scientific reports

Check for updates

OPEN

Predicting early mortality in hemodialysis patients: a deep learning approach using a nationwide prospective cohort in South Korea

Junhyug Noh[1], Sun Young Park[2], Wonho Bae[3], Kangil Kim[4], Jang-Hee Cho[5], Jong Soo Lee[2,6], Shin-Wook Kang[7], Yong-Lim Kim[5], Yon Su Kim[8,9], Chun Soo Lim[9,10], Jung Pyo Lee[9,10,11]✉ & Kyung Don Yoo[2,6,11]✉

Early mortality after hemodialysis (HD) initiation significantly impacts the longevity of HD patients. This study aimed to quantify the effect sizes of risk factors on mortality using various machine learning approaches. A cohort of 3284 HD patients from the CRC-ESRD (2008–2014) was analyzed. Mortality risk models were validated using logistic regression, ridge regression, lasso regression, and decision trees, as well as ensemble methods like bagging and random forest. To better handle missing data and time-series variables, a recurrent neural network (RNN) with an autoencoder was also developed. Additionally, survival models predicting hazard ratios were employed using survival analysis techniques. The analysis included 1750 prevalent and 1534 incident HD patients (mean age 58.4 ± 13.6 years, 59.3% male). Over a median follow-up of 66.2 months, the overall mortality rate was 19.3%. Random forest models achieved an AUC of 0.8321 for first-year mortality prediction, which was further improved by the RNN with autoencoder (AUC 0.8357). The survival bagging model had the highest hazard ratio predictability (C-index 0.7756). A shorter dialysis duration (< 14.9 months) and high modified Charlson comorbidity index scores (7–9) were associated with hazard ratios up to 7.76 (C-index 0.7693). Comorbidities were more influential than age in predicting early mortality. Monitoring dialysis adequacy (KT/V), RAAS inhibitor use, and urine output is crucial for assessing early prognosis.

Keywords End-stage kidney disease, Deep learning, Machine learning, Hemodialysis, Survival analysis

Hemodialysis (HD) is a well-established life-saving treatment for patients with end-stage kidney disease (ESKD)[1,2]. However, mortality after HD initiation remains a challenge for improving patient outcomes and longevity. With the increasing age of patients undergoing incident dialysis, accurate assessment of the risk factors associated with mortality is crucial for timely interventions and optimal patient management[3–5]. Recent studies have investigated different aspects of elderly ESKD patient care, such as the choice between HD and peritoneal dialysis (PD)[4], access placement timing, and the effect of ESKD on life expectancy and quality of life[6]. These studies have provided valuable insights into trends and factors influencing dialysis initiation, survival rates, and

[1]Department of Artificial Intelligence, Ewha Womans University, Seoul, Republic of Korea. [2]Department of Internal Medicine, Ulsan University Hospital, University of Ulsan College of Medicine, Ulsan, Republic of Korea. [3]University of British Columbia, Vancouver, Canada. [4]School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology (GIST), Gwangju, Republic of Korea. [5]Department of Internal Medicine, Kyungpook National University College of Medicine, Daegu, Republic of Korea. [6]Basic-Clinical Convergence Research Institute, University of Ulsan, Ulsan, Republic of Korea. [7]Department of Internal Medicine, Yonsei University College of Medicine, Seoul, Republic of Korea. [8]Department of Internal Medicine, Seoul National University Hospital, Seoul, Republic of Korea. [9]Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea. [10]The Division of Nephrology, Department of Internal Medicine, Seoul National University Boramae Medical Center, Seoul, Republic of Korea. [11]Jung Pyo Lee and Kyung Don Yoo contributed equally to this work. ✉email: jungpyolee@snu.ac.kr; nephrolee@gmail.com; ykd9062@gmail.com; ykd9062@uuh.ulsan.kr

the burden of comorbidities, particularly in high-risk populations. Worldwide, the mortality risk shortly after the initiation of hemodialysis is high. By analyzing mortality patterns in 86,886 patients from 11 countries, focusing on the early dialysis period using data from the Dialysis Outcomes and Practice Patterns Study (DOPPS), a prospective cohort study of in-center hemodialysis[7] observed interactions between the dialysis period and age. In addition, mortality rate trajectories after the initiation of hemodialysis have shown early peak mortality in the US[8] and Chinese populations[9]. One possible explanation for this is the high cardiovascular event (CVE) rate during the first year after HD initiation. We assessed the weekly rates of a composite of cardiovascular events (CVEs) during the first year and the monthly rates of the composite during the first two years after hemodialysis (HD) initiation in Fresenius Medical Care dialysis centers[10]. Of the 6,308 patients who started dialysis within seven days, 1449 experienced 2405 cardiovascular events over the next two years. The first-year CVE rate (30.2/100 person-years; 95% CI, 28.7–31.7) greatly exceeded the second-year rate (19.4/100; 95% CI, 18.1–20.8)[10]. This trend is more pronounced with age; consequently, early peak mortality is more evident with age[10]. However, it is not yet clear whether this is caused by age or older people having more co-morbidities. However, there is still a need for more comprehensive and accurate models to assess and predict the risk of premature mortality in patients undergoing HD, which needs to be individualized to consider ESKD treatment patterns in individual countries, particularly HD treatment patterns[11].

Machine learning models, particularly deep learning, have shown promising results in the field of nephrology, specifically in risk prediction and outcome analysis for diverse populations with kidney disease[12,13]. These advanced computational methods may enhance our understanding of the complex relationships among various factors and their impact on the diagnosis, treatment, and prognosis[14–16]. However, these studies have limitations, as they often employ different algorithms. They failed to utilize specific information on hemodialysis patients and did not incorporate repeated measurement data. These issues highlight the need for further research to address these gaps and improve risk assessment and management of hemodialysis patients.

In this study, we aimed to investigate the risk factors for premature mortality in HD patients using a deep learning approach in a nationwide prospective cohort in Korea. By leveraging the power of deep learning models, we aimed to provide a more comprehensive and accurate understanding of risk factors, ultimately contributing to better clinical decision-making and patient care for HD patients. This study proposes a novel machine learning approach for clinical outcomes to predict early mortality in HD patients using a multicenter prospective cohort.

Materials and methods

Data source and study participants

Our analysis was performed using data from the Clinical Research Center for End-Stage Renal Disease (CRC for ESRD, NCT00931970) database, which is the sole nationwide, multicenter, and prospective cohort of Korean patients with ESRD. Data were collected from 36 general and teaching hospitals across South Korea. We included 5223 patients from the CRC for the ESRD study, of whom 3284 were undergoing hemodialysis (Fig. 1). The inclusion criteria for the study specified that all participants were 18 years or older and had initiated dialysis between August 2008 and December 2014. Detailed information on the identification and enrollment of dialysis patients into the CRC for ESRD cohort has been previously described[17,18]. Data collection was conducted using a web-based platform (http://webdb.crc-esrd.or.kr), following the established methods outlined in earlier studies. Among the initially available 72 attributes, we employed machine learning algorithms to analyze the records of 3284 patients, focusing on over 50 relevant attributes. From this pool of attributes, we selected 24 independent variables that were deemed potentially influential in predicting all-cause mortality for model construction. Notably, three of these attributes are time-series variables that were tracked at two time points, 0 and 3 months: RAAS blockade usage, 24-h urine studies reflecting urine volume (ml), and dialysis adequacy as measured by KT/V. Table 1 provides a comprehensive list of the selected attributes utilized for model development. The modified Charlson Comorbidity Index (mCCI), which has been validated for dialysis patients, was determined by examining the patients' medical histories at enrollment[19]. The application of machine learning to CRC-ESRD is also discussed in our previous study[20]. All participants were informed about the study; they voluntarily participated and provided written informed consent. The study was approved by the institutional review board of each center. All the researchers conducted this study in compliance with the guidelines of the 2008 Declaration of Helsinki. This study was approved by the Institutional Review Board of Seoul National University Hospital (IRB number H-0905-047-281).

Problem statement

We endeavor to forecast the mortality risk among patients based on the aforementioned dataset, addressing two distinct objectives:

1. Classification: Our first objective is to predict whether a patient will succumb to death within a one-year timeframe. This classification problem entails developing a prediction model that estimates the probability of a patient's mortality within one year. To assess the efficacy of the models, we will employ the AUC (Area Under the Receiver Operating Characteristic Curve) as a performance metric.
2. Survival Analysis: Our second objective involves predicting the hazard ratio, which signifies the relative risk of experiencing mortality. This survival analysis task encompasses constructing a prediction model that estimates the hazard ratio based on various factors. We will evaluate the accuracy of the prediction model using the C-index (Concordance Index), a measure of the model's ability to rank the survival times of individuals.

By addressing these two distinct challenges, we aim to enhance our understanding of mortality risk among patients undergoing hemodialysis and provide valuable insights for prognosis and clinical decision-making.

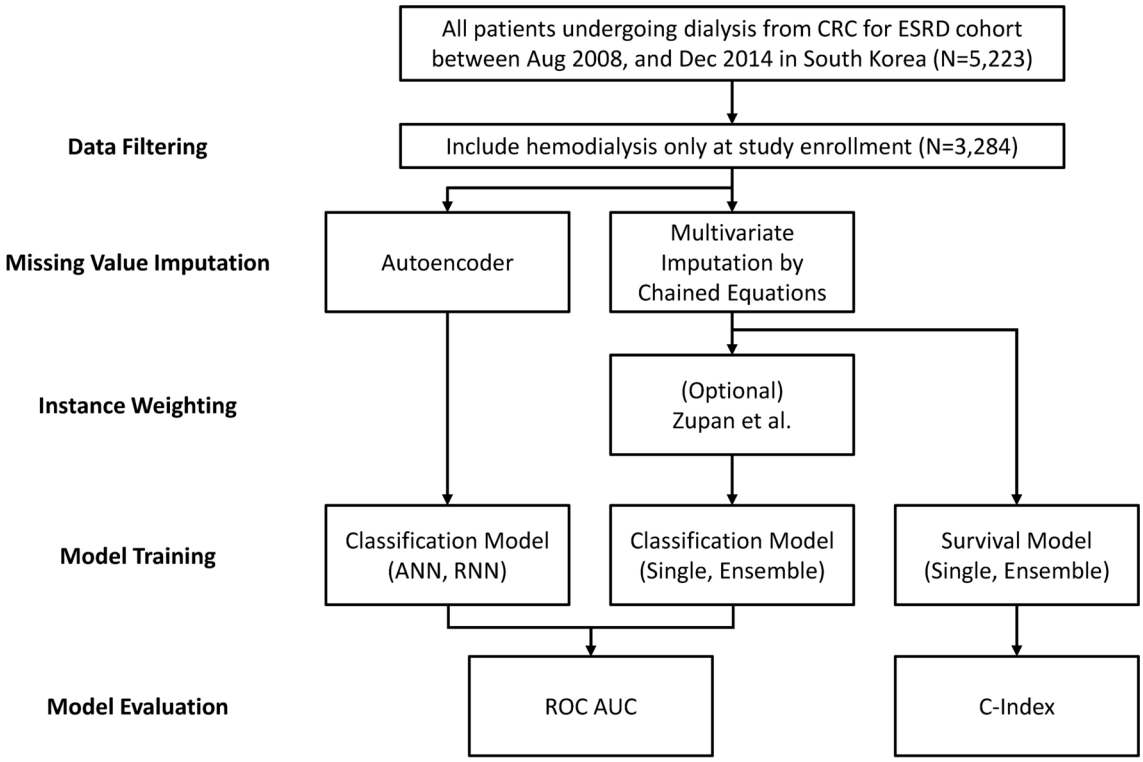

Fig. 1. Comprehensive workflow for data Preprocessing, model Training, and evaluation process. In this study, we applied the weighting methods from the study by Zupan et al.[23].

Data preprocessing

We implemented several data preprocessing steps on the given dataset, which are as follows:

1. Missing value imputation: Missing values were addressed because our data were collected over seven years, from 2008 to 2014, and it was inevitable that some values would be missing. Among the 24 variables, approximately 32.9% of patients (1,081 individuals) had at least one missing value. For patients with a small number of missing values, we performed imputation using the available variable values. We set a threshold of four or fewer missing values, resulting in 999 patients. After applying the imputation techniques, the dataset consisted of 3202 patients, accounting for 97.5% of the total dataset. The Multivariate Imputation by Chained Equations (MICE)[21] method using decision tree (CART)[22] was primarily employed for imputation, and for neural network models, an optional autoencoder approach was used. More comprehensive details on the autoencoder method will be provided in the model section.
2. Data split: To ensure the maintenance of outcome (death) proportions, we conducted a stratified random sampling when dividing the data into training, validation, and test sets. Out of the 3202 data samples, 962 samples (30%) were allocated to the test set. The remaining 2240 samples were divided into 1600 for the training set and 640 for the validation set.
3. Censored data: In the context of the classification problem, a challenge arises due to right-censored data. This means that for patients with a follow-up period of less than one year, the outcome (death) remains unknown. Within the dataset, there are 801 patients with right-censored data, comprising 264 individuals in the test set and 537 individuals in the combined train and validation sets. We excluded the 264 patients from the test set and employed either exclusion or weighting methods[23] for the training and validation sets. The weighting method duplicates each piece of right-censored data as 0 and 1, assigning weights based on the survival function. For survival analysis, the right-censored data is utilized, without impacting the total number of data samples available (Fig. 2).

Model

To address the classification and survival analysis problems, we employed a range of machine learning models. For the classification problem, we utilized single models such as logistic regression[24], ridge/lasso regression[25], and decision tree (CART)[22]. Additionally, ensemble models such as bagging[26] and random forest[27] were utilized. In the case of survival analysis, we employed corresponding models designed for survival analysis, including cox regression[28], survival ridge/lasso[25], survival decision tree[22,29], bagging[26], and survival random forest[27]. In addition to these models, we also considered the application of Artificial Neural Network (ANN)[30] and Recurrent Neural Network (RNN) for the classification problem. The RNN was particularly chosen to effectively handle the time series variables mentioned earlier: RAAS blockade usage, 24-h urine studies reflecting urine volume (ml), and dialysis adequacy measured by KT/V. These variables were concatenated with other variables at different

Variables	All participants (N = 3284)(%)	Within 1 year mortality cases (N = 183)(%)	Overall mortality cases (N = 634)(%)	Survivor (N = 2650)(%)	P*
Age (years)	58.4 ± 13.6	66.9 ± 11.4	65.9 ± 11.5	56.6 ± 13.4	< 0.001
Sex (male)	1947 (59.3)	115 (62.8)	399 (62.9)	1548 (58.4)	0.037
BMI (kg/m^2)	22.6 ± 3.4	21.7 ± 3.1	22.3 ± 3.2	22.7 ± 3.4	0.003
Primary renal disease					< 0.001
Diabetes	1565 (47.7)	103 (56.3)	380 (59.9)	1185 (44.7)	
Hypertension	546 (16.6)	25 (13.7)	100 (15.8)	446 (16.8)	
Glomerulonephritis	402 (12.2)	14 (7.7)	37 (5.8)	365 (13.8)	
Cystic kidney disease	107 (3.3)	1 (0.5)	12 (1.9)	95 (3.6)	
Unknown	153 (4.7)	9 (4.9)	21 (3.3)	132 (5.0)	
Others	511 (15.6)	31 (16.9)	84 (13.2)	427 (16.1)	
History of CVD	1010 (30.8)	88 (48.1)	303 (47.8)	707 (26.7)	< 0.001
History of DM	1728 (52.6)	110 (60.1)	409 (64.5)	1319 (49.8)	< 0.001
Smoking history	330 (10.0)	20 (10.9)	72 (11.4)	258 (9.7)	0.239
Modified CCI	5.4 ± 2.2	7.0 ± 2.3	6.9 ± 2.2	5.0 ± 2.1	< 0.001
Use of RAAS blockade	1205 (36.7)	59 (32.2)	231 (36.4)	974 (36.8)	0.891
Systolic BP (mmHg)	142 ± 22	142 ± 22	140 ± 21	142 ± 22	0.042
Diastolic BP (mmHg)	77 ± 13	75 ± 14	75 ± 13	77 ± 13	0.001
Hemoglobin (g/dL)	9.8 ± 3.2	9.3 ± 1.5	10.0 ± 4.5	9.8 ± 2.9	0.260
BUN	71.4 ± 35.4	68.8 ± 38.3	64.6 ± 32.2	73.0 ± 35.9	< 0.001
Creatinine	8.3 ± 4.4	6.9 ± 3.4	6.9 ± 3.6	8.6 ± 4.5	< 0.001
Calcium	8.3 ± 1.0	8.1 ± 0.9	8.4 ± 0.9	8.3 ± 1.1	0.007
Phosphorus	5.1 ± 2.2	4.6 ± 1.7	4.8 ± 3.3	5.2 ± 1.8	0.010
Uric acid	7.6 ± 2.2	7.3 ± 2.0	7.3 ± 1.9	7.7 ± 2.2	< 0.001
Total cholesterol	154 ± 42	149 ± 49	150 ± 41	155 ± 42	0.009
Dialysis duration (mon)	54.53 ± 55.0	30.59 ± 46.2	51.2 ± 54.4	55.3 ± 55.1	0.113

Table 1. Baseline characteristics of study participants for hemodialysis in the CRC-ESRD prospective cohort. *CVD* cardiovascular disease; *DM* diabetes mellitus; *SBP* systolic blood pressure; *DBP* diastolic blood pressure; *MCCI* modified Charlson comorbidity index; *LVH* left ventricular hypertrophy; *cTnT* cardiac troponin T; *NT pro-BNP* N-terminal pro-B-type natriuretic peptide; *RAAS* renin–angiotensin–aldosterone system. *Values are presented as n (%) for categoric variables and mean ± standard deviation for continuous variables.

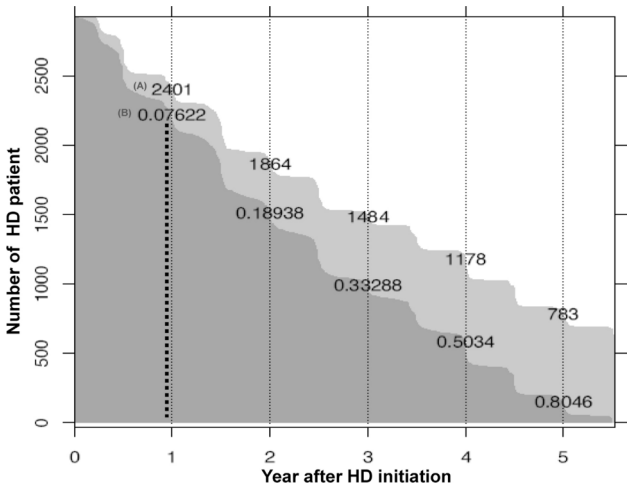

Fig. 2. Patients' follow up after HD initiation. (A) Number of patients by year of follow-up after dialysis initiation. (B) Annual survival ratio comparing non-survivors to survivors following dialysis initiation.

time steps (0/3 months) and served as inputs for each RNN cell. Among the various variants of RNN, we selected the Long Short-Term Memory (LSTM) architecture[31]. Furthermore, we incorporated an optional autoencoder (AE)[32] for the neural network models (ANN and RNN). An autoencoder is a type of neural network designed to predict input values as output values. By constraining the number of nodes in the hidden layer to be fewer than

those in the input layer, the AE can learn a compressed representation of the input data. This constraint enables efficient data representation and facilitates the utilization of the AE to handle missing values. During the training phase, some input variables were randomly removed ($p = 0.2$), and the AE was trained to reconstruct them as the original values. Figure 3 depicts the comprehensive training and inference method for the AE.

Feature importance analysis

At final, to identify the most influential predictors of patient mortality, we evaluated feature importance across five machine learning models: Logistic Regression, Ridge Regression, Lasso Regression, Decision Tree, and Random Forest. In linear models–Logistic Regression, Ridge Regression, and Lasso Regression–feature importance is determined by the absolute magnitude of the model's coefficients. Larger absolute coefficients indicate a stronger association with the target variable. Specifically, Ridge Regression employs L2 regularization, which penalizes the squared magnitude of coefficients, thereby shrinking less important feature weights towards zero without eliminating them. In contrast, Lasso Regression utilizes L1 regularization, which can force some coefficients to exactly zero, effectively performing feature selection by excluding less significant variables from the model. For tree-based models–Decision Tree and Random Forest–feature importance is calculated based on the reduction in impurity (i.e.., Gini impurity) that each feature contributes across all the splits in the trees. A feature that consistently results in a larger decrease in impurity is considered more important. Random Forest aggregates these importance scores across multiple Decision Trees within the ensemble, providing a more stable and reliable measure of feature significance compared to a single Decision Tree.

The top 10 important features identified by each model are illustrated in Figs. S1 to S5.

Implementation

The neural network models were implemented using TensorFlow version 1.13, while the remaining models were developed in R version 3.4, utilizing the glmnet, rpart, randomForestSRC, ipred, and mice packages.

Results

Demographics

The baseline characteristics utilized for modeling are shown based on mortality in HD patients (Table 1). Of the 5223 patients on dialysis in this prospective cohort, the final analysis included 3,284 patients undergoing HD. A total of 634 participants (19.3%) died during the observation period and 183 participants (5.5%) died within the first 12 months (Fig. 2). The average age was 65.9 ± 11.5 years for the overall mortality group and 56.7 ± 13.4 years for the survivor group ($p < 0.001$). The overall mortality group had 62.9% (N = 399) male patients, while the first-year mortality group had 62.8%; 59.9% of the patients began dialysis owing to diabetes.

Compared with the survivor group, the first-year mortality group had a significantly higher history of cardiovascular disease, diabetes, and increased mCCI scores. No differences between the groups regarding smoking history or the use of renin–angiotensin–aldosterone system blockade existed. Regarding laboratory

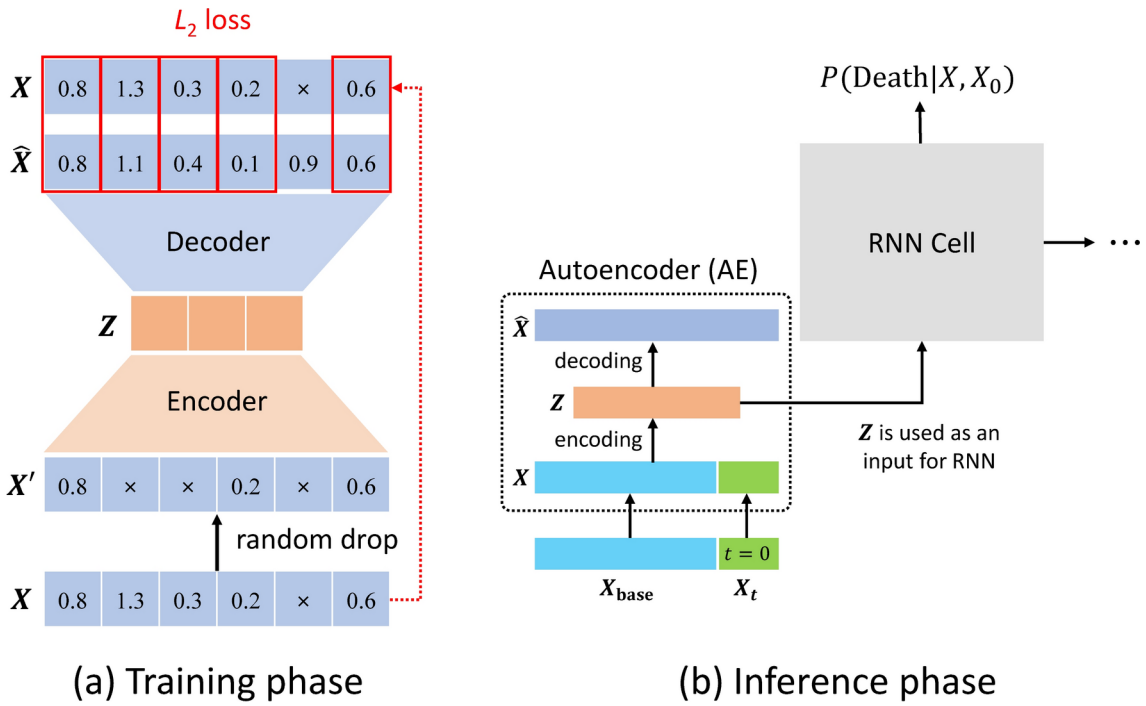

Fig. 3. Autoencoder training and inference method. During the training phase, time series variables across various time steps are merged with base variables. In the inference phase, the encoded feature (Z) derived from both base and time step variables serves as the input for the respective RNN cell.

findings at dialysis initiation, blood urea nitrogen (BUN), creatinine, and phosphorus levels were significantly lower in nonsurvivors than in survivors.

Classification of first-year mortality using traditional machine learning models

Firstly, we attempted to predict the first-year mortality of HD patients using traditional machine learning models. The ROC AUC values for the test set are presented in the lower part of Table 2. As mentioned in the data preprocessing section, all methods applied imputation for missing values via Multiple Imputation by Chained Equations (MICE)[21], and selectively applied the weighting method proposed by Zupan et al.[23] for censored examples. It was observed that tree-based ensemble models, such as random forest and bagging, performed better than single models, demonstrating the ensemble effects. Notably, the AUC for random forest was significantly higher at 0.8321 compared to 0.7571 for the single decision tree. No significant performance differences were noted with the application of the weighting method.

Classification of first-year mortality using neural networks

Next, we explored the same prediction task using neural networks, with results shown in the upper part of Table 2. The Long Short-Term Memory (LSTM) model outperformed Artificial Neural Networks (ANNs) comprised solely of fully connected layers. This indicates that LSTMs are more effective in handling time-series variables. Furthermore, in dealing with missing values, an end-to-end trained autoencoder proved to be more effective than off-the-shelf methods like MICE. Ultimately, the LSTM model utilizing an autoencoder achieved the highest AUC of 0.8357, slightly surpassing the best-performing traditional machine learning model, the random forest, which had an AUC of 0.8321.

Hazard ratio prediction

Next, we address the challenge of predicting the hazard ratio, which represents the relative risk level of mortality among HD patients. For this purpose, we adapted traditional machine learning models to facilitate survival analysis. For instance, in tree-based methods, we utilized survival analysis statistics instead of Gini or entropy indices for the splitting rules. Table 3 shows the concordance index (C-index) values for these models. Similar to the classification task, ensemble models outperformed single models, with the survival bagging model showing the highest performance, achieving a C-index of 0.7756. Figure 4 illustrates the structure of a survival decision tree (C-index 0.7693), allowing us to identify which variables and values are crucial in predicting the hazard ratio. It is noteworthy that the hazard ratio (HR) values at the leaf nodes indicate the relative HR of that group compared to the entire patient population, with higher values signifying a higher risk group. The most significant factor identified by this model was the modified Charlson comorbidity index (mCCI) for patients with a short dialysis duration of no more than 14.9 months, which was closely linked to premature mortality. Within this group, comorbidities were categorized into ≤ 6 and > 6, with the highest mortality observed across all patients, indicating a HR of 7.78, irrespective of other factors, including age.

Model	Imputation	Weighting	Hyperparameters	AUC (Val)	AUC (Test)
LSTM	Autoencoder		LSTM hnodes = 16/FC hnodes = [16]/AE hnodes = [16]	0.9351	**0.8357**
ANN	Autoencoder		hnodes = [8]/AE hnodes = [16]	0.9012	0.8160
LSTM	MICE		LSTM hnodes = 16/FC hnodes = [16]	0.9058	0.8268
ANN	MICE		hnodes = [8]	0.8927	0.8056
Random Forest	MICE		ntree = 300	0.9248	**0.8321**
Random Forest	MICE	Zupan et al.	ntree = 500	0.8894	0.8272
Bagging	MICE	Zupan et al.	nbagg = 80	0.8893	0.8109
Bagging	MICE		nbagg = 140	0.9237	0.7997
Logistic Regression	MICE			0.8861	0.7815
Logistic Regression	MICE	Zupan et al.		0.6936	0.7711
Ridge	MICE	Zupan et al.	Alpha = 0.3/lambda = 1e − 05	0.6687	0.7603
Lasso	MICE	Zupan et al.	Lambda = 1e − 04	0.6688	0.7599
Lasso	MICE		Lambda = 0.006	0.8297	0.7596
Ridge	MICE		Alpha = 0.9/lambda = 0.007	0.8291	0.7576
Decision tree	MICE		cp = − 1/maxdepth = 4	0.8776	0.7571
Decision tree	MICE	Zupan et al.	cp = − 1/maxdepth = 10	0.7551	0.7399

Table 2. Comparative performance of first-year mortality prediction models for hemodialysis in the CRC-ESRD prospective cohort. The term 'hnodes' refers to the number of nodes within the hidden layers. Names of all other hyperparameters correspond directly to those used in their respective R packages. We applied the weighting methods from the study by Zupan et al.[23]. The best performances on the test set for both deep learning models and traditional machine learning models are highlighted in bold font.

Model	Imputation	Hyperparameters	C-index (Val)	C-index (Test)
Survival Bagging	MICE	nbagg = 80	0.8430	**0.7756**
Survival Random Forest	MICE	Splitrule = logrank/ntree = 600	0.8297	0.7750
Survival Decision Tree	MICE	cp = 0.008/maxdepth = 6	0.8077	0.7693
Cox Regression	MICE		0.7716	0.7525
Survival Lasso	MICE	Lambda = 0.01	0.7784	0.7369
Survival Ridge	MICE	Alpha = 0.9/lambda = 0.04	0.7610	0.7310

Table 3. Comparative performance of hazard ratio prediction models for hemodialysis in the CRC-ESRD prospective cohort. Names of all hyperparameters correspond directly to those used in their respective R packages. The best performance on the test set is highlighted in bold font.

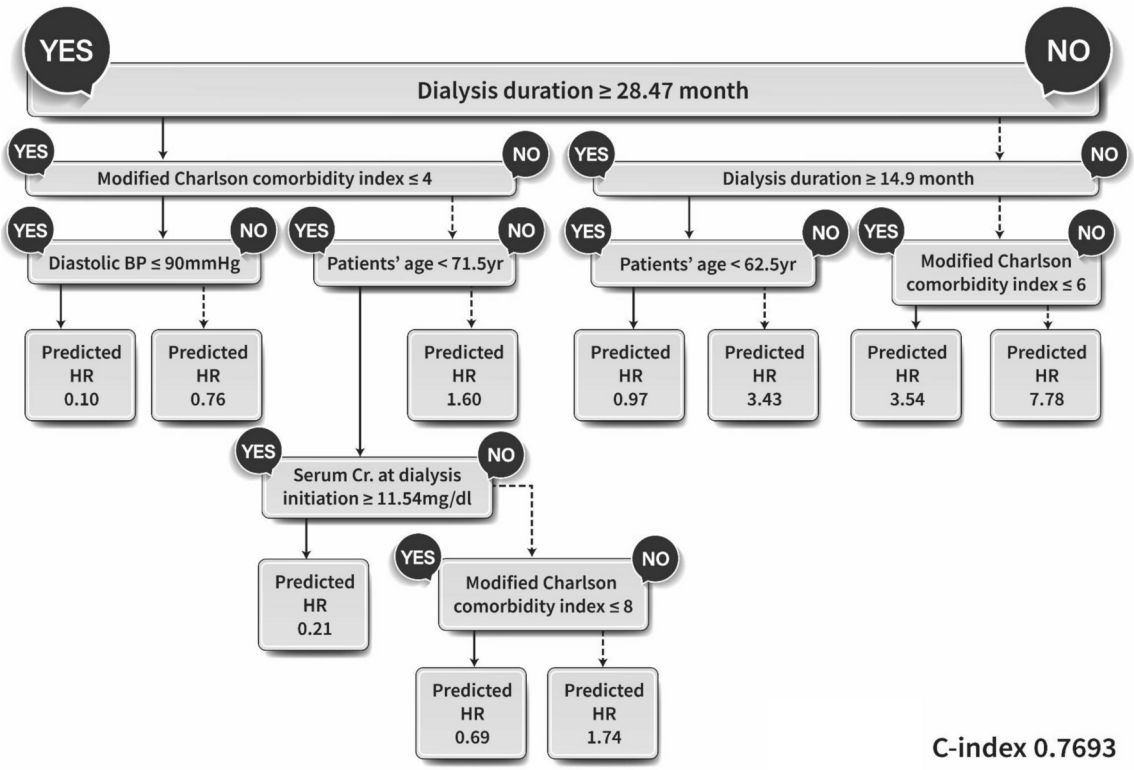

Fig. 4. Survival decision tree architecture. Each leaf node displays the relative mortality risk as a survival hazard ratio (HR). Splitting conditions are marked with "YES" for true outcomes and "NO" for false outcomes (C-index 0.7693).

Results of subgroup analysis for high-risk patients

Finally, we conducted an additional analysis of the high-risk subgroup using conventional Cox regression to identify the interconnected effects of age and mCCI risk on mortality. A multivariate analysis was performed with adjustments for confounding factors such as age, sex, primary cause of renal disease, smoking history, dialysis duration, BMI as a categorical group, history of DM, CVD, and use of RAAS blockade. As shown in Table 4, a univariate analysis revealed that age was a risk factor for overall mortality. However, the multivariate analysis showed that the age group itself did not provide a significant risk for early mortality (old age group (≥ 75 years, reference age group < 45 years), HR 2.44, 95% CI 0.89–6.68, $p = 0.082$). More importantly, the only proven prognostic factor that affected mortality in early mortality after dialysis initiation was the high mCCI group (old age group (≥ 75 years, reference age group < 45 years), HR 9.84, 95% CI 1.88–51.31, $p = 0.007$).

Discussion

In this study, we analyzed 3284 hemodialysis (HD) patients from a larger nationwide prospective cohort of 5223 individuals receiving HD. The study utilized a range of machine learning algorithms to predict the mortality within the first year of treatment for HD patients. Among these, the Recurrent Neural Networks (RNN) combined with an Autoencoder (AE) yielded the highest accuracy, achieving an Area Under the Curve (AUC)

Variables	Within 1 year mortality				Overall mortality cases			
	Unadjusted	*p*	Adjusted	*p*	Unadjusted	*p*	Adjusted	*p*
Age (≤ 44, Ref)	1		1		1		1	
45 to 59	2.08 (0.97–4.49)	0.059	1.50 (0.58–3.84)	0.395	1.82 (1.24–2.69)	0.002	1.44 (0.91–2.28)	0.113
60 to 74	4.29 (2.08–8.86)	< 0.001	1.56 (0.60–4.03)	0.358	3.97 (2.76–5.72)	< 0.001	2.05 (1.29–3.25)	0.002
≥ 75	9.71 (4.60–20.49	< 0.001	2.44 (0.89–6.68)	0.082	8.77 (5.98–12.86)	< 0.001	3.30 (2.01–5.40)	**< 0.001**
Male (Ref. female)	1.23 (0.91–1.66)	0.171	0.95 (0.65–1.38)	0.814	1.36 (1.16–1.60)	< 0.001	1.12 (0.93–1.34)	0.215
BMI (kg/m^2)								
18.5 <	0.82 (0.47–1.44)	0.507	1.08 (0.60–1.97)	0.783	0.94 (0.70–1.27)	0.711	1.43 (1.05–1.94)	**0.023**
18.5 to 22.9 (ref)	1		1		1		1	
≥ 23	0.65 (0.46–0.91)	0.012	0.48 (0.32–0.72)	**< 0.001**	0.87 (0.73–1.03)	0.124	0.75 (0.62–0.90)	**0.002**
Primary renal disease								
Diabetes (Ref.)	1		1		1		1	
Hypertension	0.68 (0.44–1.05)	0.084	0.52 (0.23–1.14)	0.105	0.76 (0.61–0.94)	0.015	0.83 (0.57–1.22)	0.355
Glomerulonephritis	0.52 (0.29–0.90)	0.022	0.71 (0.29–1.72)	0.458	0.34 (0.24–0.47)	< 0.001	0.59 (0.36–0.95)	0.032
Cystic kidney disease	0.13 (0.01–0.93)	0.043	–	–	0.37 (0.21–0.66)	0.001	0.49 (0.25–0.98)	0.044
Unknown	0.96 (0.48–1.89)	0.907	0.74 (0.27–2.04)	0.566	0.62 (0.40–0.97)	0.038	0.84 (0.48–1.48)	0.557
Others	0.89 (0.59–1.33)	0.589	1.08 (0.51–2.27)	0.829	0.66 (0.52–0.84)	0.001	0.91 (0.62–1.35)	0.669
CVD (Ref. No CVD)	2.00 (1.50–2.67)	< 0.001	1.76 (1.22–2.53)	**0.002**	2.01 (1.72–2.35)	< 0.001	1.45 (1.21–1.73)	**< 0.001**
DM (Ref. No CVD)	1.35 (1.00–1.82)	0.044	0.52 (0.26–1.03)	0.063	1.69 (1.44–1.99)	< 0.001	0.79 (0.55–1.13)	0.212
Smoking (Ref. No smoking)	1.13 (0.71–1.80)	0.596	1.46 (0.84–2.53)	0.176	1.28 (1.00–1.64)	0.045	1.36 (1.03–1.79)	0.027
Modified CCI (Ref. 0–2	1		1		1		1	
3–5	3.45 (1.07–11.16)	0.038	3.27 (0.070–15.24)	0.176	2.57 (1.51–4.38)	< 0.001	1.53 (0.82–2.86)	0.175
6–8	8.58 (2.71–27.10)	< 0.001	6.16 (1.25–30.31)	0.025	6.46 (3.85–10.85)	< 0.001	2.57 (1.33–4.95)	**0.005**
≥ 9	16.05 (4.98–51.73)	< 0.001	9.84 (1.88–51.31)	**0.007**	11.92 (7.00–20.29)	< 0.001	3.81 (1.90–7.61)	**< 0.001**
Use of RAAS blockade (Ref. No use)	0.68 (0.50–0.93)	0.018	1.09 (0.77–1.56)	0.602	0.79 (0.67–0.93)	0.006	0.90 (0.76–1.08)	0.274
Dialysis duration (per month)	0.97 (0.97–0.98)	< 0.001	0.97 (0.96–0.98)	< 0.001	0.98 (0.98–0.98)	< 0.001	0.98 (0.98–0.98)	< 0.001

Table 4. Risk factors analysis for first-year and overall mortality in the CRC-ESRD prospective cohort. CVD, cardiovascular disease; DM, diabetes mellitus; SBP, systolic blood pressure; DBP, diastolic blood pressure; MCCI, modified Charlson comorbidity index; LVH, left ventricular hypertrophy; cTnT, cardiac troponin T; NT pro-BNP, N-terminal pro-B-type natriuretic peptide; RAAS, renin–angiotensin–aldosterone system. *Values are presented as n (%) for categoric variables and mean ± standard deviation for continuous variables. mCCI group; Low, mCCI score 0–2; Moderate, mCCI score 3–5; High, mCCI score ≥ 6. Multivariate analysis was done with adjustment confounding including such as age, sex, primary renal disease, smoking history, dialysis duration, BUN, systolic BP, BMI, Hb, Calcium, history of DM, CVD, usage of RAAS blockade, and serum albumin. Significant values are in bold.

of 0.8357. Additionally, for predicting the hazard ratio, the survival bagging model emerged as the top performer with a Concordance Index (C-index) of 0.7756.

Investigation into forecasting early prognosis among ESKD patients undergoing dialysis has a long-standing history in this population. In 2009, Couchoud et al.[33] predicted the 6-month mortality rate of elderly patients initiating dialysis based solely on clinical features using French Renal Epidemiology and Information Network (REIN) cohort data. Despite selecting nine risk factors and having a good validation of the observed and expected deaths, the study only achieved a moderate C-index of 0.70. One limitation was the potential selection bias caused by missing imputations. By contrast, our study applied the same model with the same variables but used the MICE imputation method for augmentation. As a result, our study achieved an increased C-index of 0.7756 and revealed that the mCCI had a more significant impact solely on mortality than age. In a systematic review by Sanmarchi et al.[13] AI and ML techniques showed promise in predicting, diagnosing, and treating chronic kidney disease, including ESKD; however, future work is required to enhance the interpretability, generalizability, and fairness of machine learning models for clinical practice. It is noteworthy that only five studies utilized RNNs and only one study used an autoencoder for data augmentation. Based on our research findings, we anticipate improved results through the use of imputation with hemodialysis-specific information in future studies[12].

Our study results show that comorbidities are more closely related to prognosis regardless of age, and these findings remain consistent even after adjusting for dialysis efficiency, residual renal function presenting with urine volume, and RAAS blockade use. Many researchers suspect that a higher number of comorbidities could be more strongly associated with a poor prognosis in dialysis patients. Liu et al. published an excellent study on the impact of comorbidities on survival prediction in elderly dialysis patients, which modified the existing Charlson Comorbidity Index (CCI)[34] and developed a new comorbidity index (nCI) for mortality analysis in dialysis patients based on comorbid conditions used in the United States Renal Data System (USRDS) and

administrative data[35]. The nCI was developed using 2000 US incident dialysis patients and validated using the 1999 and 2001 incident US dialysis populations and the 2000 prevalent US dialysis population. Interestingly, Liu et al.[35] comorbidity index included 11 comorbidities (atherosclerotic heart disease, congestive heart failure, cerebrovascular accident/transient ischemic attack, peripheral vascular disease, arrhythmia, other heart diseases, chronic obstructive pulmonary disease, gastrointestinal bleeding, liver disease, cancer, and diabetes) but did not include age, which is a component of the original CCI. The authors demonstrated that the performance of nCI was almost identical to that of individual comorbid conditions in terms of model fit, predictive ability, and influence on inference, indicating that nCI is a better predictor than CCI. Accurate prediction requires the integration of age and comorbidities, which can be challenging. However, our study supports the findings that in high-risk patients, such as dialysis patients, comorbidities may be more critical than age.

Cohen et al. also developed a prognostic model to assess the risk of death in dialysis patients by combining selected variables, such as mCCI and serum albumin, with nephrologists' responses to a surprise question ("Would you be surprised if this patient died within six months?")[36]. This simple bedside tool for predicting six-month mortality showed an independent superior prognostic value compared with either of the two tools alone. However, one could argue that this model has limitations, as it relies on subjective parameters (for example, dementia) that are difficult to define in dialysis patients. Moreover, the surprise question may be highly subjective and variable depending on the nephrologist's training and patient knowledge. Our study has some drawbacks, such as the young average age of the patients (58.4 years) included in the prospective cohort and a lack of information on cognitive impairments. Hence, Floege et al. presented another risk prediction model developed in a European hemodialysis cohort with a mean age of 64 years using only objective measurements without information on dementia[37]. This model underwent external validation in the DOPPS cohort and demonstrated moderate discrimination (C-statistics 0.68–0.79). The similarity between this cohort and our study population has potential implications for our research.

Research introducing ML for prognostic prediction in patients has increased to overcome the limitations of previous studies that used traditional statistical methods[13]. One strength of our study is that we demonstrated improved performance by validating almost all ML algorithms for dialysis patients introduced thus far in a step-by-step manner. The AUROC for traditional statistical mortality prediction models typically falls in the range of 0.65 to 0.75[38]. The use of a sophisticated machine learning technique using electronic medical records (EMR), specifically random forest models, to forecast sudden cardiac death in elderly patients undergoing HD was an initial example of this approach. The model produced an AUROC of up to 0.79[39,40], which is comparable to our study results (Table 2). The predictive power improved with the use the RNN (AUC 0.8357), and similar research was not found in the authors' search or this systematic review. As mentioned earlier, the first reason is likely the increased imputation validity using an autoencoder[13], and the second is the simultaneous consideration of additional variable selection for the RNN, which greatly impacts prognostic prediction. In addition, dialysis adequacy as KT/V, which has been confirmed in many previous studies, including 24-h urine studies as urine volume[41] and RAAS blockade usage[17] as independent prognostic factors in this prospective cohort, likely contributed to the improved prediction. Thus, clinical applicability must be considered when applying and interpreting deep learning approaches.

Conclusion

This study utilized machine learning algorithms to predict early mortality and assess risk factors in hemodialysis patients using a nationwide prospective cohort in South Korea. The recurrent neural network with an autoencoder yielded the highest accuracy for predicting first-year mortality (AUC 0.8357), while the survival bagging model performed best in predicting hazard ratios (C-index 0.7756). Notably, comorbidities were found to be more influential than age in predicting early mortality, with high modified Charlson comorbidity index scores (≥ 7) being a significant prognostic factor, especially for patients with shorter dialysis duration (< 14.9 months). The study highlights the importance of monitoring factors such as dialysis adequacy (KT/V), renin–angiotensin–aldosterone system inhibitor use, and residual urine output for assessing early prognosis in hemodialysis patients. These findings contribute to a better understanding of risk factors and can aid in clinical decision-making and patient care for hemodialysis patients.

Data availability

To obtain access to the data, interested individuals should contact the corresponding author and submit a reasonable request explaining their need for the data. Upon receiving such a request, the corresponding author will determine whether sharing the data is appropriate and feasible.

Received: 12 April 2024; Accepted: 22 November 2024
Published online: 29 November 2024

References

1. Kovesdy, C. P. Epidemiology of chronic kidney disease: An update 2022. *Kidney Int. Suppl.* **12**(1), 7–11 (2022).
2. Suriyong, P., Ruengorn, C., Shayakul, C., Anantachoti, P. & Kanjanarat, P. Prevalence of chronic kidney disease stages 3–5 in low- and middle-income countries in Asia: A systematic review and meta-analysis. *PLoS ONE* **17**(2), e0264393 (2022).
3. Obi, Y. et al. Development and validation of prediction scores for early mortality at transition to dialysis. *Mayo Clin. Proc.* **93**(9), 1224–1235 (2018).
4. Thamer, M. et al. Predicting early death among elderly dialysis patients: Development and validation of a risk score to assist shared decision making for dialysis initiation. *Am. J. Kidney Dis.* **66**(6), 1024–1032 (2015).
5. Couchoud, C. G., Beuscart, J. B., Aldigier, J. C., Brunet, P. J. & Moranne, O. P. Development of a risk stratification algorithm to improve patient-centered care and decision making for incident elderly patients with end-stage renal disease. *Kidney Int.* **88**(5), 1178–1186 (2015).

6. Moss, A. H. Revised dialysis clinical practice guideline promotes more informed decision-making. *Clin. J. Am. Soc. Nephrol.* **5**(12), 2380–2383 (2010).
7. Robinson, B. M. et al. Worldwide, mortality risk is high soon after initiation of hemodialysis. *Kidney Int.* **85**(1), 158–165 (2014).
8. Foley, R. N., Chen, S. C., Solid, C. A., Gilbertson, D. T. & Collins, A. J. Early mortality in patients starting dialysis appears to go unregistered. *Kidney Int.* **86**(2), 392–398 (2014).
9. Zhao, X., Wang, M. & Zuo, L. Early mortality risk in incident Chinese hemodialysis patients: A retrospective cohort study. *Ren. Fail.* **39**(1), 526–532 (2017).
10. Eckardt, K. U. et al. High cardiovascular event rates occur within the first weeks of starting hemodialysis. *Kidney Int.* **88**(5), 1117–1125 (2015).
11. Jung, J. Y. et al. Executive summary of the Korean society of nephrology 2021 clinical practice guideline for optimal hemodialysis treatment. *Kidney Res. Clin. Pract.* **40**(4), 578–595 (2021).
12. Chaudhuri, S. et al. Artificial intelligence enabled applications in kidney disease. *Semin. Dial.* **34**(1), 5–16 (2021).
13. Sanmarchi, F. et al. Predict, diagnose, and treat chronic kidney disease with machine learning: A systematic literature review. *J. Nephrol.* **36**, 1101–1117 (2023).
14. Garcia-Montemayor, V. et al. Predicting mortality in hemodialysis patients using machine learning analysis. *Clin. Kidney J.* **14**(5), 1388–1395 (2021).
15. Rankin, S. et al. A machine learning model for predicting mortality within 90 days of dialysis initiation. *Kidney 360.* **3**(9), 1556–1565 (2022).
16. Yang, C. H. et al. Machine learning approaches for the mortality risk assessment of patients undergoing hemodialysis. *Ther. Adv. Chronic Disease.* **13**, 20406223221119616 (2022).
17. Yoo, K. D. et al. Effect of renin-angiotensin-aldosterone system blockade on outcomes in patients with ESRD: A prospective cohort study in Korea. *Kidney Int. Rep.* **3**(6), 1385–1393 (2018).
18. Yoo, K. D. et al. Cancer in Korean patients with end-stage renal disease: A 7-year follow-up. *PLoS One.* **12**(7), e0178649 (2017).
19. Park, J. Y. et al. Recalibration and validation of the charlson comorbidity index in Korean incident hemodialysis patients. *PLoS One.* **10**(5), e0127240 (2015).
20. Noh, J. et al. Prediction of the mortality risk in peritoneal dialysis patients using machine learning models: A nation-wide prospective cohort in Korea. *Sci. Rep.* **10**(1), 7470 (2020).
21. Buuren, S. & Groothuis-Oudshoorn, K. Mice: Multivariate imputation by chained equations in R. *JSS.* **45**, 1–67 (2011).
22. Breiman, L., Friedman, J., Stone, C. J. & Olshen, R. A. *Classification and Regression Trees* (CRC Press, 1984).
23. Zupan, B. D. J., Kattan, M. W., Beck, J. R. & Bratko, I. Machine learning for survival analysis: A case study on recurrence of prostate cancer. *Artif. Intell. Med.* **20**, 59–75 (2000).
24. Dobson, A. J. An introduction to generalized linear models. *J. Stat. Plann. Inference.* **32**(3), 418–420 (1992).
25. Friedman, J. H., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Statistical Softw.* **33**(1), 1–22 (2010).
26. Breiman, L. Bagging predictors. *Mach. Learn.* **24**, 123–140 (1996).
27. Breiman L. Random forests. *Mach. Learn.* (2001).
28. Andersen, P. K. & Gill, R. D. Cox's regression model for counting processes: A large sample study. *Ann. Stat.* **10**(4), 1100–1121 (1982).
29. LeBlanc, M. & Crowley, J. Relative risk trees for censored survival data. *Biometrics* 411–425, (1992).
30. Jain, A. K., Mao, J. & Mohiuddin, K. M. Artificial neural networks: A tutorial. *Computer.* **29**, 31–44 (1996).
31. Hochreiter, S. & Schmidhuber, J. long short-term memory. *Neural Comput.* **9**(8), 1735–1780 (1997).
32. Kramer, M. A. Nonlinear principal component analysis using autoassociative neural networks. *AIChE J.* **37**(2), 233–243 (1991).
33. Couchoud, C. et al. A clinical score to predict 6-month prognosis in elderly patients starting dialysis for end-stage renal disease. *Nephrol. Dial Transpl.* **24**(5), 1553–1561 (2009).
34. Charlson, M. E., Pompei, P., Ales, K. L. & MacKenzie, C. R. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *J. Chronic Diseases.* **40**(5), 373–383 (1987).
35. Liu, J., Huang, Z., Gilbertson, D. T., Foley, R. N. & Collins, A. J. An improved comorbidity index for outcome analyses among dialysis patients. *Kidney Int.* **77**(2), 141–151 (2010).
36. Cohen, L. M., Ruthazer, R., Moss, A. H. & Germain, M. J. Predicting six-month mortality for patients who are on maintenance hemodialysis. *Clin. J. Am. Soc. Nephrol.* **5**(1), 72–79 (2010).
37. Floege, J. et al. Development and validation of a predictive mortality risk score from a European hemodialysis cohort. *Kidney Int.* **87**(5), 996–1008 (2015).
38. Ramspek, C. L. et al. Prediction models for the mortality risk in chronic dialysis patients: A systematic review and independent external validation study. *Clin. Epidemiol.* **9**, 451–464 (2017).
39. Goldstein, B. A., Chang, T. I., Mitani, A. A., Assimes, T. L. & Winkelmayer, W. C. Near-term prediction of sudden cardiac death in older hemodialysis patients using electronic health records. *Clin. J. Am. Soc. Nephrol.* **9**(1), 82–91 (2014).
40. Akbilgic, O. et al. Machine learning to identify dialysis patients at high death risk. *Kidney Int. Rep.* **4**(9), 1219–1229 (2019).
41. Lee, M. J. et al. Prognostic value of residual urine volume, GFR by 24-hour urine collection, and eGFR in patients receiving dialysis. *Clin. J. Am. Soc. Nephrol.* **12**(3), 426–434 (2017).

Acknowledgements

We express our gratitude to all the Clinical Research Center for End-Stage Renal Disease investigators. This study was funded by grant 03-2020-2130 from the Seoul National University Hospital Research Fund and supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (Grant Numbers: HR22C1832 and HI22C1529). Junhyug Noh was supported by Ewha Womans University research grant of 2023 and the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korean government (MSIT) (No. RS-2022-00155966). It is worthwhile to mention that in 2020, Kyung Don Yoo was honored with the Outstanding Abstract Award by the Korean Society of Nephrology (KSN) for the abstract content presented herein.

Author contributions

K.D.Y., S.Y.P. and J.N. analyzed the data, drafted and revised the paper; K.D.Y., J.N., W.B., and K.K. revised the paper; J.H.C., J.S.L., S.W.K., Y.L.K., Y.S.K., and C.S.L. collected the data; W.B. and K.I.K. provided critical comments on the method; and J.N., K.D.Y., W.B., K.I.K., G.K., and J.P.L. conceived the study, participated in its design and coordination, and helped draft the manuscript. All the authors approved the final version of the manuscript.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-80900-6.

Correspondence and requests for materials should be addressed to J.P.L. or K.D.Y.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.