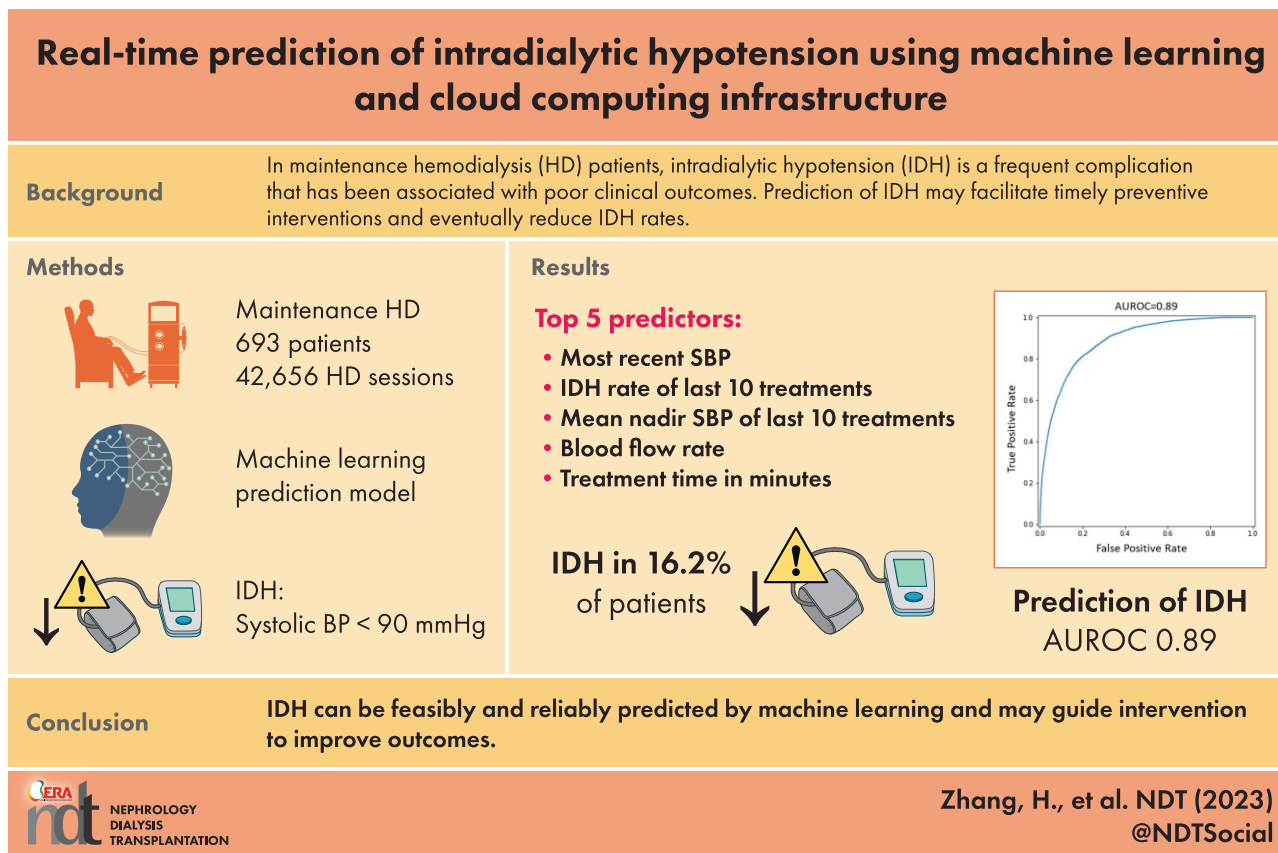# Real-time prediction of intradialytic hypotension using machine learning and cloud computing infrastructure

Hanjie Zhang[1], Lin-Chun Wang[1], Sheetal Chaudhuri[2,3], Aaron Pickering[4], Len Usvyat[2], John Larkin[2], Pete Waguespack[5], Zuwen Kuang[5], Jeroen P. Kooman[3], Franklin W. Maddux[2] and Peter Kotanko[1,6]

[1]Renal Research Institute, New York, NY, USA, [2]Fresenius Medical Care, Global Medical Office, Waltham, MA, USA, [3]Maastricht University Medical Center, Maastricht, The Netherlands, [4]Fresenius Medical Care, Data Solutions, Berlin, Germany, [5]Fresenius Medical Care, Digital Technology & Innovation, Waltham, MA, USA and [6]Icahn School of Medicine at Mount Sinai, New York, NY, USA

Correspondence to: Hanjie Zhang; E-mail: Hanjie.Zhang@RRINY.COM

GRAPHICAL ABSTRACT

Real-time prediction of intradialytic hypotension using machine learning and cloud computing infrastructure

Background
In maintenance hemodialysis (HD) patients, intradialytic hypotension (IDH) is a frequent complication that has been associated with poor clinical outcomes. Prediction of IDH may facilitate timely preventive interventions and eventually reduce IDH rates.

Methods
Maintenance HD
693 patients
42,656 HD sessions

Machine learning prediction model

IDH:
Systolic BP < 90 mmHg

Results
Top 5 predictors:
- **Most recent SBP**
- **IDH rate of last 10 treatments**
- **Mean nadir SBP of last 10 treatments**
- **Blood flow rate**
- **Treatment time in minutes**

IDH in 16.2% of patients

Prediction of IDH
AUROC 0.89

Conclusion
IDH can be feasibly and reliably predicted by machine learning and may guide intervention to improve outcomes.

Zhang, H., et al. NDT (2023)
@NDTSocial

ABSTRACT

Background. In maintenance hemodialysis patients, intradialytic hypotension (IDH) is a frequent complication that has been associated with poor clinical outcomes. Prediction of IDH may facilitate timely interventions and eventually reduce IDH rates.

Methods. We developed a machine learning model to predict IDH in in-center hemodialysis patients 15–75 min in advance. IDH was defined as systolic blood pressure (SBP) <90 mmHg. Demographic, clinical, treatment-related and laboratory data were retrieved from electronic health records and merged with intradialytic machine data that were sent in real-time to the

KEY LEARNING POINTS

What is already known about this subject?

- Intradialytic hypotension (IDH) occurs in over 10% of hemodialysis treatments.
- Timely interventions may prevent IDH.

What this study adds?

- In this manuscript, we developed a machine learning model for real-time (i.e. intradialytic) prediction of IDH, utilizing electronic health records comprising intradialytic blood pressure measurements and multiple treatment- and patient-level variables from 42 656 hemodialysis sessions in 693 in-center hemodialysis patients.
- In the training cohort, the model was optimized to generate an alert between 15 and 75 min before an IDH event. In the validation cohort, the model achieved an area under the receiver operating characteristic curve of 0.89.

What impact this may have on practice or policy?

- Based on this performance, the model could alert clinicians and trigger the timely deployment of interventions to prevent IDH.

cloud. For model development, dialysis sessions were randomly split into training (80%) and testing (20%) sets. The area under the receiver operating characteristic curve (AUROC) was used as a measure of the model's predictive performance.

Results. We utilized data from 693 patients who contributed 42 656 hemodialysis sessions and 355 693 intradialytic SBP measurements. IDH occurred in 16.2% of hemodialysis treatments. Our model predicted IDH 15–75 min in advance with an AUROC of 0.89. Top IDH predictors were the most recent intradialytic SBP and IDH rate, as well as mean nadir SBP of the previous 10 dialysis sessions.

Conclusions. Real-time prediction of IDH during an ongoing hemodialysis session is feasible and has a clinically actionable predictive performance. If and to what degree this predictive information facilitates the timely deployment of preventive interventions and translates into lower IDH rates and improved patient outcomes warrants prospective studies.

Keywords: end-stage kidney disease, intradialytic hypotension, machine learning, real-time prediction

INTRODUCTION

Intradialytic hypotension (IDH) is a common complication of hemodialysis (HD), occurring in 8%–40% of HD sessions [1–4]. IDH reduces quality of life and is associated with morbidity and mortality [1, 4–8]. The management of IDH requires staff interventions, such as patient evaluation, monitoring and adjustment of ultrafiltration rates. It is reasonable to assume that prediction of IDH followed by appropriate preventive measures may lower IDH rates.

Risk factors for IDH are diabetes, cardiovascular disease, autonomic dysfunction and high interdialytic weight gain, to name a few [4, 9, 10]. Because these risk factors are prevalent in many HD patients, IDH prediction is challenging.

Artificial intelligence (AI) and machine learning (ML), and cloud infrastructures enable the evaluation of large data volumes in a secure, reliable and efficient way. This affords an opportunity to develop and leverage models for real-time prediction of IDH among individual patients while undergoing HD. AI and ML have been used to develop models to predict IDH [11–14]. One model developed to predict IDH [defined as systolic blood pressure (SBP) <90 mmHg] based on a relatively limited number of parameters (e.g. body and dialysate temperature, ultrafiltration rate) predicted IDH with a sensitivity of 86% and a specificity of 81% [13]. It is expected that real-time models that consider a larger number of factors may result in improved IDH predictive ability.

Robust data are routinely captured before and during HD, including patient characteristics, clinical status, HD treatment parameters and intradialytic vital signs. These variables can be considered for cloud computing, advanced modeling and near real-time reporting. In this quality improvement project (QIP), we retrospectively assessed data from several HD clinics that used a secure cloud-based computer infrastructure to develop and deploy an ML prediction model that can predict IDH prior to an IDH event in real time. Clinically, prediction is only useful if it predicts an IDH event for a given patient during an ongoing dialysis treatment.

MATERIALS AND METHODS

Participants

We conducted a *post hoc* analysis of data from a convenience sample comprising patients who received in-center HD in six US Fresenius Kidney Care clinics between January 2020 and November 2020. Four clinics are located in Waltham, MA, and two clinics in New York City, NY. These clinics transfer data from dialysis machines to a secure Internet of Things (IoT) private server via Amazon Web Services (AWS) (Amazon Web Services, Inc., Seattle, WA, USA) using IoT software [15–17]. In total, data from 42 656 dialysis treatments in 693 patients were included. This project was conducted as an internal QIP and therefore was not submitted for Institutional Review Board review.

Data characteristics and cloud architecture

IDH was defined as any intradialytic SBP <90 mm Hg [18]. The outcome (i.e. dependent variable) was a binary classification of patients at risk for IDH within 15–75 min

of the given prediction time point. Variables included in the predictive model comprised demographic, clinical and laboratory data obtained from electronic health records and dialysis machine data. Dialysis machine data included variables such as sitting SBP and diastolic blood pressure (DBP), pulse, blood flow rate, dialysate flow rate, ultrafiltration rate, ultrafiltration volume removed and time on dialysis. Additional variables were derived from these data, including changes in SBP, DBP and heart rate during HD, DBP and SBP nadirs over the last 10 treatments, and IDH event rates over the last 10 treatments. Although some of these variables were assessed as frequently as every 10 s (e.g. blood flow rate, dialysate flow rate, ultrafiltration rate, ultrafiltration volume removed, time on dialysis), data points that aligned with SBP measurements were used for analysis. Therefore, whenever intradialytic SBP was measured (approximately every 30 min), a patient's IDH risk within the next 15–75 min was computed in real time. The prediction process took less than a minute.

Patients' demographics were extracted from electronic health records. Prescribed treatment time, vascular access type, pre- and post-dialysis weight, and the dialysate sodium concentration were also evaluated. Laboratory variables were captured monthly; hemoglobin was measured weekly. In total, 99 variables were evaluated in the present model (Supplementary data, Table S1).

Supplementary data, Fig. S1 shows the cloud-based infrastructure used for this QIP. Dialysis machine data were sent in real-time to the cloud-based platform. The data stream was enriched with patient and historical treatment data, flowed through streaming analytical processes, and the data were segmented into windows of time. Features were calculated for each time window and passed to the ML model. The model results were then visualized on a dashboard, plotting intradialytic SBP and predicted probability of IDH over treatment time. This dashboard could be accessed by the clinical staff and provide insights into patients who are predicted to experience IDH during the ongoing treatment. All streamed machine data and model results could be stored and monitored.

Model design and development

Model design, development and evaluation were substantially guided by inputs from healthcare professionals (HCP) who are experienced in the care of HD patients. HCP requested an actionable 15–75 min prediction time window preceding IDH. Data from the 15 min before an IDH were not used, because an alert within such a short timeframe would not provide sufficient lead time for HCP to respond to a predicted IDH event. Post-IDH data were not considered by the model. HCP indicated that a 10% rate of false IDH alarms (=false positive rate) is clinically and operationally acceptable.

The model was built using the AWS SageMaker development platform and was trained and tested on data collected in the dialysis clinics [19].

To develop the model, we created training (80%) and test (20%) datasets in two ways, depending on the unit of randomization. In the session randomization design, we randomly assigned HD sessions to the respective training and test sets. In the patient randomization design, we randomly assigned patients to training and test sets.

We used Python version 3.7.7 (Python Software Foundation, Fredericksburg, VA, USA) to build the ML model using the XGBoost (Extreme Gradient Boosting) package [20]. Model-specific internal parameters were set to predetermined default values to cover general use cases and were tuned for the specific task to obtain optimal performance using a Bayesian strategy [21]. The number of weak learners was 200, the learning rate was 0.05 and the maximal depth was 7 for the final model. The XGBoost used input variables from the training dataset to construct multiple decision trees, giving each a random sample, and established a series of thresholds that split variables to maximize the information gain. Decision trees were constructed iteratively, and new decision trees were added to predict prior errors. XGBoost decision trees handle missing values without imputation, by including their presence when determining the splits. To show the integrity of the dataset, the percentage of missing values for each variable was summarized in Supplementary data, Table S2.
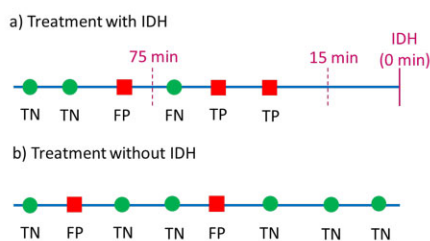
Evaluation of model performance

Evaluation of model performance was based on the definitions of true positive (TP), false positive (FP), true negative (TN) and false negative (FN) as defined in Fig. 1. The performance analysis comprised two aspects, one per SBP measurement (usually multiple SBP measurements per HD session) and one that considered the entire HD session. This approach allowed us to look at the model performance from different angles. For both analyses we assessed sensitivity (also termed recall) as [TP/(TP+FN)], and specificity as [TN/(TN+FP)]. For the model performance based on SBP measurements, we also calculated the area under the receiver operating characteristics curve (AUROC). To that end, the rates of TP and FP were computed by the prediction model across the entire IDH probability spectrum (i.e. zero to one). Selected IDH prediction thresholds are presented in the Results section. We considered an AUROC of at least 0.85 as clinically meaningful. We also calculated the area under the precision-recall curve (AUPRC). Precision (also called positive predictive value) is defined as TP/(TP + FP).

We also conducted a patient-level analysis of the model's performance. For the respective analysis of model sensitivity, only patients with IDH events were included, because only they contribute TP and FN predictions. Conversely, for the analysis of model specificity, patients who experienced IDH during all treatments had to be excluded because they did not contribute TN and FP results.

In addition to these performance indices, we constructed calibration plots. To that end, we plotted the probability of IDH as predicted by the model versus the true IDH event rate. To increase insights from the calibration plots, we used two binning methods, one with 10 equal bin widths and one with equal sample size per bin (200 bins).

Shapley values were calculated using the SHAP Python package (Python Software Foundation) to determine a
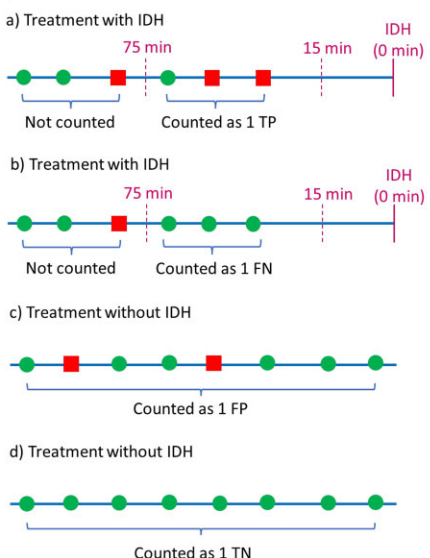
Analysis per SBP measurement

a) Treatment with IDH

75 min — 15 min — IDH (0 min)

TN TN FP FN TP TP

	TP	TN	FP	FN
Count	2	2	1	1

b) Treatment without IDH

TN FP TN TN FP TN TN TN

	TP	TN	FP	FN
Count	0	6	2	0

Analysis per HD session

a) Treatment with IDH

75 min — 15 min — IDH (0 min)

Not counted | Counted as 1 TP

	TP	TN	FP	FN
Count	1	0	0	0

b) Treatment with IDH

75 min — 15 min — IDH (0 min)

Not counted | Counted as 1 FN

	TP	TN	FP	FN
Count	0	0	0	1

c) Treatment without IDH

Counted as 1 FP

	TP	TN	FP	FN
Count	0	0	1	0

d) Treatment without IDH

Counted as 1 TN

	TP	TN	FP	FN
Count	0	1	0	0

■ Predicted Probability ≥ Threshold
● Predicted Probability < Threshold

Figure 1: Evaluation of model performance. The performance analysis comprised two aspects, one per SBP measurement (usually multiple SBP measurements per HD session) and one that considered the entire HD session. Evaluation of model performance was based on the definitions of TP, FP, TN and FN.

particular feature's contribution to the outcome [22, 23]. The SHAP (SHapley Additive exPlanations) value method assigns each feature an importance value for a particular prediction. Shapley values are calculated for each variable and each observation, representing a measure of effect (positive or negative value) of the observed value on each individual prediction. SHAP methods withhold and include individual inputs in all possible combinations, and compare differences between withheld and included data, to compute the mean value of all possible differences for attributing the feature importance. SHAP values are output as log-odds, meaning they are additive explanations of feature importance. Therefore, positive SHAP values increase the predicted probability of IDH, whereas negative SHAP values decrease the probability of IDH. We calculated the overall feature importance for individual variables in the model with the SHAP value method using the mean absolute values for each variable across all observations.

RESULTS

Patient characteristics

We studied 693 in-center HD patients during 42 656 dialysis sessions. Their mean age was 60.5 ± 16.1 years. Most patients were males (58.2%), 47.8% were white and 40.1% had diabetes (Table 1). Of the 693 patients, approximately 80% experienced at least one IDH episode (Supplementary data, Fig. S2). About 16.2% of all HD sessions were complicated by IDH. A total of 355 693 intradialytic SBP were recorded. About 4% of all SBP measurements fell within the window of 75–15 min before an IDH event (Table 2).

Model development, performance and top features

All variables used in the prediction model and percentage of missing values for each variable are shown in Supplementary data, Tables S1 and S2, respectively.

Session randomization design

Here we randomized 42 656 HD sessions to training (34 124 sessions) and test (8532 sessions) sets. Supplementary data, Table S3 shows descriptive characteristics of selected model input variables for the entire patient population (N = 42 656), the training (N = 34 124) and the test datasets (N = 8532), respectively. In the test sample, an AUROC of 0.887 [95% confidence interval (CI) 0.881–0.892] was achieved (Fig. 2A). The AUPRC was 0.334 (Supplementary data, Fig. S3).

The performance analysis per SBP measurement is shown in Table 3. Based on discussions with HCP, a false positive rate of 10% was deemed acceptable. Therefore, we choose an IHD probability of ≥0.09 as the threshold for classifying

Table 1: Patient baseline demographics (N = 693 patients).

Variable	
Age, years [mean (SD)]	60.5 (16.1)
Dialysis vintage, years [mean (SD)]	3.8 (4.6)
Male [N (%)]	403 (58.2)
Race [N (%)]	
Black	140 (20.2)
White	331 (47.8)
Other	222 (32.0)
Hispanic [N (%)]	101 (15.9)
Comorbidities [N (%)]	
Chronic obstructive pulmonary disease	44 (6.3)
Diabetes	278 (40.1)
HIV/AIDS	27 (3.9)
Hyperparathyroidism	648 (93.5)
Myocardial infarction	34 (4.9)
Peripheral vascular disease	38 (5.5)
Ischaemic heart disease	4 (0.5)
Hepatitis	55 (7.9)
Vascular access type (%)	
Arteriovenous fistula	57.5
Arteriovenous graft	10.7
Catheter	31.8

SD, standard deviation; HIV/AIDS, human immunodeficiency virus/acquired immunodeficiency syndrome.

Table 2: Characteristics of HD sessions with respect to IDH and BP.

Variable	All	Training data	Test data
HD sessions [N (% of total)]	42 656 (100)	34 124 (80)	8532 (20)
HD sessions with IDH [N (% of HD sessions)]	6922 (16.2)	5548 (16.3)	1374 (16.1)
BP measurements [N (% of total)]	355 693 (100)	284 268 (79.9)	71 425 (20.1)
BP measurements in the window of 75–15 min before IDH [N (% of BP measurements)]	14 452 (4.06)	11 544 (4.06)	2908 (4.07)

predictions. With this probability threshold, a sensitivity and specificity of 0.65 and 0.90, respectively, were attained (Table 3). Model performance was also assessed considering the entire HD session. The resulting sensitivities and specificities with respect to three selected IDH probability thresholds (0.09; 0.15; 0.2) are shown in Table 4. Using a threshold of ≥0.09, sensitivity and specificity were 0.734 and 0.780, respectively.

We also assessed the model performance at the patient level. For the patient-level sensitivity analysis, we had to exclude 288 patients (44%) from the test dataset because these patients did not experience any IDH and thus contributed neither TP nor FN results. The mean patient-level sensitivity in the remaining 360 patients was 0.48. Only 2 of the 648 patients in the test dataset experienced IDH during every session. These two patients were not considered for the computation of specificity, as they contributed—by definition—neither TN nor FP results. The mean specificity calculated from the remainder of 646 patients is 0.86. In addition, a subgroup analysis in the 288 patients without any IDH resulted in a mean specificity of 0.96.
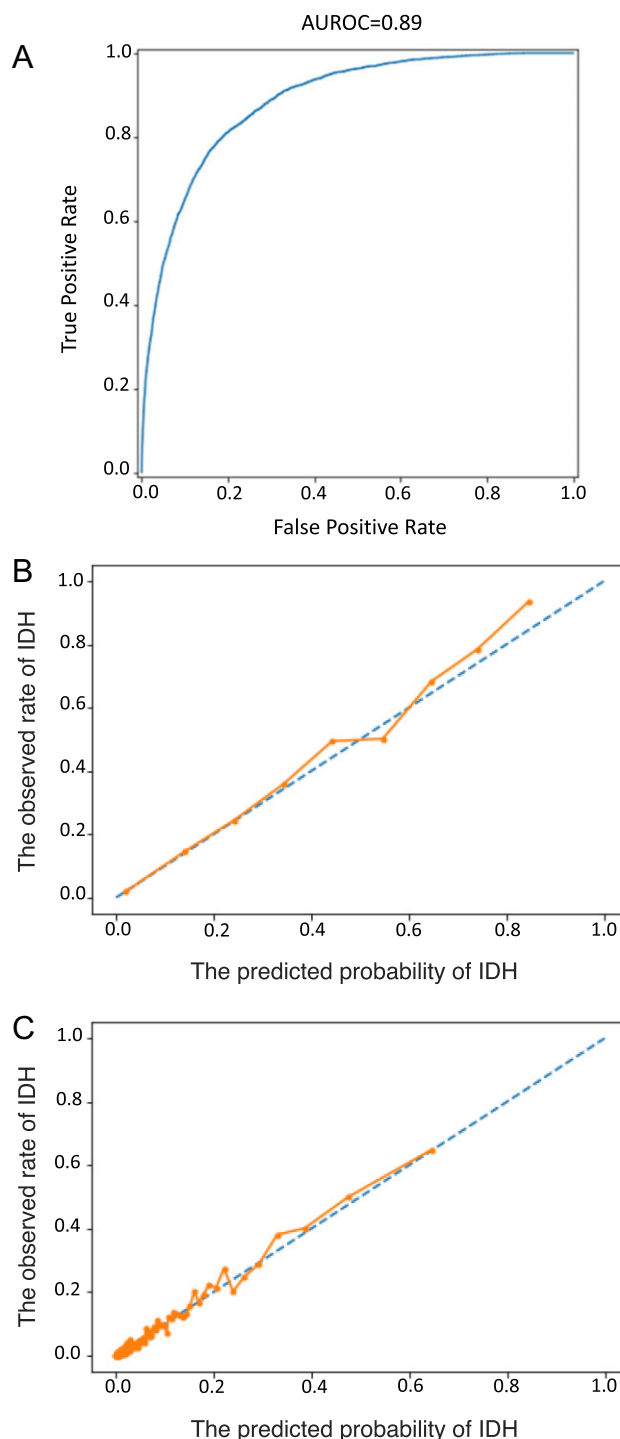

Figure 2: (**A**) AUROC. The true positive rate is shown on the y-axis, equal to sensitivity. The false positive rate on the x-axis is calculated as 1 – specificity. The 95% CIs of the AUROC are 0.881–0.892. (**B**) Calibration plot with equal bin width. Probability calibration plot shows the predicted probability against observed events. (**C**) Calibration plot with equal number of samples per bin (200 bins in total). Here, instead of equal bin-width, we set the bin width based on the number of samples to account for the data distribution (most probabilities are under 0.5).
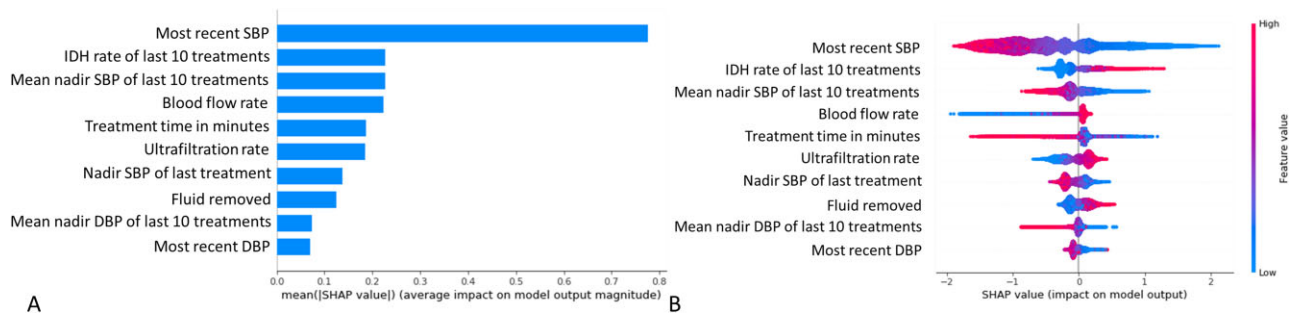
Figure 3: Top 10 predictors for IDH in the ML model. (**A**) Mean absolute SHAP values. (**B**) The SHAP summary plot shows the degree of each measurement's positive or negative effect on the prediction (x-axis). Warmer colors represent higher observed values for that measurement; cooler colors indicate lower values. For example, the higher (warmer color) the "most recent SBP," the more the negative impact it has on the model (less chance of IDH).

Table 3: Sensitivity, specificity and precision as a function of IDH probability thresholds.

IDH probability threshold	Sensitivity	Specificity	Precision
0.00	1.00	0.00	0.05
0.05	0.79	0.82	0.16
0.09	0.65	0.90	0.22
0.1	0.62	0.91	0.23
0.15	0.50	0.95	0.30
0.2	0.38	0.97	0.35
0.3	0.25	0.99	0.47
0.4	0.16	0.99	0.56
0.5	0.10	1.00	0.62
0.6	0.06	1.00	0.73
0.7	0.02	1.00	0.81
0.8	0.01	1.00	0.94

This analysis considered individual SBP measurements.

Table 4: Descriptive statistics and sensitivity, specificity and precision as a function of IDH probability thresholds.

Descriptives		IDH probability thresholds		
	N	0.09	0.15	0.20
Patients	648			
HD sessions	8532			
HD sessions with IDH	1374			
HD sessions without IDH	7158			
Correctly predicted sessions with IDH (TP)		1009	797	646
Falsely predicted sessions with IDH (FP)		1574	901	550
Correctly predicted sessions without IDH (TN)		5584	6257	6608
Falsely predicted sessions without IDH (FN)		365	577	728
Sensitivity		0.734	0.580	0.470
Specificity		0.780	0.874	0.923
Precision		0.391	0.469	0.540

This analysis considered entire HD sessions.

Figure 2B and C shows calibration plots. Our model is reliable and very likely to return the true probability of an IDH event. The Shapley values (Fig. 3) indicate that the input variables most predictive of IDH within the next 15–75 min were the "most recent SBP," "IDH rate during the last 10 treatments," "mean nadir SBP during the last 10 treatments,"

"blood flow rate," "treatment time," "ultrafiltration rate," "lowest SBP during the previous treatment," "fluid removed," "mean nadir DBP during the last 10 treatments," and "most recent DBP".

Patient randomization design

Here we randomized patients to training (554 patients; 33 816 sessions) and test (139 patients; 8840 sessions) sets. The model performance was materially identical to the session randomization design [AUROC of 0.876 (95% CI 0.870–0.882)] (Supplementary data, Fig. S4).

Prediction examples

Figure 4A depicts an example dashboard for a patient having IDH, with an SBP of 70 mmHg, at 100 min into treatment. This was the fifth assessment of BP during the session. The dashboard displays a low prediction probability for the first two SBP measurements, then a higher-than-threshold prediction probability at both SBP measurements before the occurrence of IDH.

Figure 4B shows a patient with stable SBP and without IDH. The prediction probability associated was well below the threshold of ≥0.09.

DISCUSSION

Predicting IDH in HD patients is challenging due to the numerous patient- and treatment-related factors that affect IDH risk. Clinically, prediction is only useful if it predicts an IDH event for a given patient during an ongoing dialysis treatment. Our analysis shows that IDH can be predicted using real-time dialysis data and ML algorithms during an ongoing dialysis session. Our model was developed and validated on 42 656 HD sessions from a large and diverse patient cohort. It achieved AUROCs of 0.887 (95% CI 0.881–0.892) (session randomization design) and 0.876 (95% CI 0.870–0.882) (patient randomization design), both numbers are significantly greater than 0.85. The model's high performance in the testing dataset suggests that the tool may potentially assist clinicians in intervening proactively and in real-time in patients at risk for IDH.

Although previous studies have evaluated the ability of ML models to predict IDH, most did not consider the impact of time-varying factors, such as blood pressure (BP)
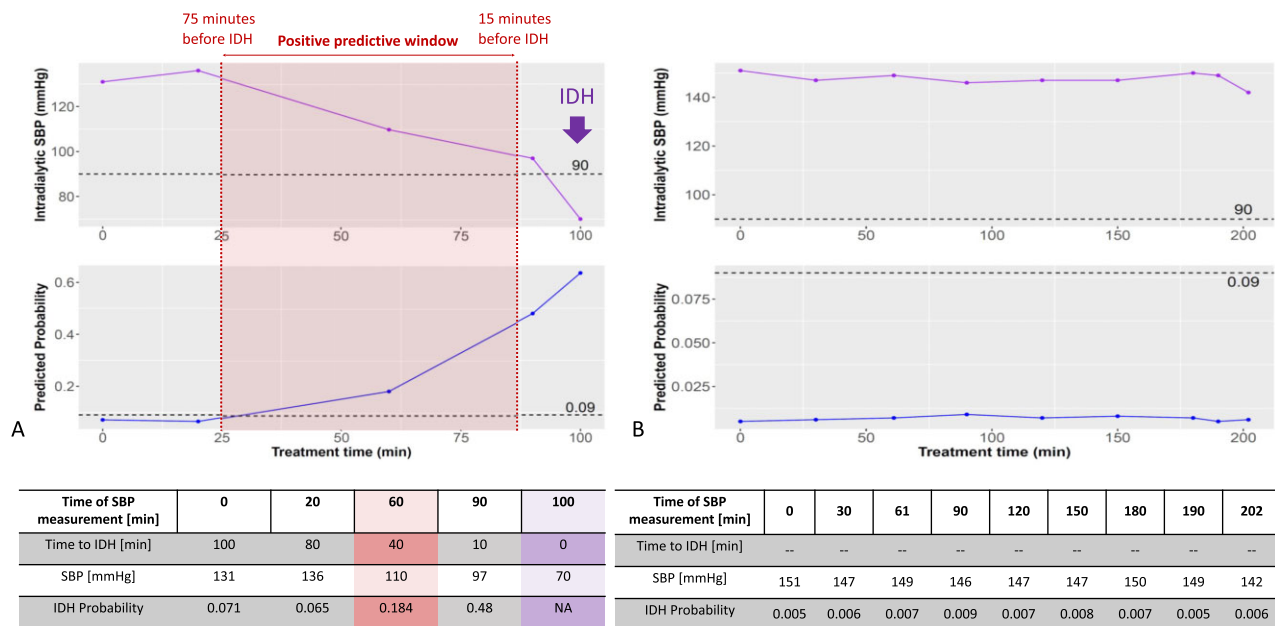
Figure 4: Two examples demonstrating a patient with (**A**; left panel) and without (**B**; right panel) IDH. (A) An example of SBP (top row) and IDH probability (bottom row) in a patient who experienced IDH. The model predicted the probability of IDH throughout HD with each SBP measurement (approximately every 30 min during a regular HD treatment); the model was trained to predict IDH 15–75 min before the event. An IDH probability ≥0.09 was set as IDH alert threshold (dashed horizonal lines in the lower panels). Data after the IDH event were not considered. (B) An example of SBP and predicted IDH probability in a patient who did not experience IDH; the IDH probability remained below the IDH alert threshold throughout the treatment.

[11–14]. For example, Barbieri *et al.* developed an AI model to guide the management of BP, fluid volume and dialysis dose in chronic HD patients [11]. Their multiple endpoint model predicted Kt/V, fluid volume removal, BP and heart rate, based on patient characteristics, prior hemodynamic responses and dialysis prescriptions. While the model predicted post-dialysis weight and Kt/V accurately, the authors reported that modeling minimum SBP was more difficult, possibly due to measurement errors by machine sensors or to the natural variations in the physiology of hemodynamic reactions.

A more recent deep learning (recurrent neural network) model developed by Lee *et al.* to predict the real-time risk of IDH using a timestamp-bearing dataset appeared to perform better than previous models [12]. The model predicted IDH, defined as intradialytic SBP <90 mmHg within 1 h, with an AUROC of 0.94 (95% CI 0.94–0.94). According to the authors, the improved performance of their model could be attributed to the fact that it considered the impact of patients' vital signs and HD settings, which vary over time and are known to affect the risk of IDH [12, 24, 25]. In addition, the results of the analysis' feature set–ablation analysis and feature ranking analysis, both of which confirmed the value of considering real-time changes in vital signs and HD settings, supported the inclusion of these data in the IDH prediction model [12].

For any real-time prediction, a stable and sophisticated IT infrastructure is essential. A proof-of-concept analysis of real-time data on intradialytic relative blood volume (RBV) using ML confirmed that a cloud-based framework can be used to predict intradialytic RBV changes [26].

The XGBoost model is a robust choice well suited to general ML problems. The model is typically excellent at identifying and modeling complex feature interactions and performs well on tabular data. Although the given model is very accurate, experimentation with other algorithms that natively model time lag dependencies (e.g. recurrent neural network, long short-term memory networks, gated recurrent units) may provide some benefit in the future. In addition, although modeling as a binary classifier simplifies the problem, we may also consider predicting IDH risk directly. In the present model, whereas SBP during the current and previous treatments had the highest SHAP values, the relative importance of blood flow rate as an important contributor to the model was somewhat unexpected. Recent literature showed no relation between blood flow rate and hemodynamics during HD [27, 28]. In a sub-analysis of the Hemodialysis (HEMO) study, an increased incidence of IDH was observed among patients randomized to higher Kt/V targets resulting from increased blood flow, dialysate flow, dialyzer surface area and, if needed, session length [29]. The authors of this study hypothesized that rapid shifts in osmolality explain this observation. Whereas it should be recognized that ML models cannot provide mechanistic explanations for observed associations, the results of our program might provide rationale for future intervention studies assessing the effects of lowering blood flow rate in patients with impending IDH.

The true value of IDH prediction tools depends on the ability of clinicians to promptly intervene to prevent IDH. After evaluating our results, we had extensive discussions with clinical practitioners (i.e. nephrologists and nursing staff) who

indicated that a 10% false positive rate (90% specificity) is an operationally reasonable number while still having a decent sensitivity of 65%. The probability threshold was set to ≥ 0.09 to avoid too many false alarms. Our prediction tool records data in real time and sends them to the cloud, where the ML algorithm computes the probability of IDH. Based on the IDH risk, alerts are generated and reported to the dashboard. Our system and other similar tools that may provide real-time prediction represent a valuable advance in the preemptive management of IDH. Preventive interventions include, e.g. adjustments of the dialysis prescription (such as lowering ultrafiltration rate), the patient's position and fluid administration [6].

While predicting IDH during an ongoing HD session is highly relevant clinically, there is also interest in analyzing the model's performance on a patient level. For the patient-level analysis of model sensitivity, note that we had to exclude 288 patients (44%) because they did not experience an IDH and thus—by definition—contributed neither TP nor FN results. To better understand and quantitate the model sensitivity on a patient level, a larger population needs to be studied. This would also inform an important aspect of the model's patient-level usefulness in clinical practice. In contrast, for the patient-level specificity analysis only the two patients (0.3%) who experienced IDH during all sessions had to be excluded, so the patient number does not impact this analysis.

One limitation of our approach is the *post hoc* analysis of routinely collected data. Also, we did not calculate a specific sample size for our research; instead, we used a large convenience sample. An additional limitation relates to general drawbacks of ML models, including the fact that it is difficult to understand how the model arrives at its prediction or how any individual factor influences its results [12]. Because the model was created based on a specific US dataset, its design may need to be adjusted to populations with different demographic, clinical or dialysis treatment characteristics. Another shortcoming is the lack of oral medication information in our dataset.

In conclusion, cloud-based ML enables real-time intradialytic IDH prediction. If and to what degree this predictive information facilitates the timely deployment of preventive interventions and translates into lower IDH rates and improved patient outcomes warrants prospective studies.

SUPPLEMENTARY DATA

Supplementary data are available at *ndt* online.

FUNDING

No external funding was provided for the conduct of the program. Analysis was supported by Fresenius Medical Care employee salaries and company infrastructure.

AUTHORS' CONTRIBUTIONS

Design was performed by H.Z., P.K., P.W. and Z.K. The interpretation, drafting and revision of this manuscript was conducted by all authors. The decision to submit this manuscript for publication was jointly made by all authors, and the manuscript was confirmed to be accurate and approved by all authors.

DATA AVAILABILITY STATEMENT

As this was an internal quality improvement project, data will not be shared publicly.

CONFLICT OF INTEREST STATEMENT

S.C., A.P., L.U., J.L., P.W., Z.K. and F.W.M. are employees of Fresenius Medical Care. P.K., H.Z. and L.-C.W. are employees of the Renal Research Institute, a wholly owned subsidiary of Fresenius Medical Care. J.P.K. is an employee of Maastricht University Medical Center. L.U., S.C., F.W.M. and P.K. have share options/ownership in Fresenius Medical Care. S.C., J.L., L.U., P.K., H.Z. and F.W.M. are inventors on patent(s) in the field of dialysis. P.K. receives honorarium from Up-To-Date, is on the Editorial Board of *Blood Purification* and *Kidney and Blood Pressure Research*. J.L. is a guest editor on the Editorial Board of *Frontiers in Physiology*. F.W.M. has directorships in Fresenius Medical Care Management Board, Goldfinch Bio and Vifor Fresenius Medical Care Renal Pharma.

REFERENCES

1. Keane DF, Raimann JG, Zhang H *et al.* The time of onset of intradialytic hypotension during a hemodialysis session associates with clinical parameters and mortality. *Kidney Int* 2021;**99**:1408–17. https://doi.org/10.1016/j.kint.2021.01.018

2. 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. https://doi.org/10.1159/000500877

3. Sands JJ, Usvyat LA, Sullivan T *et al.* Intradialytic hypotension: frequency, sources of variation and correlation with clinical outcome. *Hemodial Int* 2014;**18**:415–22. https://doi.org/10.1111/hdi.12138

4. Kanbay M, Ertuglu LA, Afsar B *et al.* An update review of intradialytic hypotension: concept, risk factors, clinical implications and management. *Clin Kidney J* 2020;**13**:981–93. https://doi.org/10.1093/ckj/sfaa078

5. Caplin B, Kumar S, Davenport A. Patients' perspective of haemodialysis-associated symptoms. *Nephrol Dial Transplant* 2011;**26**:2656–63. https://doi.org/10.1093/ndt/gfq763

6. Chou JA, Streja E, Nguyen DV *et al.* Intradialytic hypotension, blood pressure changes and mortality risk in incident hemodialysis patients. *Nephrol Dial Transplant* 2018;**33**:149–59. https://doi.org/10.1093/ndt/gfx037

7. Jansen MA, Hart AA, Korevaar JC *et al.* Predictors of the rate of decline of residual renal function in incident dialysis patients. *Kidney Int* 2002;**62**:1046–53. https://doi.org/10.1046/j.1523-1755.2002.00505.x

8. Stefansson BV, Brunelli SM, Cabrera C *et al.* Intradialytic hypotension and risk of cardiovascular disease. *Clin J Am Soc Nephrol* 2014;**9**:2124–32. https://doi.org/10.2215/CJN.02680314

9. Cho A, Lee YK, Oh J *et al.* The relationship between intradialytic hypotension and vascular calcification in hemodialysis patients. *PLoS One* 2017;**12**:e0185846. https://doi.org/10.1371/journal.pone.0185846

10. Gul A, Miskulin D, Harford A *et al.* Intradialytic hypotension. *Curr Opin Nephrol Hypertens* 2016;**25**:545–50. https://doi.org/10.1097/MNH.0000000000000271

11. Barbieri C, Cattinelli I, Neri L *et al.* Development of an artificial intelligence model to guide the management of blood pressure, fluid volume, and dialysis dose in end-stage kidney disease patients: proof of concept and first clinical assessment. *Kidney Dis* 2019;**5**:28–33. https://doi.org/10.1159/000493479

12. Lee H, Yun D, Yoo J *et al.* Deep learning model for real-time prediction of intradialytic hypotension. *Clin J Am Soc Nephrol* 2021;**16**:396–406. https://doi.org/10.2215/CJN.09280620

13. Lin CJ, Chen CY, Wu PC *et al.* Intelligent system to predict intradialytic hypotension in chronic hemodialysis. *J Formos Med Assoc* 2018;**117**:888–93. https://doi.org/10.1016/j.jfma.2018.05.023

14. Thakur SS, Abdul SS, Chiu HS *et al.* Artificial-intelligence-based prediction of clinical events among hemodialysis patients using non-contact sensor data. *Sensors* 2018;**18**:2833. https://doi.org/10.3390/s18092833

15. Amazon. *What is AWS* https://aws.amazon.com/what-is-aws/ (21 March 2023, date last accessed)

16. Amazon. *AWS HIPAA eligible services* https://aws.amazon.com/compliance/hipaa-eligible-services-reference/ (21 March 2023, date last accessed)

17. HP Enterprise. *Edge device* https://www.hpe.com/us/en/servers/edgeline-systems.html (21 March 2023, date last accessed)

18. Flythe JE, Xue H, Lynch KE *et al.* Association of mortality risk with various definitions of intradialytic hypotension. *J Am Soc Nephrol* 2015;**26**:724–34. https://doi.org/10.1681/ASN.2014020222

19. Amazon. SageMaker. 2020. https://aws.amazon.com/sagemaker/ (21 March 2023, date last accessed).

20. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* 2016;785–94. https://doi.org/10.1145/2939672.2939785

21. Dernoncourt F, Nemati S, Kassis EB *et al.* Hyperparameter selection. In: MIT Critical Data (ed.), *Secondary analysis of electronic health records.* Cham (CH): Springer, 2016, 419–27.

22. Lundberg SM, Erion G, Chen H *et al.* from local explanations to global understanding with explainable AI for trees. *Nat Mach Intell* 2020;**2**:56–67. https://doi.org/10.1038/s42256-019-0138-9

23. Lundberg SMLS. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst* 2017;**30**. NeurlPS Proceedings.

24. Park J, Rhee CM, Sim JJ *et al.* A comparative effectiveness research study of the change in blood pressure during hemodialysis treatment and survival. *Kidney Int* 2013;**84**:795–802. https://doi.org/10.1038/ki.2013.237

25. Wang F, Wang Y, Tian Y *et al.* Pattern recognition and prognostic analysis of longitudinal blood pressure records in hemodialysis treatment based on a convolutional neural network. *J Biomed Inform* 2019;**98**:103271. https://doi.org/10.1016/j.jbi.2019.103271

26. Chaudhuri S, Han H, Monaghan C *et al.* Real-time prediction of intradialytic relative blood volume: a proof-of-concept for integrated cloud computing infrastructure. *BMC Nephrol* 2021;**22**:274. https://doi.org/10.1186/s12882-021-02481-0

27. Schytz PA, Mace ML, Soja AM *et al.* Impact of extracorporeal blood flow rate on blood pressure, pulse rate and cardiac output during haemodialysis. *Nephrol Dial Transplant* 2015;**30**:2075–9. https://doi.org/10.1093/ndt/gfv316

28. Sherman RA. We lower blood flow for intradialytic hypotension. *Semin Dial* 2016;**29**:295–6. https://doi.org/10.1111/sdi.12486

29. Mc Causland FR, Brunelli SM, Waikar SS. Dialysis dose and intradialytic hypotension: results from the HEMO study. *Am J Nephrol* 2013;**38**:388–96. https://doi.org/10.1159/000355958