time. The data output thresholds were set as 0.1, 0.3, 0.5, 0.7, and 0.9. (A) IDH-1. (B) IDH-2. (C) IDH-3.

Supplemental Figure 4. Decision curve analysis of recurrent neural network (RNN) and three other models. (A) Intradialytic hypotension (IDH)-1. (B) IDH-2. (C) IDH-3.

Supplemental Figure 5. Platt scaling plot used to calibrate the models. (A) Intradialytic hypotension (IDH)-1. (B) IDH-2. (C) IDH-3. Bin size = 0.05.

Supplemental Methods

Study variables

The baseline clinical information was age, sex, vital signs (systolic blood pressure, diastolic blood pressure, heart rate, and body temperature), hemodialysis settings (type of hemodialysis [hemodialysis, hemodiafiltration, hemoperfusion, hemofiltration, hemofiltration reinfusion, and supplementary ultrafiltration], blood flow rate, dialysate flow rate, target and time-varying amounts of ultrafiltration, time setting, access route [arteriovenous fistula and graft, temporary catheterization via internal jugular and femoral vein, subcutaneously-tunneled catheter, sheath for coronary angiography, and Hickman catheter], pre-dialysis weight, use of anti-coagulant [heparin and nafamostat mesilate], priming fluid [normal saline, half saline, and red blood cells], dialysate [e.g., Hemo B Dex 0.1% and 0.15%, and Hemotrate-B1] and dialyzer [e.g., APS-15U, APS-21U, Rexeed-13LX, Rexeed-18LX, BLS 812G, BLS 812SD, BLS 814SD, BLS 816SD, BLS 819SD, NC 1485, PHF0714, SG30, Adsorba, polyflux 14, polyflux 14H, polyflux 14L, polyflux 14S, polyflux 170H, polyflux 17L, polyflux 17S, polyflux 6H, polyflux 8L, polyflux S, Theranova 400, F4 HPS, F5 HPS, F6 HPS, FX, FX paed, FX5, FX8, FX40, FX50, FX80, FB 130T, and Sureflux 130E-GA]), the dialysate temperature and concentrations of sodium, potassium, calcium, and bicarbonate, incident or prevalent sessions, admission status, the presence of comorbidities (e.g., diabetes mellitus, hypertension, cardiovascular disease, and kidney transplantation), the number of session per week, the history of IDH within one week, total number of sessions with IDH within one week, and medications used before initiating hemodialysis. Laboratory blood findings were measured at the beginning of the hemodialysis sessions, including white blood cells, hemoglobin, platelet, cholesterol, albumin, glucose, calcium, phosphate, uric acid, blood urea nitrogen, creatinine, sodium, potassium, chloride, and total carbon dioxide. There were no missing variables.

Echocardiographic information before hemodialysis sessions including left ventricular ejection fraction, left ventricular end-diastolic dimension, left ventricular end-systolic dimension, interventricular septum thickness, and left ventricular mass was available for 227,640 (87%) sessions and 7256 (78%) patients. This information was used in a sensitivity analysis.

Model development

Statistical analyses were performed using R software (version 3.5.1; The Comprehensive R Archive Network: http://cran.r-project.org) and Python (version 3.6.8; Python Software Foundation: http://www.python.org). The PyTorch 1.3 was used as a deep learning framework throughout this process (1).

The categorical and continuous variables of the baseline characteristics are presented as proportions and means ± standard deviation, respectively. The dataset was treated as follows: S = [(x1,1, y1,1), …, (x1,L1, y1,L1), (x2,1, y2,1), …, (xd,Ld, yd,Ld)], where d and Ld indicate the number of dialysis cases and the frame number of the dth dialysis, respectively. The ground truth labels were denoted as yp,q = [yp,q,IDH-1, yp,q,IDH-2(initial), yp,q,IDH-2(present)], where 0 was normal and 1 was abnormal (i.e., IDH). When converting the data in the training dataset into vectors, the continuous features were standardized with a mean of 0 and a variance of 1, and the categorical features were transformed into binary variables (i.e. 0 or 1) by one-hot encoding. The dataset S was used as the training data to train the recurrent neural network, multilayer perceptron, Light Gradient Boosting Machine, and logistic regression models.

Binary cross-entropy loss was used as the loss function for the recurrent neural network to calculate the difference between actual labels and predictions. We used the Adam optimization method as the optimizer (2). The pseudocode for the recurrent neural network is given below.

Pseudocode 1: Predicting IDH Using recurrent neural network Evaluation

Function IDH-prediction (Trained Network, $X_{p,1:L_p}$, σ)

  Input: Trained Network: BatchNorm, RNN, MLP-1, MLP-2, MLP-3;

    σ: logistic function;

    $X_{p,1:L_p} = (x_1, x_2, \ldots, x_{L_p})$: entries of p-th dialysis of test data;

  Output : $\hat{y}_{p,1:L_p} = [(\hat{y}_{1,IDH-1}, \hat{y}_{1,IDH-2(init)}, \hat{y}_{1,IDH-3(present)}), \ldots, (\hat{y}_{p,IDH-1}, \hat{y}_{p,IDH-2(init)}, \hat{y}_{p,IDH-3(present)})]$: predictions for p-th dialysis

  Timestamp index t ← 1

  for t =1 to L_p:

    $x_{t,batchnorm} \leftarrow$ BatchNorm(x_t)

    $h_t \leftarrow$ RNN ($x_{t,batchrnom}$, h_{t-1})

    $\hat{y}_{t,IDH-1} \leftarrow \sigma($ MLP-1(h_t))

    $\hat{y}_{t,\mathrm{IDH-2(init)}} \leftarrow \sigma($ MLP-2(h_t))

    $\hat{y}_{t,\mathrm{IDH-3(present)}} \leftarrow \sigma($ MLP-3(h_t))

    t ← t+1

$y_{p,1:L_p} = [(y_{1,IDH-1}, y_{1,IDH-2(init)}, y_{1,IDH-3(present)}), \ldots, (y_{L_p,IDH-1}, y_{L_p,IDH-2(init)}, y_{L_p,IDH-3(present)})]$: ground truth labels of p-th dialysis

12	Quantitative evaluation using the actual labels $y_{p,1:L_p}$ and the prediction $\hat{y}_{p,1:L_p}$

RNN, recurrent neural network; MLP, multilayer perceptron

The multilayer perceptron algorithm consists of a series of non-linear functions and fully-connected layers that are affine transforms as follows: $\hat{y}_{p,q} = \sigma_n \circ f_n \circ \sigma_{n-1} \circ f_{n-1} \circ \ldots \circ \sigma_1 \circ f_1(x_{p,q})$ (3). The σ_j is the jth non-linear function (e.g., $\sigma_j(x) = \max(0, x)$) and the f_j is the jth fully-connected layer (i.e., affine transform). Throughout this calculation, the multilayer perceptron can extract meaningful information on higher dimensions of the input vector. For a probability model, the last σ_n is a logistic function. The binary cross-entropy loss and the Adam optimization methods were used (2). The architecture of the multilayer perceptron is shown below.

Layer	Shape
BatchNorm	260 (input feature size)
Fully connected layer + ReLU	260 × 256
Fully connected layer + BatchNorm + ReLU	256 × 256
Fully connected layer + ReLU	256 × 256
Fully connected layer + ReLU	256 × 256
Fully connected layer	256 × 3
Logistic function	3

BatchNorm, batch normalization; ReLU, rectified linear unit

Light Gradient Boosting Machine combines learned ‘Learners’ after learning several weak ‘Learners’ (4). Throughout the learning and predicting process for weak ‘Learners’, the error is lowered by gradient boosting on residuals of incorrectly predicted results. The Light Gradient Boosting Machine method is faster than the other Gradient Boosting Machine methods such as extreme gradient boosting (4). Logistic regression calculates the weighted sum of the feature vector that is derived from regression coefficients and feature values.

Feature ranking analysis

To estimate how much features contribute to the prediction of IDH, we use the feature ranking method proposed in the previous paper (5). This method drops each feature one by one from the test dataset when the model inference and compares the prediction results to the reference prediction result which is gained without losing any features. Large prediction differences between dropped data and full-featured data represent that the dropped features have contributed much more when the model makes predictions.

$$score_{fdrop} = \frac{1}{N}\sum_{i=1}^{N} |p_{i,fdrop} - p_i| \quad \ldots (A)$$

To apply this feature ranking method to our approach, we modify it to suit our settings as shown in equation (A). In eq (A), f_{drop} is the feature we focus on and drop from the input data. p_i means the reference prediction result of the i-th data which the model inferences using all features and $p_{i,fdrop}$ means that the prediction result of the i-th data when the feature f_{drop} is dropped. These absolute values are averaged over all dataset of size N. The scores were calculated for IDH-1, IDH-2, and IDH-3 respectively.

The batch norm layer serves as a standardization. To drop each feature, we have set each output of the batch norm layer as 0. The prediction result may be higher or lower than reference prediction, however, to measure the degree of difference between dropped feature data and full featured data we average the absolute value of differences.

References

1. Paszke A, Gross S, Chintala S, Chanan G, Yang E, DeVito Z, Lin Z, Desmaison A, Antiga L, Lerer A: Automatic differentiation in PyTorch. *In NIPS-W* 2017.

2. Kingma DP, Ba J: Adam: A method for stochastic optimization. *arXiv preprint arXiv:14126980*, 2014

3. Pal SK, Mitra S: Multilayer perceptron, fuzzy sets, and classification. *IEEE Trans Neural Netw* 3: 683-697, 1992.

4. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q, Liu T-Y: LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *31st Conference on Neural Information Processing Systems (NIPS 2017)* 3148-3156, 2017.

5. Ginley B, Lutnick B, Jen KY, Fogo AB, Jain S, Rosenberg A, et al.: Computational Segmentation and Classification of Diabetic Glomerulosclerosis. *J Am Soc Nephrol*, 30**:** 1953-1967, 2019.

Supplemental Table 1. Characteristics of hemodialysis sessions that were included and excluded from the analyses

Variables	Excluded (n = 20,752)	Included (n = 261,647)
Age (years)	19 ± 20	62 ± 15
Male (%)	58	58
Diabetes mellitus (%)	15	48
Hypertension (%)	57	68

Supplemental Table 2. Description of the four models used in the present study

Models	Descriptions	Comparison with other models
Logistic regression	Output the probability of a certain event happening by using a linear classifier followed by a logistic function. It is able to model the linear relationship between the input features and the target.	No inherent way of modeling temporal information
Light Gradient Boosting Machine	Composed of ensemble of tree-based models. It is able to make accurate predictions by iteratively boosting the errors of individual models.	Able to compute feature importance of the input data Able to model non-linear relationships No inherent way of modeling temporal information
Multilayer perceptron	A class of deep neural network that can model complex relationships by using multiple hidden layers. Similar to above methods, only fixed-size input is taken.	No explicit way of computing feature importance (black-box model). Feature selection method should be applied. No inherent way of modeling temporal information
Recurrent neural network	A class of deep neural network that assumes sequential data as inputs with temporal relationship. It does this by continuously utilizing previous sequences for predicting the target of current sequence.	Inherently models input sequentially No explicit way of computing feature importance (black-box model). Feature selection method should be applied.

Supplemental Table 3. Summarized table of feature sets of the model input

Feature sets	Variables
A (age and sex)	Age and sex
B (hemodialysis-related features)	Type of hemodialysis, dialysate flow rate, blood flow rate, target and time-varying amounts of ultrafiltration, target ultrafiltration, time setting, vascular access route, pre-dialysis weight, anti-coagulants, priming fluids, dialyzer type, dialysate sodium concentration, dialysate potassium concentration, dialysate calcium concentration, dialysate bicarbonate concentration, dialysate temperature, and incident or prevalent hemodialysis.
C (vital signs)	Systolic blood pressure, diastolic blood pressure, heart rate, and body temperature.
D (clinical information)	Diabetes mellitus, hypertension, cardiovascular disease, kidney transplant donor, kidney transplant recipient, number of dialysis session per week, previous history of IDH within 7 days, and total number of sessions with IDH within 7 days.
E (laboratory findings)	White blood cell count, hemoglobin, platelet, cholesterol, albumin, glucose, calcium, phosphate, uric acid, blood urea nitrogen, creatinine, sodium, potassium, chloride, and bicarbonate.
F (medications)	Beta-blockers, calcium channel blockers, angiotensin-converting enzyme inhibitors, aldosterone receptor blockers, diuretics, lipid-lowering agents, minoxidil, aspirin, adenosine diphosphate receptor inhibitors, warfarin, oral hypoglycemic agents, insulin, allopurinol, febuxostat, erythropoietin-stimulating agents, calcium-based phosphate binder, and non-calcium-based phosphate binders.

Supplemental Table 4. F1 score for predicting intradialytic hypotension in deep and other machine learning, and logistic regression models

		Outcome		
Threshold	Models	IDH-1	IDH-2	IDH-3
0.1	RNN	0.5028	0.6116	0.4391
	MLP	0.5096	0.6114	0.4391
	LightGBM	0.5025	0.6127	0.4301
	LR	0.4788	0.5754	0.4102
0.3	RNN	0.5844	0.6950	0.4981
	MLP	0.5717	0.6916	0.4846
	LightGBM	0.5759	0.6963	0.4800
	LR	0.5498	0.6763	0.4484
0.5	RNN	0.5197	0.6732	0.3682
	MLP	0.4765	0.6637	0.3522
	LightGBM	0.4912	0.6702	0.3186
	LR	0.4523	0.6271	0.2612
0.7	RNN	0.3542	0.5631	0.1714
	MLP	0.2829	0.5479	0.1685
	LightGBM	0.3002	0.5589	0.0961
	LR	0.2546	0.4873	0.0829
0.9	RNN	0.1275	0.3357	0.0163
	MLP	0.0706	0.3120	0.0102
	LightGBM	0.0367	0.2689	0.0014
	LR	0.0454	0.2547	0.0071

IDH-1, intradialytic hypotension defined as nadir systolic blood pressure <90 mmHg; IDH-2, intradialytic hypotension defined as decrease in systolic blood pressure ≥20 mmHg and/or decrease in mean arterial pressure ≥10 mmHg based on blood pressure at initial time point; IDH-3, intradialytic hypotension defined as decrease in systolic blood pressure ≥20 mmHg and/or decrease in mean arterial pressure ≥10 mmHg based on blood pressure at prediction time point.

BP, blood pressure; RNN, recurrent neural network, MLP, multilayer perceptron; LightGBM, Light Gradient Boosting Machine; LR, logistic regression.

Supplemental Table 5. F1 scores of the recurrent neural network after ablation of the feature set

		Outcome		
Threshold	Models	IDH-1	IDH-2	IDH-3
0.1	Remove set A	0.5051	0.6066	0.4366
	Remove set B	0.5002	0.5950	0.4174
	Remove set C	0.4140	0.5086	0.3683
	Remove set D	0.4965	0.6133	0.4321
	Remove set E	0.5043	0.6072	0.4317
	Remove set F	0.5041	0.6103	0.4354
0.3	Remove set A	0.5818	0.6918	0.4983
	Remove set B	0.5700	0.6836	0.4766
	Remove set C	0.4413	0.5592	0.3716
	Remove set D	0.5733	0.6896	0.4813
	Remove set E	0.5809	0.6943	0.4972
	Remove set F	0.5834	0.6948	0.4965
0.5	Remove set A	0.5129	0.6762	0.3705
	Remove set B	0.4760	0.6594	0.3266
	Remove set C	0.2851	0.4156	0.1161
	Remove set D	0.5120	0.6560	0.3315
	Remove set E	0.5028	0.6697	0.3569

	Remove set F	0.5154	0.6739	0.3630
0.7	Remove set A	0.3438	0.5697	0.1614
	Remove set B	0.2991	0.5270	0.1257
	Remove set C	0.1018	0.1605	0.0122
	Remove set D	0.3652	0.5340	0.1286
	Remove set E	0.3328	0.5538	0.1454
	Remove set F	0.3471	0.5647	0.1657
0.9	Remove set A	0.1190	0.3358	0.0105
	Remove set B	0.0880	0.2398	0.0051
	Remove set C	0.0064	0.0046	0.0000
	Remove set D	0.1381	0.3119	0.0068
	Remove set E	0.1119	0.3264	0.0095
	Remove set F	0.1263	0.3295	0.0154

IDH-1, intradialytic hypotension defined as nadir systolic blood pressure <90 mmHg; IDH-2, intradialytic hypotension defined as decrease in systolic blood pressure ≥20 mmHg and/or decrease in mean arterial pressure ≥10 mmHg based on blood pressure at initial time point; IDH-3, intradialytic hypotension defined as decrease in systolic blood pressure ≥20 mmHg and/or decrease in mean arterial pressure ≥10 mmHg based on blood pressure at prediction time point.

Each set contains features as follows: A, age and sex; B, hemodialysis-related features; C, vital signs; D, comorbidities; E, laboratory findings;

and F, medications.

Supplemental Table 6. Area under the curves for predicting intradialytic hypotension in incidence and prevalent hemodialysis sessions

Outcomes	Groups	AUROC (95% CI)	AUPRC (95% CI)
IDH-1	Incident	0.944 (0.941–0.947)	0.652 (0.650–0.653)
	Prevalent	0.936 (0.935–0.938)	0.610 (0.608–0.612)
IDH-2	Incident	0.872 (0.868–0.876)	0.783 (0.781–0.784)
	Prevalent	0.869 (0.868–0.871)	0.780 (0.779–0.782)
IDH-3	Incident	0.772 (0.766–0.777)	0.473 (0.471–0.475)
	Prevalent	0.792 (0.790–0.795)	0.516 (0.515–0.518)

AUROC, area under the receiver operating characteristic curve; CI, confidence interval; AUPRC, area under the precision-recall curve; IDH, intradialytic hypotension.

Supplemental Table 7. Area under the curves for predicting intradialytic hypotension according to the tertiles of ultrafiltration

Outcomes	Ultrafiltration	AUROC (95% CI)	AUPRC (95% CI)
IDH-1	1st tertile	0.951 (0.948–0.953)	0.667 (0.665–0.668)
	2nd tertile	0.935 (0.932–0.937)	0.593 (0.591–0.594)
	3rd tertile	0.926 (0.924–0.929)	0.590 (0.588–0.592)
IDH-2	1st tertile	0.870 (0.868–0.872)	0.745 (0.744–0.747)
	2nd tertile	0.874 (0.872–0.876)	0.789 (0.787–0.790)
	3rd tertile	0.862 (0.860–0.864)	0.799 (0.797–0.800)
IDH-3	1st tertile	0.788 (0.784–0.792)	0.477 (0.475–0.479)
	2nd tertile	0.799 (0.796–0.803)	0.525 (0.523–0.527)
	3rd tertile	0.780 (0.777–0.784)	0.525 (0.523–0.526)

AUROC, area under the receiver operating characteristic curve; CI, confidence interval; AUPRC, area under the precision-recall curve; IDH, intradialytic hypotension.

Supplemental Table 8. Area under the curves for predicting intradialytic hypotension based on a history of intradialytic hypotension

Outcomes	History of IDH	AUROC (95% CI)	AUPRC (95% CI)
IDH-1	Absent	0.938 (0.936–0.941)	0.502 (0.500–0.504)
	Present	0.860 (0.857–0.863)	0.664 (0.662–0.665)
IDH-2	Absent	0.884 (0.881–0.886)	0.748 (0.746–0.749)
	Present	0.858 (0.856–0.860)	0.791 (0.789–0.792)
IDH-3	Absent	0.792 (0.788–0.796)	0.446 (0.445–0.448)
	Present	0.784 (0.781–0.786)	0.528 (0.526–0.530)

IDH, intradialytic hypotension; AUROC, area under the receiver operating characteristic curve; CI, confidence interval; AUPRC, area under the precision-recall curve.

Supplemental Table 9. Performance of the recurrent neural network model after using sessions-stratified randomization

Outcome	AUROC (95% CI)	AUPRC (95% CI)
IDH-1	0.943 (0.941–0.944)	0.659 (0.657–0.661)
IDH-2	0.877 (0.876–0.878)	0.788 (0.787–0.789)
IDH-3	0.793 (0.791–0.794)	0.508 (0.506–0.510)

AUROC, area under the receiver operating characteristic curve; CI, confidence interval; AUPRC, area under the precision-recall curve; IDH, intradialytic hypotension.

Supplemental Table 10. Performance of the recurrent neural network model in sessions with and without echocardiographic information

	Model without echocardiographic information		Model with echocardiographic information	
Outcomes	AUROC (95% CI)	AUPRC (95% CI)	AUROC (95% CI)	AUPRC (95% CI)
IDH-1	0.938 (0.936–0.939)	0.643 (0.642–0.645)	0.938 (0.936–0.939)	0.643 (0.642–0.645)
IDH-2	0.875 (0.874–0.876)	0.783 (0.781–0.784)	0.875 (0.873–0.876)	0.782 (0.781–0.784)
IDH-3	0.800 (0.797–0.802)	0.522 (0.521–0.524)	0.799 (0.797–0.801)	0.522 (0.520–0.524)

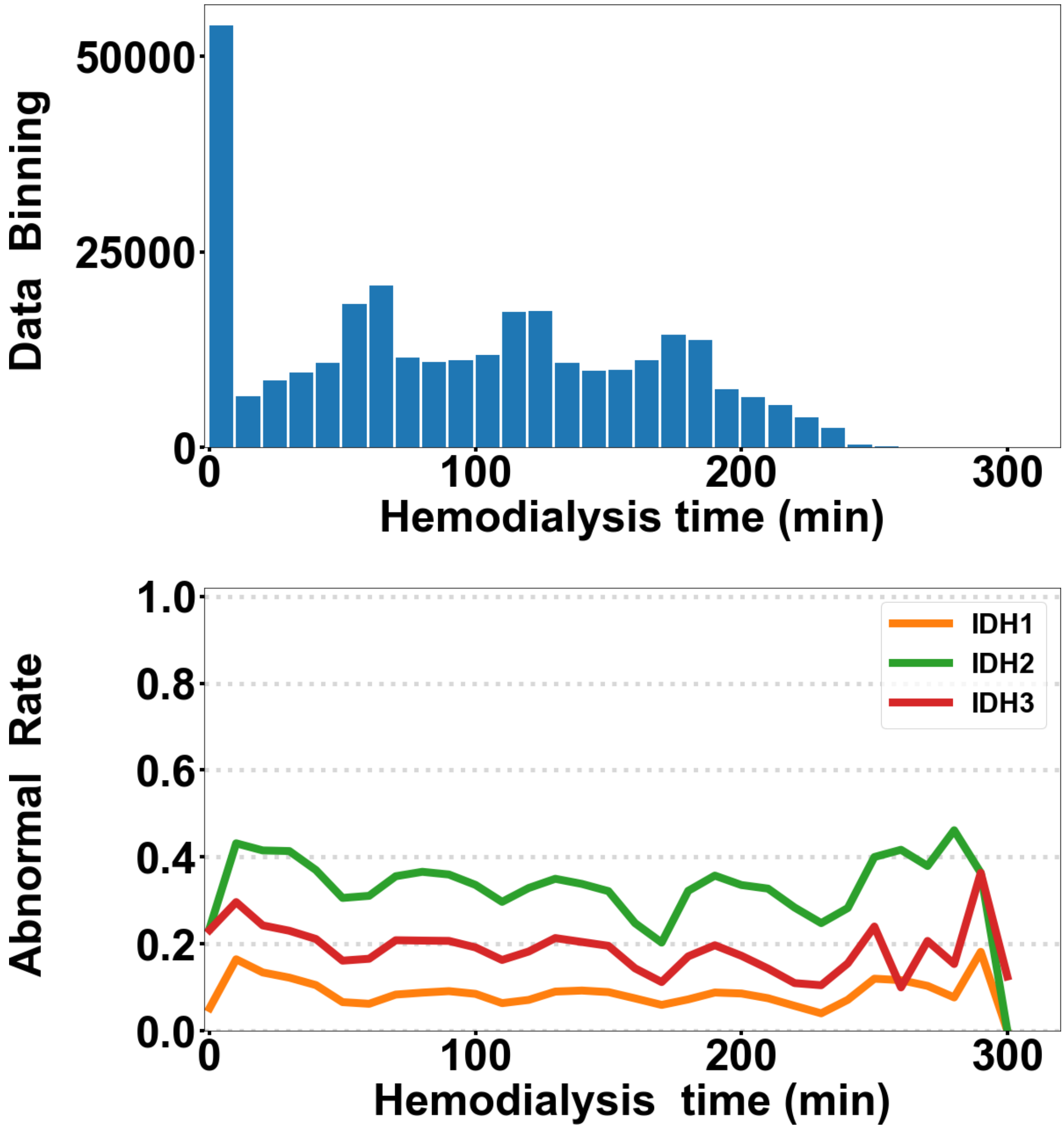
AUROC, area under the receiver operating characteristic curve; CI, confidence interval; AUPRC, area under the precision-recall curve; IDH, intradialytic hypotension.

Supplemental Table 11. Area under the curves for predicting the differently defined intradialytic hypotension

Outcomes	Models	AUROC (95% CI)	*P* value	AUPRC (95% CI)
IDH-4	RNN	0.930 (0.929–0.932)		0.742 (0.740–0.744)
	MLP	0.928 (0.926–0.929)	<0.001	0.731 (0.729–0.732)
	LightGBM	0.928 (0.927–0.929)	<0.001	0.731 (0.730–0.733)
	Logistic regression	0.916 (0.914–0.917)	<0.001	0.694 (0.692–0.696)
IDH-5	RNN	0.888 (0.887–0.890)		0.724 (0.722–0.726)
	MLP	0.884 (0.882–0.885)	<0.001	0.715 (0.714–0.717)
	LightGBM	0.887 (0.885–0.888)	<0.001	0.715 (0.714–0.717)
	Logistic regression	0.872 (0.871–0.874)	<0.001	0.687 (0.685–0.688)

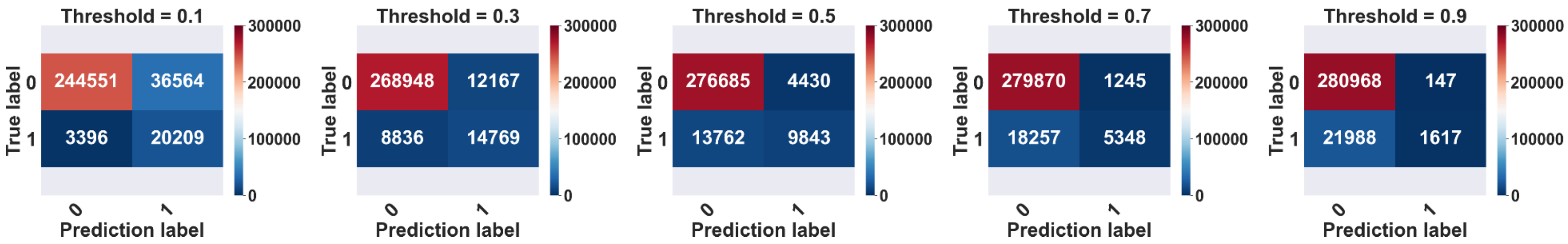
AUROC, area under the receiver operating characteristic curve; CI, confidence interval; AUPRC, area under the precision-recall curve; IDH, intradialytic hypotension; RNN, recurrent neural network, MLP, multilayer perceptron; LightGBM, Light Gradient Boosting Machine; LR, logistic regression.

Supplemental Figure 1. Exploratory data analysis of the rate of intradialytic hypotension (IDH).

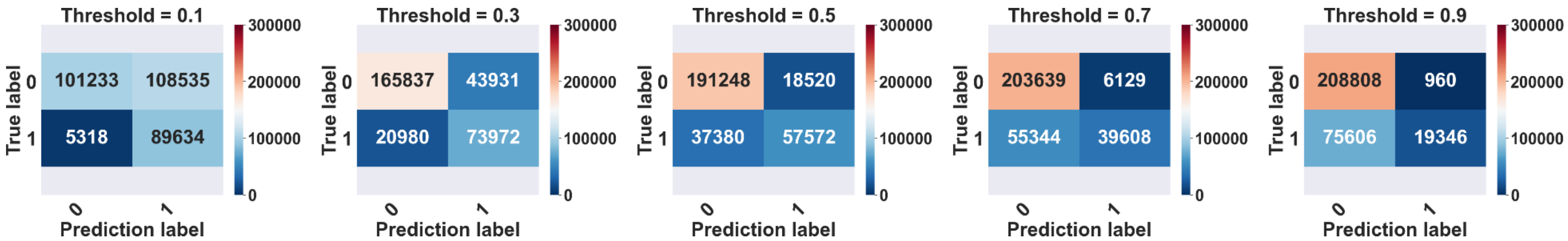


Supplemental Figure 2. Confusion matrix plot for (A) IDH-1, (B) IDH-2, and (C) IDH-3. The case numbers are given in each cell.

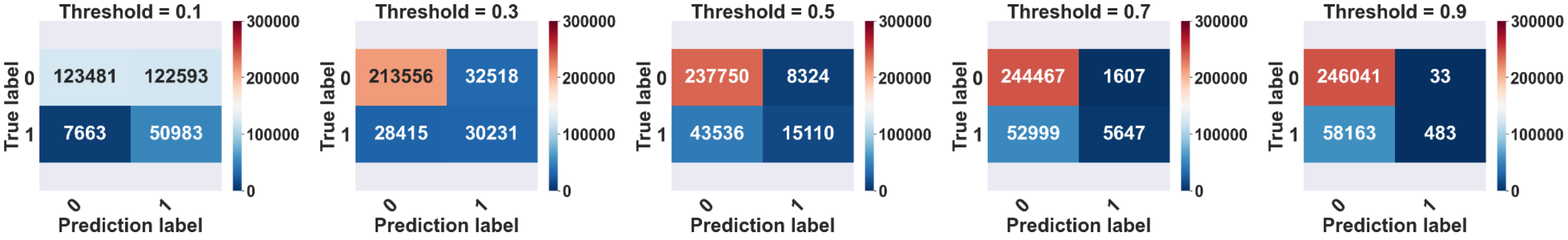
(A)



(B)



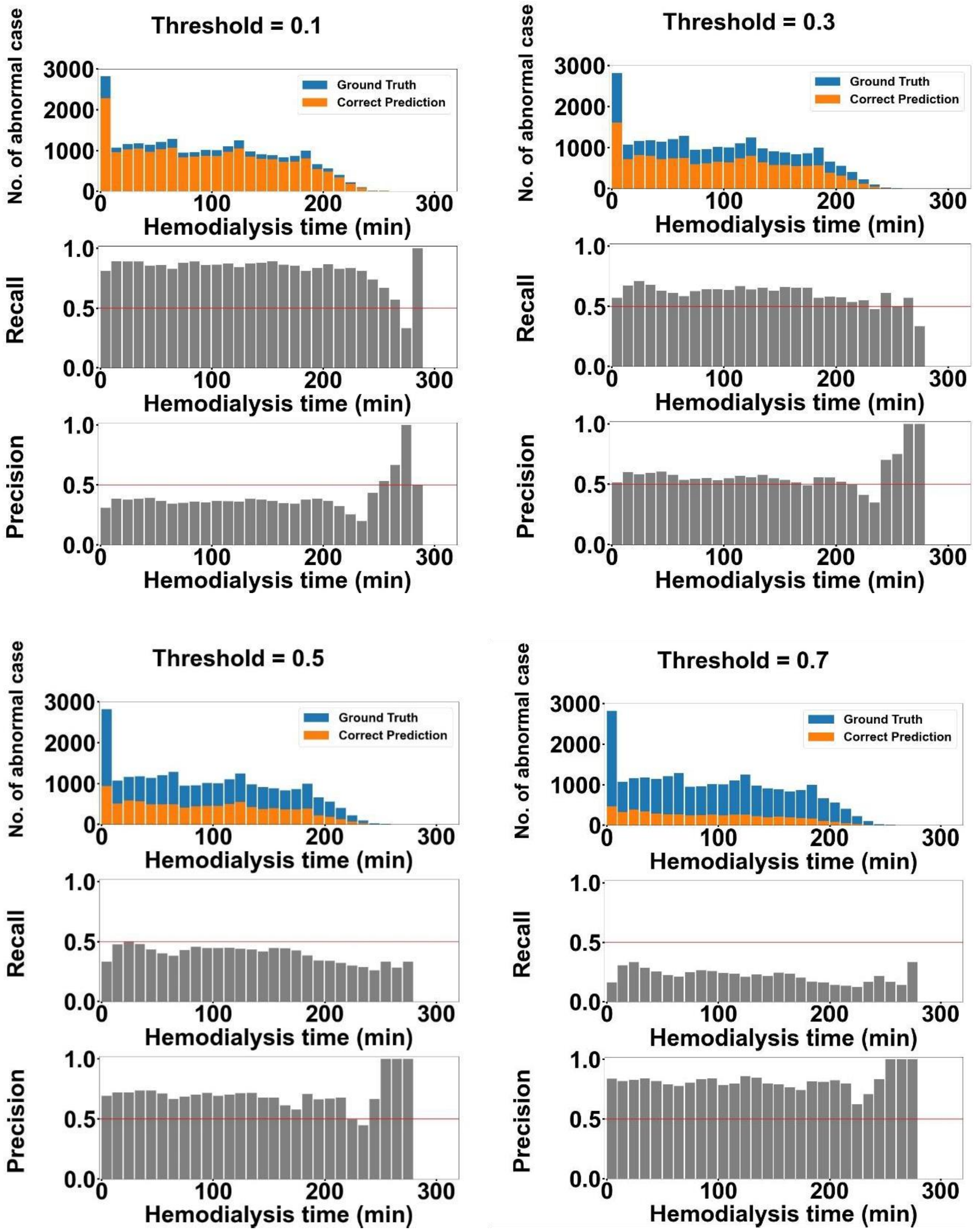
(C)

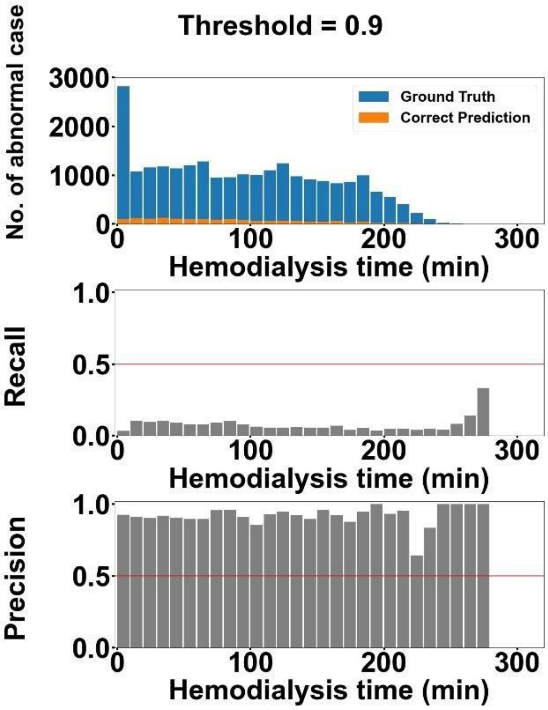


Threshold = 0.1
123481
122593
7663
50983
Threshold = 0.3
213556
32518
28415
30231
Threshold = 0.5
237750
8324
43536
15110
Threshold = 0.7
244467
1607
52999
5647
Threshold = 0.9
246041
33
58163
483
True label
Prediction label
300000
200000
100000
0

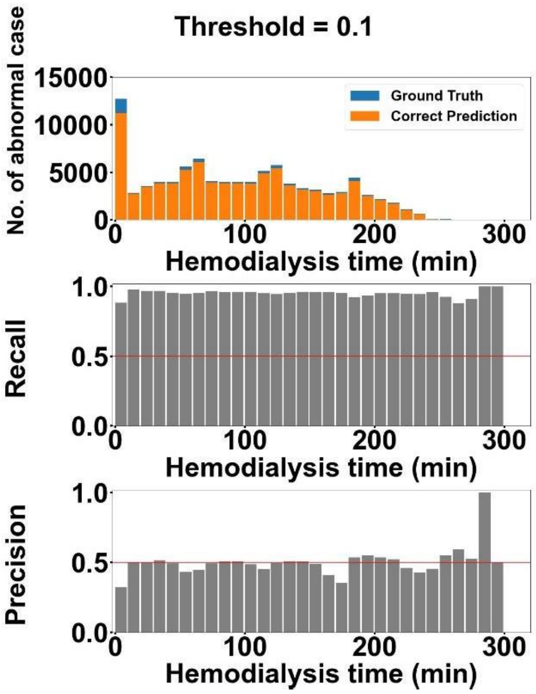
Supplemental Figure 3. Precision-recall graph according to the hemodialysis time. The data output thresholds were set as 0.1, 0.3, 0.5, 0.7, and 0.9. (A) IDH-1. (B) IDH-2. (C) IDH-3.

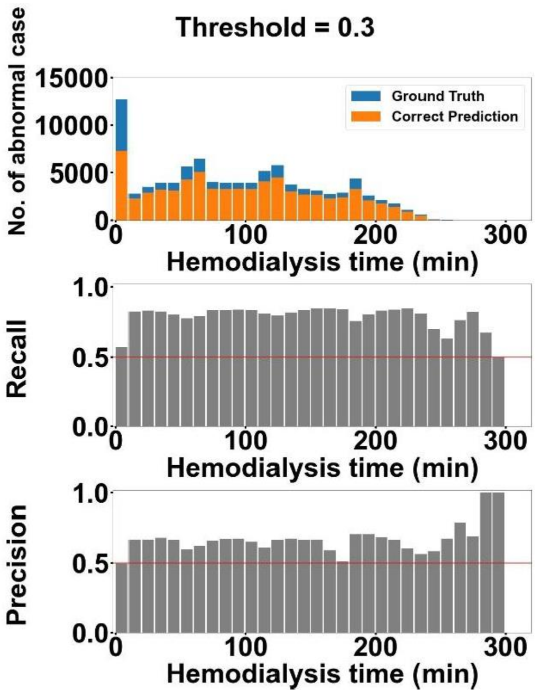
(A)

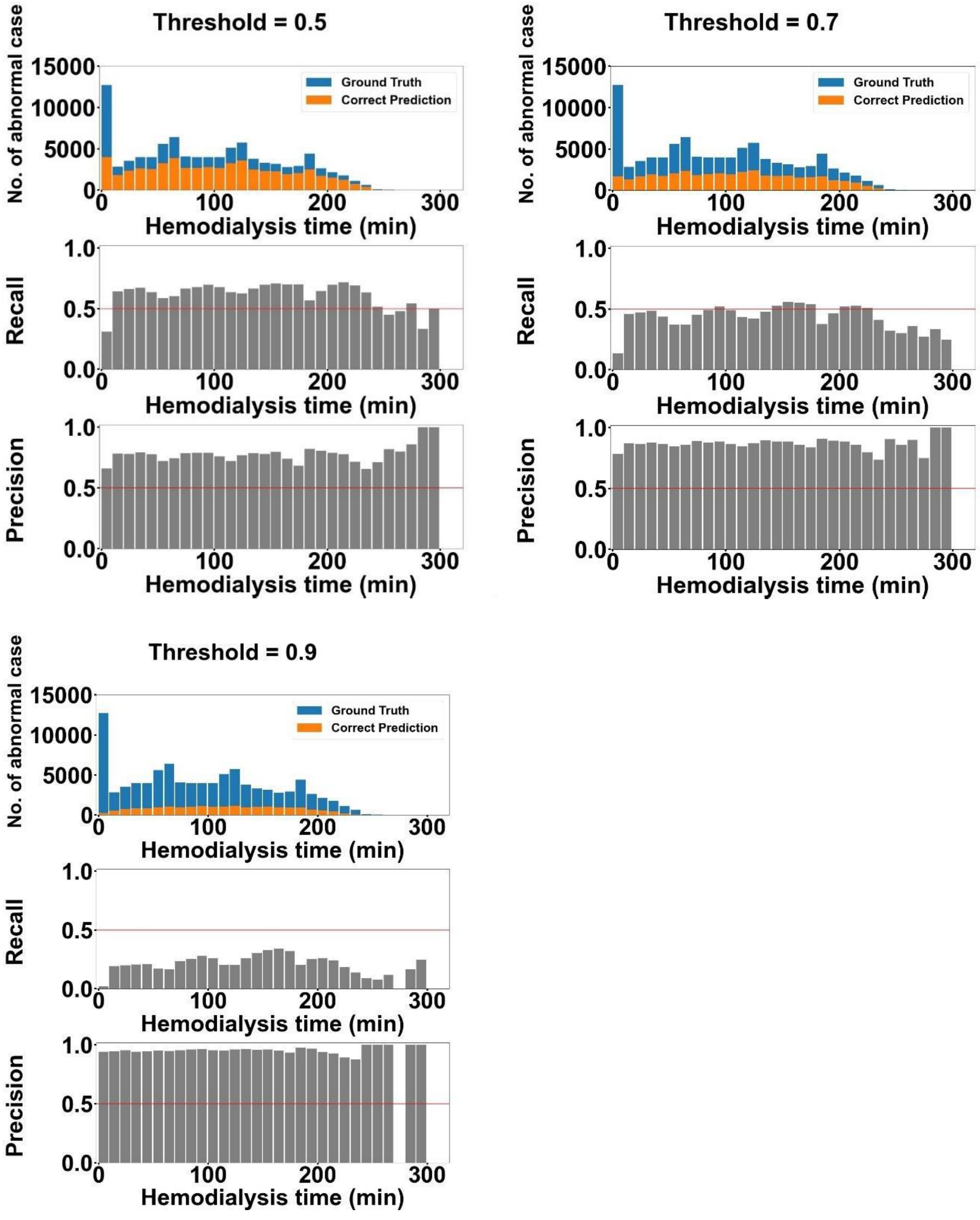



Threshold = 0.9
Ground Truth
Correct Prediction
No. of abnormal case
Recall
Precision
Hemodialysis time (min)

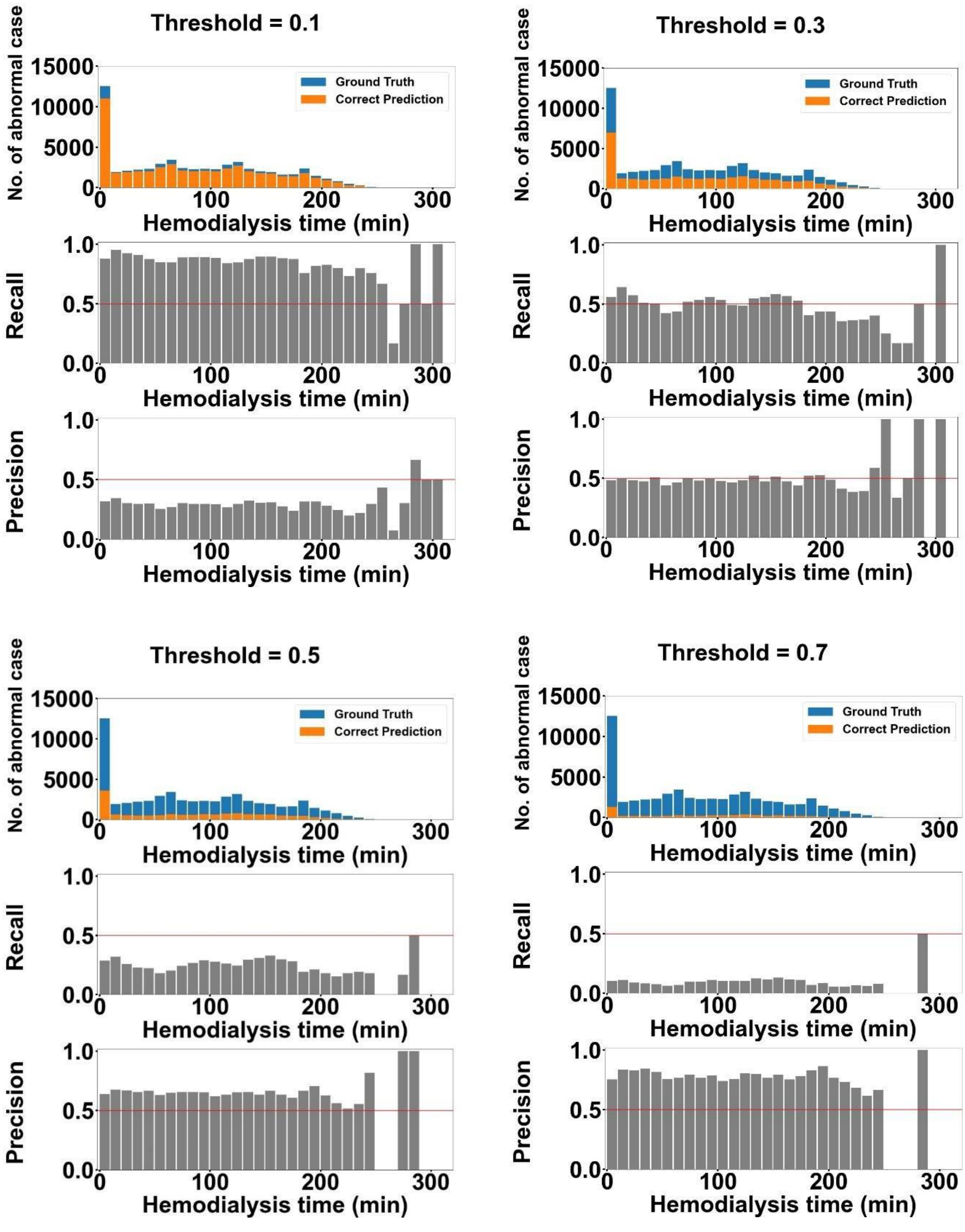
(B)

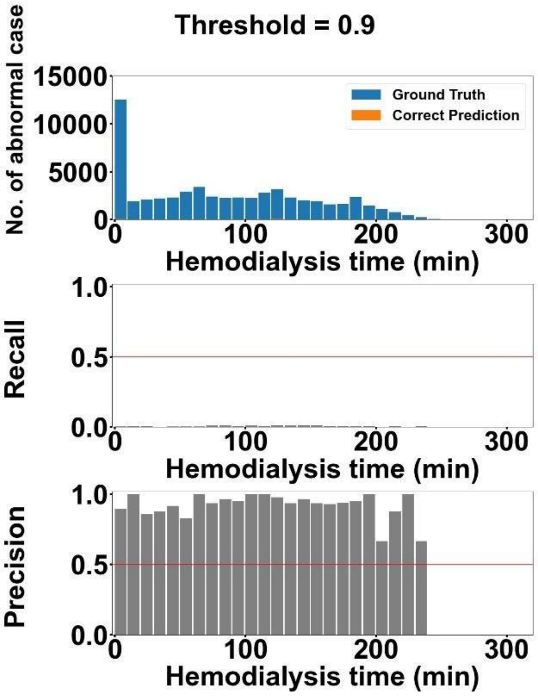

Threshold = 0.1
Ground Truth
Correct Prediction
No. of abnormal case
Recall
Precision
Hemodialysis time (min)


Threshold = 0.3
Ground Truth
Correct Prediction
No. of abnormal case
Recall
Precision
Hemodialysis time (min)


Threshold = 0.5
Threshold = 0.7
Threshold = 0.9
No. of abnormal case
Ground Truth
Correct Prediction
Recall
Precision
Hemodialysis time (min)
15000
10000
5000
0
1.0
0.5
0.0
100
200
300

(C)

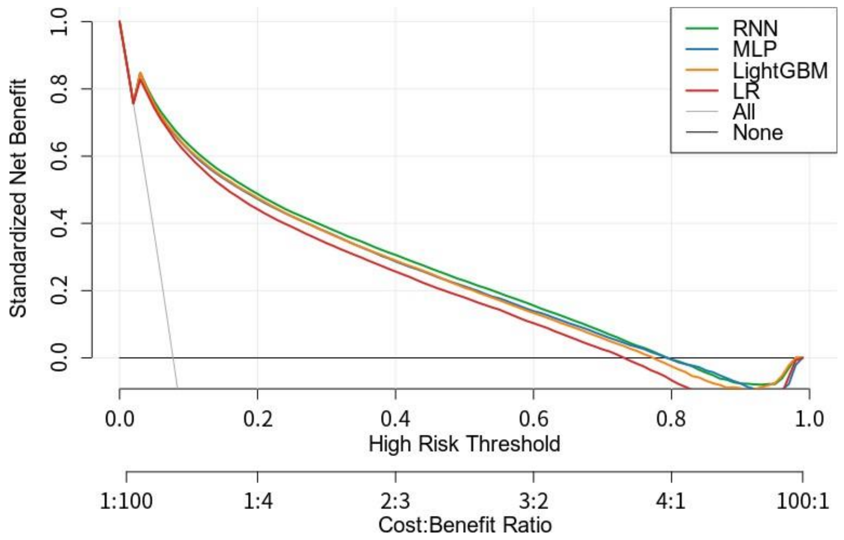

Threshold = 0.1
Threshold = 0.3
Threshold = 0.5
Threshold = 0.7
Ground Truth
Correct Prediction
No. of abnormal case
Recall
Precision
Hemodialysis time (min)
0
5000
10000
15000
0.0
0.5
1.0
100
200
300



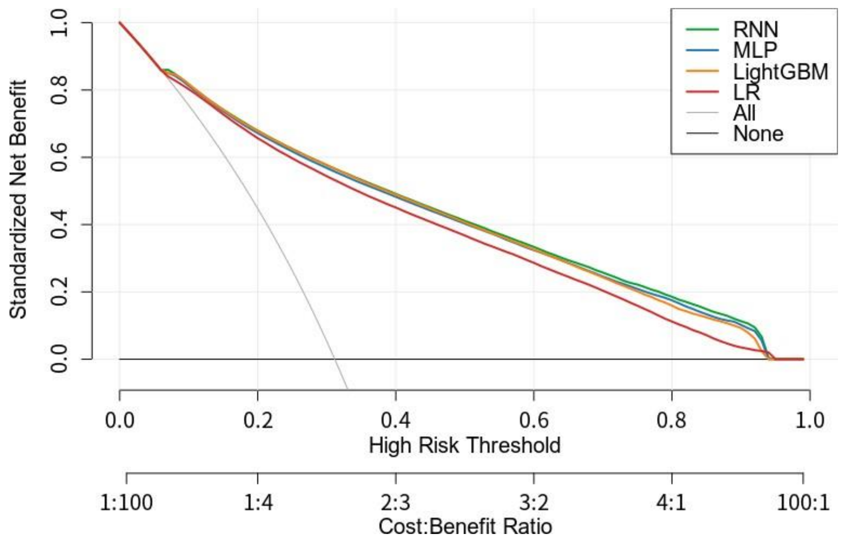
Threshold = 0.9
No. of abnormal case
15000
10000
5000
0
Ground Truth
Correct Prediction
0
100
200
300
Hemodialysis time (min)
1.0
Recall
0.5
0.0
0
100
200
300
Hemodialysis time (min)
1.0
Precision
0.5
0.0
0
100
200
300
Hemodialysis time (min)

Supplemental Figure 4. Decision curve analysis of recurrent neural network (RNN) and three other models. (A) Intradialytic hypotension (IDH)-1. (B) IDH-2. (C) IDH-3. MLP, multilayer perceptron; LightGBM, Light Gradient Boosting Machine; LR, logistic regression.

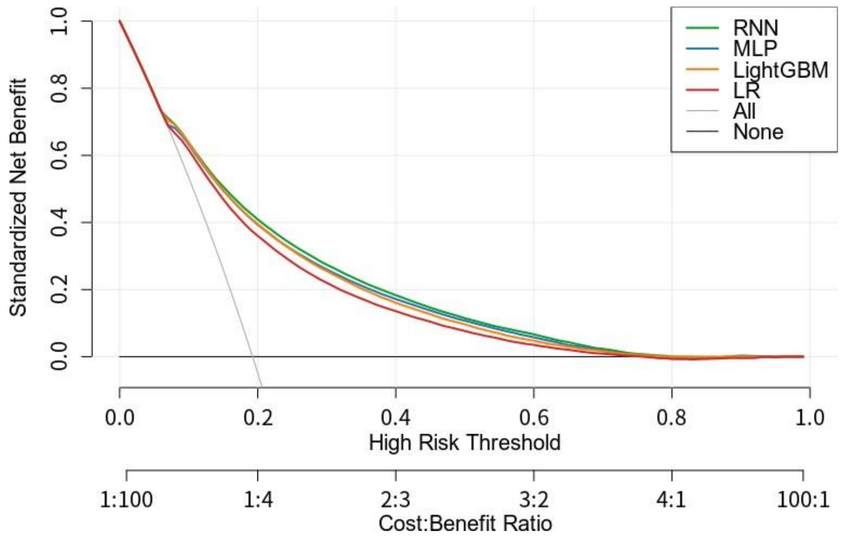
(A)



(B)

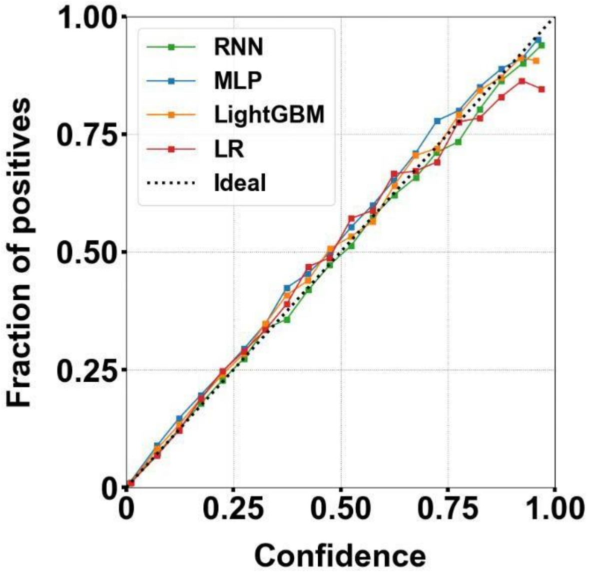


(C)

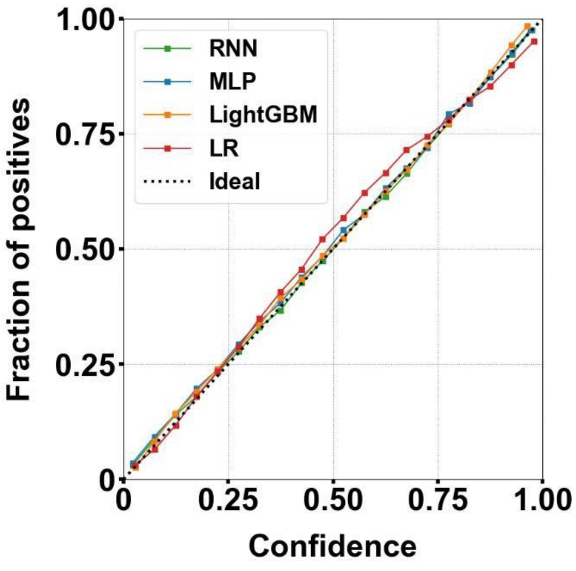

RNN
MLP
LightGBM
LR
All
None
Standardized Net Benefit
0.0
0.2
0.4
0.6
0.8
1.0
High Risk Threshold
1:100
1:4
2:3
3:2
4:1
100:1
Cost:Benefit Ratio

Supplemental Figure 5. Platt scaling plot used to calibrate the models. (A) Intradialytic hypotension (IDH)-1. (B) IDH-2. (C) IDH-3. Bin size = 0.05. RNN, recurrent neural network; MLP, multilayer perceptron; LightGBM, Light Gradient Boosting Machine; LR, logistic regression.

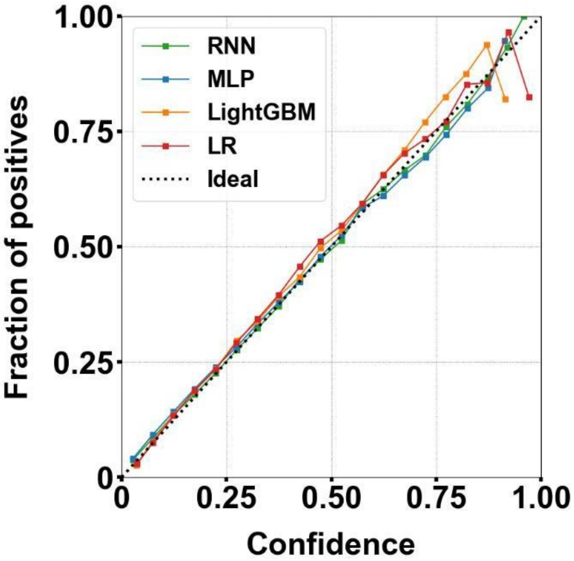
(A)



(B)



(C)



1.00
0.75
0.50
0.25
0
Fraction of positives
0
0.25
0.50
0.75
1.00
Confidence
RNN
MLP
LightGBM
LR
Ideal