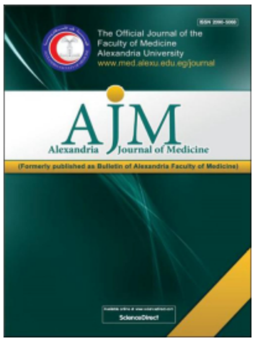

Alexandria Journal of Medicine

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tajm20

A step towards the application of an artificial intelligence model in the prediction of intradialytic complications

Ahmed Mustafa Elbasha, Yasmine Salah Naga, Mai Othman, Nancy Diaa Moussa & Hala Sadik Elwakil

To cite this article: Ahmed Mustafa Elbasha, Yasmine Salah Naga, Mai Othman, Nancy Diaa Moussa & Hala Sadik Elwakil (2022) A step towards the application of an artificial intelligence model in the prediction of intradialytic complications, Alexandria Journal of Medicine, 58:1, 18-30, DOI: 10.1080/20905068.2021.2024349

To link to this article: https://doi.org/10.1080/20905068.2021.2024349

Published online: 09 Mar 2022.

ALEXANDRIA JOURNAL OF MEDICINE
2022, VOL. 58, NO. 1, 18–30
https://doi.org/10.1080/20905068.2021.2024349

A step towards the application of an artificial intelligence model in the prediction of intradialytic complications

Ahmed Mustafa Elbasha[a], Yasmine Salah Naga[a], Mai Othman[b], Nancy Diaa Moussa[b] and Hala Sadik Elwakil[a]

[a]Department of Internal Medicine, Nephrology Unit, Faculty of Medicine, Alexandria, Egypt; [b]Department of Biomedical Engineering, Medical Research Institute, Alexandria, Egypt

ABSTRACT

Introduction: Acute intradialytic complications remain a major burden in end stage renal disease (ESRD) patients on hemodialysis (HD). They often lead to early termination of the HD session affecting dialysis adequacy and patient overall health. The aim of the study was to create an artificial intelligence model and to assess its performance in the prediction of the occurrence of intradialytic clinical events.

Methods: We studied 6000 HD sessions performed for 215 ESRD patients, recording many predictors that included: patient, machine, and environmental factors. These data were collected within 24 weeks, including 12 weeks in the COVID 19 era and were used to develop and train an artificial neural network model (ANN) to predict the occurrence of intradialytic clinical events such as: hypotension, headache, hypertension, cramps, chest pain, nausea, vomiting, and dyspnea.

Findings: Our ANN model showed mean precision and recall of 96% and AUC of 99.3% in binary ANN to predict occurrence of an intradialytic complication (event or no event), while the accuracy of the categorical ANN in predicting the type of event was 82%. We found that heart rate changes, mean systolic pressure, ultrafiltration rate, dialyzate sodium, meal, urea reduction ratio, room humidity and dialysis session duration most strongly influence occurrence of an intradialytic complication.

Discussion: Our ANN model can be used to predict the risk of intradialytic clinical events among HD patients and can support decision-making for healthcare in the frequently understaffed dialysis units, especially in COVID 19 era.

ARTICLE HISTORY
Received 2 November 2021
Accepted 27 December 2021

KEYWORDS
Hemodialysis; intradialytic complications; artificial intelligence; COVID 19

1. Introduction

In 2017, the worldwide prevalence of ESRD reached 2,203 per million with an increase of 1.7% since 2016 and 65.0% since 2000 [1]. The number of patients undergoing renal replacement therapy (RRT) worldwide has increased to 2.5 million and is expected to rise to 5.4 million by 2030 [2]. In the main North African countries including; Egypt, Sudan, Libya, Tunisia, Algeria, and Morocco, the average incidence of ESRD is estimated to be 182 patients per million populations (pmp) [3], which is around 31,000 new cases per year, including 1,700 cases of renal transplantation [3].

Hemodialysis is the main RRT modality in all North African countries [4]. In 2006, the estimated annual incidence of ESRD was around 74 per million [5] and the prevalence of dialysis patients has increased from 225 pmp in 1996 to 483 pmp in 2008 (according to the Egyptian renal registry) [6]. In 2019, the Egyptian renal data system estimated that dialysis patients exceed 50,000 (an estimated prevalence of around 50,000/Egypt's population in millions) [7]. In Alexandria (the second-largest city in Egypt) the estimated HD prevalence rate was around 710 pmp in 2019 [8].

Hemodialysis is associated with many intradialytic as well long-term complications; cardiovascular complications are the most common in both categories. Among intradialytic complications, intradialytic hypotension is the most frequent event, occurring in 20–50% of patients [9]. Muscle cramps are also common during dialysis and two or more interdialytic cramps per week occur in 25% of patients [10]. Other common complications include nausea, vomiting with a rate of 5%–15% and headache with a rate of 5%–10% [11].

In Egypt, ESRD mainly affects patients in their economically productive years with some predilection to male gender [8]. Therefore, inadequate dialysis in these patients renders them incapable of continuing work. In addition, there is understaffing in dialysis units, making close monitoring of the over-crowded dialysis units difficult and leading to more intradialytic complications.

CONTACT Ahmed Mustafa Elbasha a_elbasha14@alexmed.edu.eg Department of Internal Medicine, Nephrology Unit, Faculty of Medicine, El Azareeta Medical Campus Champollion Street, El-Khartoum Square, Alexandria Governorate, Alexandria, Egypt

Artificial intelligence (AI) is a field of science and engineering concerned with the computational information of what's generally referred to as intelligent behavior, and with the development of objects that exhibit such behavior [12]. Machine learning (ML) is one of the prime branches of AI. Machine learning can be described as a collection of algorithms that have the capability of learning and improving from experience, without being specifically programmed for a particular task. Random forest, support vector machine and artificial neural network (ANN) are examples of machine learning algorithms [13]. Many artificial intelligence-based algorithms have been accepted by the Food and Drug Administration (FDA) to be used in clinical practice, they do not replace the role of physician but they may complement it in a wide range of applications in medicine. This ranges from predicting patient outcomes such as diagnosis and treatment efficacy, to discovering patterns in large databases and to understanding disease pathogenesis.

Artificial neural network has been widely used in many medical fields including computer-aided surgery operation navigation systems [14], virtual tools for surgery [14], pain management, and psychiatric disorders [15]. Artificial neural network has effectively been used for solving complex and unpredictable problems in many fields of clinical medicine without the need for statistical models [16]. Examples of the use of artificial intelligence in nephrology are diagnosis of chronic kidney diseases (e.g. IgA nephropathy [17], glomerular vs tubular renal disease [18]), evaluation of hemodialysis efficiency by urea kinetic modeling [19] and diagnosis of renal transplant rejection [20].

Large quantities of data can be collected during HD sessions. Biostatistical methods have been traditionally used in medical data analysis. However, exploring large sets of data and complex non-linear associations is better achieved by artificial neural network. Artificial neural network has been successfully used to detect relevant connections in a database and has been commonly used for the detection, evaluation, and estimation of outcomes in many clinical settings [21].

Most developing countries suffer from insufficient number of medical personnel and limited resources that may not allow adequate monitoring of patients during the dialysis procedure and that raise the risk of intradialytic complications. Application of a predictive artificial intelligence -generated model may aid in identifying patients who need closer monitoring and allows timely and appropriate intervention [22].

The aim of this work was to develop an artificial neural network model that can predict the risk of intradialytic clinical events among regular HD patients in order to prevent its occurrence and improve patients' quality of life and suggesting proper individualized management.

2. Materials and methods

2.1. Setting, design and study sample

In this study, we included 215 patients on regular HD sessions at El-Mowasah University Hospital, Alexandria, Egypt with exclusion of patients on hemodialysis for less than 6 months, patients younger than 18 years-old and patients who received renal transplantation during the study course. El-Mowasah hospital is one the Alexandria University hospitals and is considered as a tertiary referral dialysis center receiving dialysis patients with multiple comorbidities. Approval of the ethics committee of research of the Alexandria Faculty of Medicine and an informed consent from each patient were obtained.

2.2. Measures

2.2.1. Endpoint definitions

Clinical event was defined as experience on of the following complications: muscle cramps, headache, dyspnea or nausea and vomiting even if not requiring session termination.

Intradialytic hypotension (IDH) was defined according to the Kidney Disease Outcomes Quality Initiative (KDOQI) as a decrease in SBP ≥20 mm Hg or a decrease in MAP by ≥10 mm Hg associated with a clinical event and the need for nursing intervention [23].

Because no standard definition of intradialytic hypertension exists, we used a definition used in prior clinical studies. Intradialytic hypertension was defined as any of the following: (a) an increase of 15 mmHg or more in (MAP) of within or after dialysis [24], (b) an increase of 10 mmHg in the systolic blood pressure before and after dialysis [25,26].

HD sessions were categorized into two groups, sessions in which the patient experienced a clinical event (CE group) and sessions without any significant event (no-event group), depending on presence or absence of the following intradialytic clinical events: intradialytic hypotension, hypertension, muscle cramps, headache, dyspnea, nausea, and vomiting. If multiple events occurred, only the first event was recorded.

2.2.2. Input variables

During each HD session, the following data were captured: patient factors including: heart rate, blood pressure, weight gain, demographic data, laboratory investigations, original kidney disease, and other comorbidities; machine factors including: modality, duration, clotting, ultrafiltration rate, dialyzate sodium, bicarbonate, flow and temperature; and other factors namely, room temperature and humidity, medications, meal, fluid and caffeine intake (Table 1).

Our raw data included collection of 50 variable in 6000 HD sessions counting about 300,000 data entry.

Table 1. List of variables used and frequency of collection.

Predictor	Frequency Collected
Demographic factors:Age (years) Gender (male)Dialysis vintage	OnceOnceOnce
Patient comorbidities:Original kidney diseaseDiabetes, Heart disease and Peripheral vascular diseases	OnceOnce
Dialysis-specific factors:Modality (HDF, HD)Duration and frequency of the sessionsMean dialyzer blood flow (ml/min) Mean Blood pump flow (ml/min)Ultrafiltration rate (ml/min) Dialyzer membrane charactersDialyzate fluid temperature (0 C) and componentsOccurrence of dialysis lines clottingType of vascular accessWeight loss percentage (%)Interdialytic weight gain	Every sessionEvery sessionEvery sessionEvery sessionEvery sessionEvery sessionEvery sessionEvery sessionEvery sessionOnceEvery sessionEvery session
Dialysis hemodynamics:Heart rate (HR) Mean heart rate (bpm) HR changes (increase, decrease or steady)**Blood pressure (BP)**: Systolic pressure (mmHg) Diastolic pressure (mmHg)Mean arterial pressure MAP (mmHg) Pulse pressure (mmHg)Blood pressure changes (increase, decrease or steady)	Predialysis,1st and 2nd hourAfter 1st hour of sessionPredialysis,1st and 2nd hourPredialysis,1st and 2nd hourPredialysis,1st and 2nd hourPredialysis,1st and 2nd hourAfter 1st hour of session
Laboratory investigations:Urea reduction rate (%)Creatinine (Pre HD)Phosphate (mg/dl) Calcium (mg/dl)Hemoglobin (g/dl)White cell count (cells per liter)Platelet cell count (cells per liter)Serum Albumin (mg/dl)	Every 4 weeksEvery 4 weeksEvery 8 weeksEvery 8 weeksEvery 4 weeksEvery 4 weeksEvery 4 weeksEvery 8 weeks
Intravenous medications: Epoetin alfaHeparinIron sucrose	Every sessionEvery sessionEvery session
Other factors:Room temperature and humidityMeal and coffee intake	Every sessionEvery session

2.2.3. Data engineering (data cleaning, feature generation and handling of missing data)

The machine-learning estimation depends mainly on the quality of the used data. The initial dataset must be structured in a way that simulates the real-world environment, to make sure that the model will correctly predict unseen data.

Three analyses were performed on the collected data, namely:

(I) The first analysis (A1): The 6000 dialysis sessions were divided into two groups based on the occurrence of an intradialytic event: a clinical event occurred in 47.9% (2,874 sessions), while no event occurred in 52.1% (3,126 sessions).

(II) The second analysis (A2): The 6000 dialysis sessions were divided into two groups based on the occurrence of an intradialytic hypotension, as it is the most frequently recorded event: an intradialytic hypotension event occurred in 16.4% (982 sessions), while no hypotension occurred in 83.6% (5,018 sessions).

(III) The third analysis (A3): The 2,874 events cases were divided into the 7 recorded outcomes in the third analysis, no event was also included, so 8 outcomes were examined (Table 2).

As data were unsuitable to be processed by machine learning algorithms, preliminary data preparation procedures were carried out. Before starting, categorical features were checked and converted into numerical form. The missing data were recovered by using standard statistical methods. To deal with missing data, we applied the following imputation strategy:

(1) If the missing information was categorical, it was replaced with the median.
(2) If the missing information was continuous, it was replaced with the mean.

Since different magnitudes of data would lead to domination of the higher features over the smaller ones, depending on the units adopted and the design of the operation, we normalized the data before the study. Using this method, the features were transformed so that their values were located within the specified range, between 0 (lower bound) and 1 (upper bound).

After preprocessing and dealing with missing values, the dataset was divided into 80% training set and a 20% test set from different dialysis sessions in the training set patients. The training subset was fed as input to the artificial neural network classifier. Finally, the test subset was used to check the performance of the trained classifier network.

2.2.4. Statistical analysis (ML framework and analytic strategy)

After data was collected, it was revised, coded and entered to the statistical software SPSS version 21. Descriptive statistics were used for summarization of data using frequency distribution tables and graphs. SPSS and SAS software were used for statistical analysis and construction.

For quantitative variables, mean and standard deviation (SD) were calculated. Quantitative variables were expressed as percentage. Several testes were used and P value of <0.05 was considered significantly.

The following statistical tests were used:

Table 2. The percentage of the target outcomes.

The output	The rate in the dataset
No event	52.1%
Hypotension	16.4%
Hypertension	7.4%
Headache	7%
Cramps	5.8%
Nausea Vomiting	4.6%
Chest Pain	3.4%
Dyspnea	3.3%

(1) Chi-square (X2) test (Fisher or Monte Carlo) was used for analysis of categorical data. The 5% level was used as the cut-off value for Statistical significance.
(2) Student t-test was used to compare two groups for normally distributed quantitative variables.

ANN is composed of a hierarchical organization set of artificial neurons organized in successive layers. It is trained on data to predict the target outcome for a new input of similar data. To compute a specific input-output non-linear relation, the model should be trained with a sufficiently large set of input and output data pairs [27].

To design the ANN framework, the number of neurons in each layer, the required activation function, the optimizer used, and a few other network parameters were adjusted [28].

The sequential ANN model was implemented. It consisted of a dense input layer, hidden layer, and output layer. The input layer included 50 neurons (corresponding to the 50 variables), as well as 128 neurons in the hidden layer. The number of neurons in the hidden layer follows the following rule: If the final number of input attributes in each training subset is x, we should use at least the closest number to 2x in the power of 2. It is good to have the number in the power of 2, as it helps the computation of the network to be faster. For example, the 50 input attributes in our training subset, preferably started with ($2 \times 50 = 100$) and used the closest power of 2, so 128, as shown in Figures 1 and 2.

The number of one hidden layer and two hidden layers were tested, but the evaluation metrics showed that artificial neural network with 1 layer had the same evaluation metrics as with 2 layers. Therefore, there was no need to use 2 layers as it is more complex.

For the hidden layer, an activation function should be implemented to enable non-linearity. For this reason, Rectified Linear Unit (ReLU) was applied to the hidden layer. Finally, the output layer, responsible for the final classification, had a value of one neuron for the binary classification in the first two analyses, where class 0 referred to the no event sessions and class 1 indicated the presence of clinical event during the dialysis session as shown in Figure 1, sigmoid activation was added to the output layer. Binary cross-entropy was used as a loss function.

Furthermore, in the third analysis, eight neurons (7 events and no event) in the output layer were applied for the categorical classification as shown in Figure 2. In this case, the SoftMax function was applied as the output layer activation function, as it is a multi-classification problem. Categorical cross-entropy was used as a loss function.

The grid search estimator was used to optimize and find the model hyper-parameters (epochs and batch size) that give the highest accuracy. We used grid

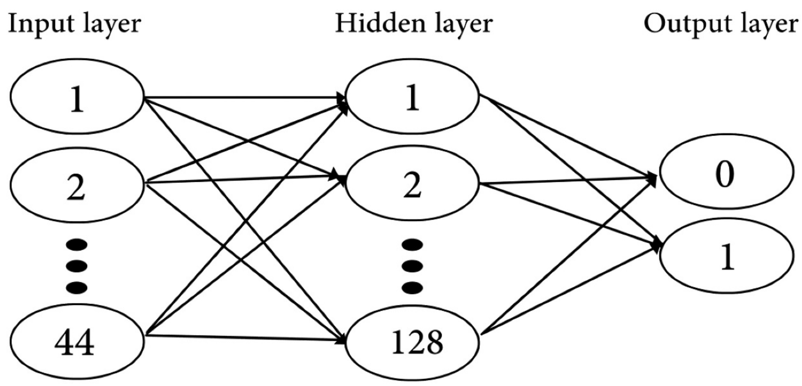

Figure 1. The neural network structure for binary classification.

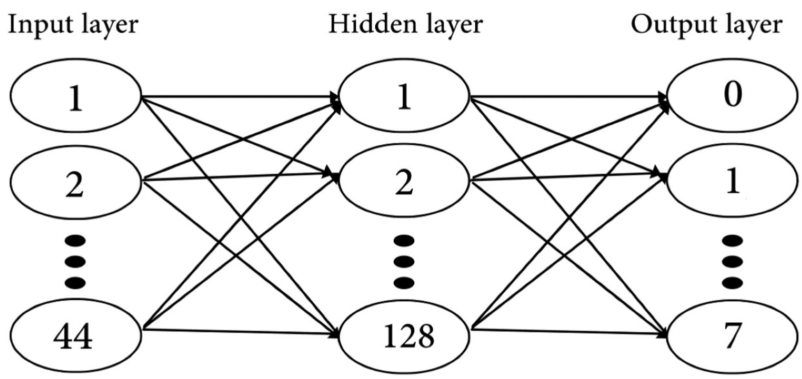

Figure 2. The neural network structure for categorical classification.

search to optimize batch size and epochs. The range for batch size was = [16,32,64] and for epochs was = [25,50,100]. For the first analysis, the best values were 32 and 50 for batch size and epochs respectively. For the first analysis, the best values were 32 and 50, for the second, 64 and 100, while for the third analysis it was 32 and 100 for batch size and epochs, respectively.

All experiments were based on computer programming Python interpreted language. Jupyter notebook Python 3.6 is an open-source programming language that was used to develop the proposed techniques. The experiments were performed with a 2 GHz Intel Core i7 Lenovo with 8GB of RAM, running on Windows 10 operating system. NumPy, Pandas, Scikit learn, Keras, and matplotlib libraries were used in preprocessing training, and evaluation of the used models.

The performance of artificial neural network in the three analyses was presented, and the most effective features in prediction were analyzed. Multiple measures were used to judge the performance of the model. In this study, the precision, recall, accuracy, and f1-score were calculated. In the binary classification, we also estimated receiver operating characteristic (ROC), and Precision-Recall curves as evaluation measures of the model.

2.2.5. Features selection

The most effective 26 features were selected based on filter feature selection technique. It is based on ranking the features according to their usefulness in the prediction and evaluating the importance of features based on the properties of data. Then, the output of the selected features is applied to the machine learning algorithm. The filter approach uses a variety of methods, such as Pearson's correlation coefficient, and mutual information. Pearson's determines the correlation between the features and the output class, while mutual information shows the amount of information between the feature, and the output class. For feature selection, we first removed the correlated input features, assuming that the features were highly correlated if their correlation coefficient is greater than 0.75. In this step, we removed seven features. Second, we chose only features that are highly correlated with the output, the feature that had no impact on the output was removed. If the correlation coefficient was between −0.05 and 0.05, we neglected it. In this step we removed 15 features. Third, we chose features that had high mutual information with the output. In this step, we removed two features.

3. Results

A total of 6,000 HD sessions were observed within 24 weeks (from November 2019 to May 2020). Prospective clinical data were collected manually every session, while laboratory investigations were reported from the routine monthly investigations. Our study included 215 regular hemodialysis patients, the main demographic and clinical characteristics of the studied patients are summarized in (Table 3). Of note, 56.8% of these patients had comorbidities including: uncontrolled HTN and DM, cardiac or liver diseases, stroke and malignancy. 55.3% of patients were males and 70.6% aged between 18 and 60 years old. The most common cause of renal failure was hypertension

Table 3. Demographic and clinical description of (215) HD patients.

Variable	Mean (± SD), Number (Percent)
Age (year)Percentage of patients aged (18–60 year)	55.11 ± 12.9 years old152 (70.6%)
Gender:	
MaleFemale	119 (55.3%)96 (44.7%)
Employment:	
UnemployedFull timePart timeRetired	111 (51.6%)22 (10.2%)35 (16.3%)47 (21.9%)
Original kidney disease	
HypertensionDiabetes mellitusGlomerulonephritisPyelonephritisContrast induced nephropathyAnalgesic nephropathyObstructive uropathySystemic lupus erythematosusAdult polycystic kidney disease Unknown	82 (38.1%)42 (19.5%)19 (9.0%)8 (3.7%)4 (1.9%)8 (3.7%) 10 (4.6%)5 (2.3%)10 (4.6%)27 (12.6%)
Comorbidities	
No comorbiditiesUncontrolled hypertensionUncontrolled diabetes mellitusCardiac diseaseHepatic diseaseStrokeMalignancy	93 (43.2%)35 (16.3%)29 (13.5%)36 (16.7%)13 (6.1%)5 (2.3%)4 (1.9%)
Virology	
NegativeHepatitis C virus positiveHepatitis B virus positiveHepatitis C & B virus positive	171 (79.5%)39 (18.1%)4 (1.9%)1 (0.5%)
Vascular access	
Arterio-venous fistula (AVF)Arterio-venous graft (AVG)Permanent catheterTemporary catheter	173 (80.5%)4 (1.9%)29 (13.5%)9 (4.1%)
Type of dialysis	
Hemodialysis (HD)Hemodiafiltration (HDF)	4281 session (71.3%)1719 session (28.7%)
Dialysis vintage (years)	6.1452 ± 4.63 years
Frequency of intradialytic events	
NoLess than one event /weekOne event or more/week	67 (31.2%)57 (26.5%)91 (42.3%)

Table 4. Comparison between the event and no event group according to the 26 important variables.

Predictor	No events sessions(n = 3126)	Events sessions(n = 2874)	p-value
Demographic factors:			
•Age (years)	42.4 ± 14.7	59.4 ± 13.2	t(p): 47.039 (<0.001*)
Patient comorbidities:			
•Heart disease	409 (13.1%)	1132 (39.4%)	χ2(p): 542.780 (<0.001*)
•Peripheral vascular diseases	303 (9.7%)	757 (26.3%)	χ2(p): 285.266 (<0.001*)
Dialysis-specific factors:			
•Modality HDHDF	1794 (57.4%)1332 (42.6%)	2487 (86.5%)387 (13.5%)	χ2(p): 622.198 (<0.001*)
•Duration of the sessions (hours)	3.9 ± 0.5	3.8 ± 0.6	t(p): 4.751 (<0.001*)
•Interdialytic period (days)	1.4 ± 0.3	1.9 ± 0.5	t(p): 2.658 (<0.001*)
•Ultrafiltration rate (l/hour)	0.5 ± 0.4	0.8 ± 0.3	t(p): 37.266 (<0.001*)
•Dialyzer surface area (m^2)1.31.51.71.82.1	390 (12.5%)813 (26%)910 (29.1%)626 (20%)387 (12.4%)	322 (11.2%)570 (19.8%)474 (16.5%)770 (26.8%)738 (25.7%)	χ2(p): 2.317 (0.128)χ2(p): 32.187 (<0.001*)χ2(p): 134.340 (<0.001*)χ2(p): 38.398 (<0.001*)χ2(p): 173.821 (<0.001*)
•Dialyzate fluid temperature (0 C)	36.7 ± 1.9	36.7 ± 0.8	t(p): 1.486 (0.137)
•Dialyzate sodium (mEq/l)	136.4 ± 2	135.3 ± 3	t(p): 16.849 (<0.001*)
● Mean dialyzate flow (ml/min)	633.5 ± 149.1	589.5 ± 137.4	t(p): 11.900 (<0.001*)
•Occurrence of dialysis lines clotting	103 (3.3%)	614 (21.4%)	χ2(p): 464.849 (<0.001*)
Vital signs			
Heart rate changes after one hour of HDStableDecreasingIncreasing	2225 (71.2%)511 (16.3%)390 (12.5%)	545 (19%)1303 (45.3%)1026 (35.7%)	χ2(p): 1642.56(<0.001*χ2(p): 596.635 (<0.001*)χ2(p): 447.888 (<0.001*)
BP changes after one hour of HDStableDecreasingIncreasing	3046 (97.4%)43 (1.4%)37 (1.2%)	137 (4.8%)1811 (63.0%)926 (32.2%)	χ2(p): 5164.709(<0.001*)
● Pre-dialysis mean arterial pressure	96.8 ± 16.6	103.4 ± 17.1	t(p): 15.229 (<0.001*)
Laboratory investigations:			
● Urea reduction ratio (%)	60.3 ± 10.7	66.9 ± 9.2	t(p): 25.803 (<0.001*)
•Creatinine mg/dl (Pre HD)	8.2 ± 2.3	9.1 ± 2.7	t(p): 13.369 (<0.001*)
•Calcium (mg/dl)	8.5 ± 1.2	8.6 ± 1.5	t(p): 2.932 (0.003*)
•Hemoglobin (g/dl)	10.2 ± 1.5	9.7 ± 2.2	t(p): 9.153 (<0.001*)
•White cell count (cells per liter)	8.4 ± 2.7	9.7 ± 4.3	t(p): 13.441 (<0.001*)
•Serum Albumin (mg/dl)	3.4 ± 0.6	3.1 ± 0.8	t(p): 16.645 (<0.001*)
Medications:			
•Antihypertensive before dialysis	447 (14.3%)	1078 (37.5%)	χ2(p): 425.488 (<0.001*)
Other factors:			
● Room temperature	24.1 ± 3.9	25.1 ± 5.2	t(p): 9.089 (<0.001*)
● Room humidity	44.2 ± 7.4	50.9 ± 10.1	t(p): 29.059 (<0.001*)
Meal intake during the sessionNoYes	2542 (81.3%)544 (17.4%)	1110 (38.6%)1543 (53.7%)	χ2(p): 1145.960 (<0.001*)χ2(p): 869.097 (<0.001*)

Quantitative data was expressed using Mean ± SD. While Qualitative data was expressed using Number (%)
χ^2: **Chi square test: Student t-test**
p: p value for comparing between **No event** and event
*: Statistically significant at p ≤ 0.05

followed by DM then glomerulonephritis (GN) and adult polycystic kidney disease (APKD). Most patients are dialyzing using a functioning AVF. 28.7% of dialysis sessions were online HDF.

Statistical analysis was done comparing between event and nonevent group regarding different parameters and the comparison between the event and no event group according to the 26 important variables was shown in (Table 4).

Primary analysis of the data identified that 26 of the variables had the greatest impact on the risk of a clinical event. These feature are ranked in descending order: blood pressure, heart rate changes (increase or decrease in blood pressure (BP) and heart rate (HR) after one hour of HD session), age, session duration, ultrafiltration rate, hemoglobin level, room humidity, white blood cells (WBCs), dialyzate sodium (that was variable in sessions ranged from 133 to 142 mEq/L), dialyzate temperature, urea reduction ratio (URR), serum creatinine, interdialytic period, serum albumin, dialysis lines clotting, heart disease, dialyzer surface area, meal intake, peripheral vascular disease (PVD), HD modality, antihypertensive medications, DM, dialyzate flow, room temperature, serum calcium, predialysis mean arterial pressure (MAP).

The developed artificial neural network model had an accuracy of 96% in predicting event and nonevent, 94% in predicting hypotension, and 82% in the multi-classes prediction to specify which event will happen.

(Table 5) shows the performance of the neural network algorithm applied to two different analyses with binary target output. The proposed artificial neural network had better performance for the first analysis, as the dataset was balanced (52.1% nonevent, and 47.9% event) while the second analysis was imbalanced (83.6% non-hypotension, and 16.4% hypotension). This shows that the performance of an artificial neural network in general is better in balanced binary analyses.

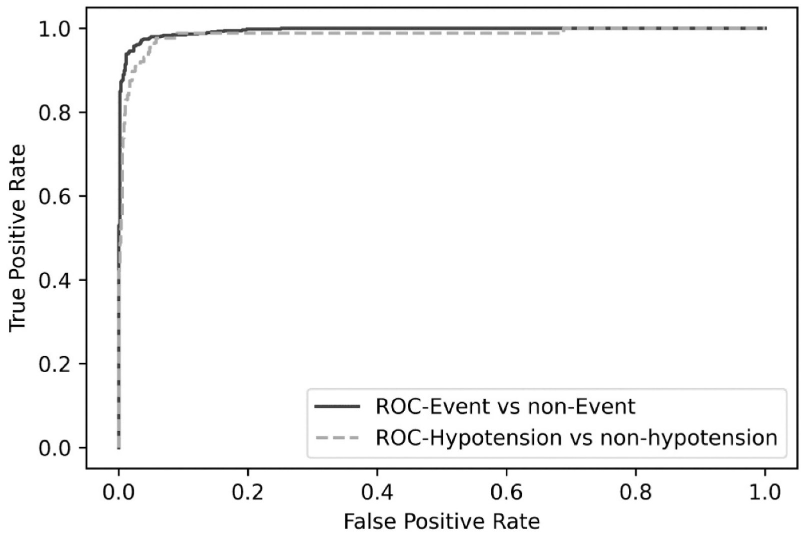

Figure 3. The ROC curve for the two binary analyses A1 and A2.

Figure 3 shows the ROC curve of A1 and A2. We can observe that in the first analysis, the line approaches its goal, which is to obtain the curve nearest to one on the Y-axis because the ROC value is

Table 5. The performance of the binary and categorical ANN.

Binary ANN					
	Accuracy	Specificity	Sensitivity	F1- score	AUC-ROC
Event vs no-Event(**first analysis**)	96%	96%	96%	96%	99.3%
Hypotension vs no-hypotension(**second analysis**)	94%	94%	94%	94%	97.8%
Categorical ANN					
	Accuracy	Specificity	Sensitivity		F1- score
0 (No event)	82%	95%	97%		96%
1 (Hypotension)		76%	89%		82%
2 (Headache)		74%	50%		60%
3 (Dyspnea)		69%	52%		59%
4 (Chest pain)		40%	27%		32%
5 (Nausea-Vomiting)		67%	21%		32%
6 (Cramps)		47%	53%		49%
7 (Hypertension)		68%	98%		80%

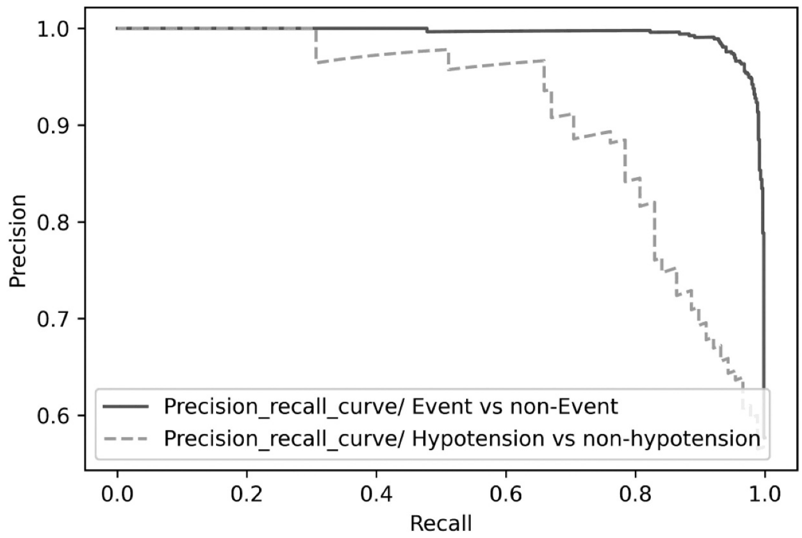

Figure 4. The precision-recall curve for the two binary analyses.

99.1% for both classes. In the second analysis, its performance decreased slightly but still gives high results of 98.3%. The performance of the second analysis is lower because of the lower event rate making the data imbalanced.

Finally, the precision-recall curve for the two binary analyses is showed in Figure 4. We can observe that in the first analysis, the line approaches its goal, which is to obtain the curve nearest to one on the Y-axis because the AUC value is 0.99 for both classes. This value indicates that the classifier performs well in detecting the occurrence of complications during the dialysis session. In the second analysis, its value is 92%.

After features selection, the performance of the two binary analyses was the same with the selected features (Table 5). While reducing the number of features, reduced the complexity of the model. However, the total accuracy of multi-classification was decreased by only 1%, the sensitivity, f1-score of hypotension were increased by 2%. The result of this model is shown in (Table 6).

4. Discussion

Despite improvement in the dialysis techniques and machines, the occurrence of intradialytic complications in HD patients is still common. These complications are attributed to many factors and most of which are preventable. The prediction of these complications could help staff at dialysis units to identify those patients at higher risk, who need closer monitoring and timely intervention to prevent the intradialytic complication and its effect on the session duration and quality.

Table 6. The performance of the categorical ANN with the most significant 26 features.

	Accuracy	Specificity	Sensitivity	F1- score
0 (noevent)	81%	93%	98%	96%
1 (Hypotension)		78%	91%	84%
2 (Headache)		68%	62%	65%
3 (Dyspnea)		38%	35%	37%
4 (Chest pain)		21%	18%	20%
5 (Nausea-Vomiting)		60%	16%	25%
6 (Cramps)		67%	45%	54%
7 (Hypertension)		69%	85%	76%

In our study, we selected a model using a neural network for the prediction of complications in regular hemodialysis patients because it is simple, accurate, and relatively stable to outliers and noise. In addition, it offers useful internal error estimates, strength, correlations and variable significance [29].

Our proposed model (artificial neural network with grid search) allowed the prediction of intradialytic complications during the dialysis with high accuracy.

Our ANN model used routinely measured variables as predictors. Indeed, we used patient characteristics, co-morbidities, treatment parameters, laboratory investigations, and some environmental variables that are easily collectible by nurses or physicians. Therefore, this model may be easily applied in under-equipped dialysis units.

This study focused on the most common seven complications occurring during the HD session. We applied our model in three analyses with different targets. It is quite accurate in all aspects of binary and multi-classification. Our result in binary classification showed accuracy of 96% with area under ROC of 99.3% and accuracy in multi-classification reached 82%.

Intradialytic hypotension was the most common complication in our study followed by hypertension, headache, and muscle cramps. This findings are similar to the results of other studies [30,31].

The accuracy of the artificial neural network applied to the third analysis with eight categorical outcomes was 82% which is less than the performance of the artificial

Table 7. The performance of the binary ANN with the most significant 26 features.

	Accuracy	Specificity	Sensitivity	F1-score	AUC-ROC
Event vs nonevent **(first analysis)**	96%	96%	96%	96%	99.3%
Hypotension vs non-hypotension **(second analysis)**	94%	94%	94%	94%	97.8%

neural network model on the binary analyses, as it is more complex for the model to distinguish between eight different classes. Despite the low accuracy in multiclassification, the high precision can help identify patients at risk of the complication and may allow timely intervention. The precision and f1-score vary among the different outcomes as the ratio of each event in the data differs. Indicating that artificial neural network may give better results in cohorts where the studied outcome is more common.

The mean values of all predictors are shown in Table 7. The basic statistical analysis shows statistically significant differences in most of the studied parameters between the sessions with an event and those without event.

Another method to improve artificial neural network output is to reduce the number of features. Feature selection allowed us to identify 26 of the 50 features with the highest impact on outcome. The most significant 26 features were extracted to reduce the total number of features. The advantage of reducing the total number of features is: it makes the model easier to apply as it reduces the number of variables that dialysis unit staff have to collect 26 features rather than 50 and the complexity of the artificial neural network model was reduced.

Other recent studies have also attempted to establish an artificial intelligence model to predict intradialytic events. In 2019, Putra FR et al. [32] conducted a similar artificial neural network model in which 3,237 sessions were analyzed in 109 patients over 23 weeks. They used three main predictors including heart rate, respiration rates, and movement data measured by a special device that is not widely available, in addition to the demographic details of patients, such as gender, age, height, and weight. They then combined the five events in a binary classification problem (event vs. nonevent). Patients who reported emergency visit, muscle spasm, inpatient, emergency visit and inpatient or sudden death during observation were classified in the class-event. Their model had a mean precision and recall of 93.45%, and area under ROC was 96.7%. however their model could not predict which event would occur [32].

Thakur SS et al. [33] also used a non-contact sensor device to monitor vital parameters like the heart rate, respiration rate, and heart rate variability of 109 hemodialysis (HD) patients, then developed a supervised machine-learning-based prediction model to predict event or no-event based on the sensor data and demographic information. A mean area under ROC of 90.16% with 96.21% mean precision, and 88.47% mean recall was achieved. These two studies used advanced devices like non-contact sensor that are not widely available in dialysis units so their model cannot be widely applied unlike our model which utilizes easily measured variables. Also, they used only 3 vital parameters while we used more than 50 parameters.

In addition, Barbieri C et al. [34] formed an artificial neural network model analyzing 766,000 HD session records during hemodialysis sessions in 2019 that included roughly 60 variables representing patient characteristics including session-specific Kt/V, ultrafiltration volume, heart rate, and BP. This study used large number of sessions with a wide variety of variables but despite the large number of data used, they did not consider dialysis machine parameters or environmental factors in their collected variables. They created a neural network model to predict the individualized, session-specific patient reaction to dialysis-related prescriptions on multiple relevant hemodynamic parameters (e.g. intradialytic heart rate and BP changes and trends) and dialysis adequacy parameters (e.g. Kt/V and fluid removal), but no other complications were studied in contrast to our study, which included seven common intradialytic complications.

Cheng J et al. [35] also has formed a new predictive model for prediction of intradialytic hypotension in chronic hemodialysis using AI based on a database of 55,516 HD sessions of 653 HD patients, resulting in 285,705 valid BP records. Their logistic regression model showed the sensitivity of 86% and specificity of 81% for both nadir systolic BP (SBP) of <90 mmHg and <100 mmHg, suggesting good performance in prediction of IDH. This model used dialysis settings (including machine temperature, conductivity, and UF rate), baseline demographic variables (such as age, sex, DM, and dry weight) as predictors, but our model showed higher specicifty and sensitivity in predicting intradialytic hypotension. The common limitation of this study and ours that we defined IDH based on SBP values, without considering symptoms and interventions.

Hojun L et al. [36] created a model capable of predicting intradialytic hypotension. They applied a deep learning model using data from 261,647 hemodialysis sessions, divided them into training (70%), validation (5%), calibration (5%), and testing (20%) sets. Their artificial neural network model achieved an

AUC of 0.94 (95% confidence intervals, 0.94 to 0.94). They used three definitions for IDH. IDH-1 was defined when intradialytic nadir systolic BP was <90 mm Hg within 1 hour [2]. When IDH was defined as a decrease in systolic BP of ≥20 mm Hg and/or a decrease in mean arterial pressure of ≥10 mm Hg within 1 hour, the reference BPs were determined at initial (IDH-2) or prediction (IDH-3) time point. Variables recorded included age, sex, vital signs, comorbidities, medications, and laboratory findings. However, our prediction model not only included intradialytic hypotension but also examined overall intradialytic complications instead of only focusing on the hypotension.

Most recently, in 2021 2021, Liu Y et al. [37] developed machine learning algorithms to predict intradialytic adverse events. Data were collected from 108 patients on regular HD in a total of 4221 HD sessions. Dialysis data were collected automatically by HD devices, and physiological data were recorded by medical staff but in our study whole data are collected manually by physicians and nurses. Their developed algorithm predicted overall intradialytic adverse events, with an area under the curve (AUC) of 0.83, sensitivity of 0.53, and specificity of 0.96. The algorithm also predicted muscle cramps, with an AUC of 0.85, and blood pressure elevation, with an AUC of 0.93. The most common adverse event in their study was muscle cramps but, in our study, IDH was the most common complication. Their results shown the top 16 features that majorly contributed to predicting muscle cramps included patient characteristics, venous pressure, trans-membranous pressure, ultrafiltration, blood flow rate, and pulse pressure, we didn't include venous pressure and trans-membranous pressure as predictors in our study.

Artificial intelligence has also been used to predict long-term outcome in HD patients. Titapiccolo J et al. [38]trained a Lasso logistic regression model and a random forest model and used it for predicting the cardiovascular outcome of incident hemodialysis (HD) patients. Data relating to the dialysis machine properties and the vital signs of 4246 incident hemodialysis patients were collected during the first 18 months of HD and then used to predict the occurrence of cardiovascular events within a 6-month time interval. Random forest showed higher performance with AUC of the ROC curve and sensitivity higher than 70%. The most important variables in the model were blood test variables such as the total protein content, percentage value of albumin, total protein content, creatinine, C reactive protein, age of patients and weight loss in the first six months. Although the studied outcome is different from our study, some variables like serum albumin level, patient age, and ultrafiltration rate were also predictive in our study, indicating that they play a role in short-term as well as long-term outcomes in HD patients.

As early as in 2001, Akl et al. [19] in a study conducted in Mansoura university applied an artificial neural network model to study and predict concentrations of urea during a hemodialysis session. Urea blood concentrations, patient weight and total urea clearance were calculated at 30-minute intervals during the HD session by direct dialysis quantification (in 15 chronic hemodialysis patients), then they trained an artificial neural network model to recognize the evolution of measured urea concentrations and were then able to anticipate the time taken for the hemodialysis session to achieve a target solute removal index.

Another study, exploring outcome in HD patient was published by Akibilgic et al. in 2019 [39]. This study included 27,615 US veterans who had developed ESRD. To predict 30-, 90-, 180-, and 365-day all-cause mortality after start of HD, they used a random forest method on 49 variables obtained before dialysis initiation to predict the outcomes. The final random forest model provided C-statistics of 0.7185, 0.7446, 0.7504, and 0.7488 for predicting risk of death within the 4 different time windows. However, their study did not include multiple variables known to affect mortality early in dialysis including: vascular access at the time of dialysis transition, the use of certain medications or the length of dialysis period, highlighting the fact that models may miss important factors not included in the analysis by their designers.

Most mentioned models did not reach the same level of specificity and accuracy that our model reached in the binary classification analyses. The lower accuracy in multi-classification analysis is due to the limited number of patients who faced some of the described events in each class. Despite the relatively lower accuracy, it may prove very useful.

Our study has several strengths:

- We used a large dataset comprised prospectively collected data of 50 variables recorded during 6000 sessions, in a diverse cohort of patients recorded every session for 24 weeks.
- We used variables routinely and easily collected by limited number of nurses and without the need for special devices or advanced dialysis machines that record patient's vitals continuously, where this study was performed in COVID 19 era, so our model will be easily applicable in the underequipped clinics and units.
- We analyzed the occurrence of seven common intradialytic complications.

- We used a multiclassification artificial neural network model that can identify and predict which event of the 7 events can happen which was not previously attempted in other studies.
- We included some unique environmental predictors as room humidity and temperature that can be important factors in a hot humid climate country like Egypt and could be applied in various developing countries.

Our study also has some limitations: Although our dataset is relatively dense, there are many other factors that may increase the probability of intradialytic complications like CRP, parathormones and dialyzate calcium and potassium. This study shows effects of several variables simultaneously and no single factor can be blamed for the intradialytic event, so further studies comparing each factor separately can better delineate causality. We developed our prediction models from data in a single tertiary referral center including a high percentage of patients with comorbidities, which precludes the direct application to other patient groups as routine dialysis care, and patient characteristics may vary from center to center.

We will soon start to externally validate our artificial neural network model in a multi-center trial set to include a larger number of patients and sessions. Use of a larger dataset will also allow us to increase the accuracy of the model especially in multiclassification analysis. Another future step for artificial intelligence models in this field would be interventional studies to examine the role of artificial intelligence in preventing intradialytic complications and in the prescription of HD.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

[1] Saran R, Robinson B, Abbott KC, et al. US renal data system 2017 annual data report: epidemiology of kidney disease in the United States. Am J Kidney Dis. 2018;71(3 Suppl 1):A7.

[2] Liyanage T, Ninomiya T, Jha V, et al. Worldwide access to treatment for end-stage kidney disease: a systematic review. Lancet. 2015;385(9981):1975–1982.

[3] Barsoum RS. Burden of end-stage kidney disease: North Africa. Clin Nephrol. 2016;86(7):14.

[4] Barsoum RS. End stage renal disease (ESKD) in Egypt and North Africa. Chron Kidney Dis Disadvantaged Popul. 2017;113–123.

[5] Barsoum RS. Chronic kidney disease in the developing world. N Engl J Med. 2006;354(10):997–999.

[6] El-Zorkany KM. Maintenance hemodialysis in Menoufia governorate, Egypt: is there any progress? J Egypt Soc Nephrol Transplant. 2017;17(2):58.

[7] Ahmed HA, Zahran AM, Issawi RA. Prevalence and etiology of end-stage renal disease patients on maintenance hemodialysis. Menoufia Med J. 2020;33(3):766.

[8] Megahed AF, Saleh RMF, Fahmy AM, et al. Hemodialysis status in an Egyptian Coastal City, Alexandria: an epidemiological study with a 3-year prospective mortality observation.

[9] Coca A, González P, Sánchez J, et al. FP639 PREVALENCE OF INTRADIALYTIC HYPOTENSION USING VARIOUS DEFINITIONS, TIME FOR CONSENSUS? Nephrol Dialysis Transplantation. 2019;34(Supplement_1):gfz106. FP639.

[10] Khajehdehi P, Mojerlou M, Behzadi S, et al. A randomized, double-blind, placebo-controlled trial of supplementary vitamins E, C and their combination for treatment of haemodialysis cramps. Nephrol Dialysis Transplantation. 2001;16(7):1448–1451.

[11] Jesus A, Oliveira HA, Paixão MOR, et al. Clinical description of hemodialysis headache in end-stage renal disease patients. Arq Neuropsiquiatr. 2009;67(4):978–981.

[12] Ramesh A, Kambhampati C, Monson JR, et al. Artificial intelligence in medicine. Ann R Coll Surg Engl. 2004;86(5):334.

[13] Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. Jama. 2017;318(6):517–518.

[14] Overley SC, Cho SK, Mehta AI, et al. Navigation and robotics in spinal surgery: where are we now? Neurosurgery. 2017;80(3S):S86–S99.

[15] Mishkind MC, Norr AM, Katz AC, et al. Review of virtual reality treatment in psychiatry: evidence versus current diffusion and use. Curr Psychiatry Rep. 2017;19(11):80.

[16] Reggia JA. Neural computation in medicine. Artif Intell Med. 1993;5(2):143–157.

[17] Geddes CC, Fox JG, Allison ME, et al. An artificial neural network can select patients at high risk of developing progressive IgA nephropathy more accurately than experienced nephrologists. Nephrol Dialysis Transplantation. 1998;13(1):67–71.

[18] Van Biesen W, Sieben G, Lameire N, et al. Application of Kohonen neural networks for the non-morphological distinction between glomerular and tubular renal disease. Nephrol Dialysis Transplantation. 1998;13(1):59–66.

[19] Akl AI, Sobh MA, Enab YM, et al. Artificial intelligence: a new approach for prescription and monitoring of hemodialysis therapy. Am J Kidney Diseases. 2001;38(6):1277–1283.

[20] Furness P, Levesley J, Luo Z, et al. A neural network approach to the biopsy diagnosis of early acute renal transplant rejection. Histopathology. 1999;35(5):461.

[21] Peng Y, Zhang Y, Wang L. Guest editorial: artificial intelligence in biomedical engineering and informatics: an introduction and review. Artif Intell Med. 2010;48(2–3):71–73.

[22] Topol EJ. A decade of digital medicine innovation. Sci Transl Med. 2019;11(498):eaaw7610.

[23] Workgroup KD. K/DOQI clinical practice guidelines for cardiovascular disease in dialysis patients. Am J Kidney Dis. 2005;45:S1–S153.

[24] Locatelli F, Cavalli A, Tucci B. The growing problem of intradialytic hypertension. Nat Rev Nephrol. 2010;6(1):41–48.

[25] Van Buren PN, Kim C, Toto R, et al. Intradialytic hypertension and the association with interdialytic ambulatory blood pressure. Clin J Am Soc Nephrol. 2011;6(7):1684–1691.
[26] Van Buren PN, Inrig JK. Mechanisms and treatment of intradialytic hypertension. Blood Purif. 2016;41(1–3):188–193.
[27] Almansour NA, Syed HF, Khayat NR, et al. Neural network and support vector machine for the prediction of chronic kidney disease: a comparative study. Comput Biol Med. 2019;109:101–111.
[28] Niel O, Bastard P. Artificial intelligence in nephrology: core concepts, clinical applications, and perspectives. Am J Kidney Diseases. 2019;74(6):803–810.
[29] Shahid N, Rappon T, Berta W. Applications of artificial neural networks in health care organizational decision-making: a scoping review. PloS One. 2019;14(2):e0212356.
[30] Agrawal RK, Khakurel S, Hada R, et al. Acute intradialytic complications in end stage renal disease on maintenance hemodialysis. J Nepal Med Assoc. 2012;52(187). DOI:10.31729/jnma.358
[31] Mehmood Y, Ghafoor S, Ashraf M, et al. Intradialytic complications found in patients at a tertiary care hospital. Austin J Pharmacol Ther. 2016;4(1):1079.
[32] Putra FR, Nursetyo AA, Thakur SS, et al. Prediction of clinical events in hemodialysis patients using an artificial neural network. Stud Health Technol Inform. 2019;264:1570–1571.
[33] Thakur SS, Abdul SS, Chiu H-YS, et al. Artificial-intelligence-based prediction of clinical events among hemodialysis patients using non-contact sensor data. Sensors. 2018;18(9):2833.
[34] 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(1):28–33.
[35] Lin C-J, Chen C-Y, Wu P-C, et al. Intelligent system to predict intradialytic hypotension in chronic hemodialysis. J Formosan Med Assoc. 2018;117(10):888–893.
[36] 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(3):396–406.
[37] Liu Y-S, Yang C-Y, Chiu P-F, et al. Machine learning analysis of time-dependent features for predicting adverse events during hemodialysis therapy: model development and validation study. J Med Internet Res. 2021;23(9):e27098.
[38] Titapiccolo JI, Ferrario M, Cerutti S, et al. Artificial intelligence models to stratify cardiovascular risk in incident hemodialysis patients. Expert Syst Appl. 2013;40(11):4679–4686.
[39] Akbilgic O, Obi Y, Potukuchi PK, et al. Machine learning to identify dialysis patients at high death risk. Kidney Int Rep. 2019;4(9):1219–1229.

Appendix

True negatives (TN): The negative labels which are correctly classified as negative.

False positives (FP): The negative labels which are incorrectly classified as positive.

True positives (TP): The positive labels which are correctly classified as positive.

False negatives (FN): The positive labels which are incorrectly classified as negative.

Classification Accuracy: For any ML model, this is one of the most intuitive and standard metrics to measure the performance if the dataset is balanced, it is defined as the rate of true prediction [1].

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{2}$$

Sensitivity (Recall): is a critical measure in medical studies because it identifies all real positive cases that have a disease. It is the number of positive classes that are predicted correctly. (e.g. the percentage of sick people who are correctly identified as having some illness) [2].

$$Sensitivity = \frac{TP}{TP + FN} \tag{3}$$

Precision: It is the ratio of the true positive label over the positive prediction [3].

$$Precision \ = \frac{TP}{FP + TP} \tag{4}$$

F1- score: It is the harmonic mean of confidence and sensitivity [2].

$$F1 - score = \frac{2 * TP}{2 * TP + FP + FN} \tag{5}$$

False Positive Rate (FPR): The number of Positive class that has been predicted wrong [1].

$$Falsepositiverate = \frac{FP}{FP + TN} \tag{6}$$

Receiver Operating Characteristic (ROC): It is one of the widely accepted evaluation criteria in medical applications. It represents true positive rate against the false-positive rate [1].

Precision-Recall curve: It is a graph that represents Precision against Recall. The best scenario that the trained model has both high precision and recall. But most machine learning algorithms often involve a trade-off between the two. A perfect PR curve has greater AUC (area under the curve) [4].

Reference

1. Thakur SS, Abdul SS, Chiu H-YS, Roy RB, Huang P-Y, Malwade S Artificial-intelligence-based prediction of clinical events among hemodialysis patients using non-contact sensor data. Sensors. 2018;18 (9):2833.

2. Mezzatesta S, Torino C, Meo PD, Fiumara G, Vilasi A. A machine learning-based approach for predicting the outbreak of cardiovascular diseases in patients on dialysis. Computer Methods and Programs in Biomedicine. 2019;177:9–15.

3. Bradley AP, Duin RPW, Paclik P, Landgrebe TCW, editors. Precision-recall operating characteristic (P-ROC) curves in imprecise environments. 18th International Conference on Pattern Recognition (ICPR'06); 2006 20–24 Aug. 2006.

4. Boyd K, Eng KH, Page CD, editors. Area under the Precision-Recall Curve: Point Estimates and Confidence Intervals. Machine Learning and Knowledge Discovery in Databases; 2013 2013//; Berlin, Heidelberg: Springer Berlin Heidelberg.