Journal of Multidisciplinary Healthcare

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Ensemble Machine Learning for Predicting 90-Day Outcomes and Analyzing Risk Factors in Acute Kidney Injury Requiring Dialysis

Tzu-Hao Wang[1,2], Chih-Chin Kao[3–5,*], Tzu-Hao Chang[2,6,*]

[1]Division of General Medicine, Department of Medical Education, Shuang-Ho Hospital, Taipei Medical University, New Taipei City, Taiwan, Republic of China; [2]Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, Taipei, Taiwan, Republic of China; [3]Division of Nephrology, Department of Internal Medicine, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan, Republic of China; [4]Division of Nephrology, Department of Internal Medicine, Taipei Medical University Hospital, Taipei, Taiwan, Republic of China; [5]Taipei Medical University-Research Center of Urology and Kidney (TMU-RCUK), Taipei Medical University, Taipei, Taiwan, Republic of China; [6]Clinical Big Data Research Center, Taipei Medical University Hospital, Taipei City, Taiwan, Republic of China

*These authors contributed equally to this work

Correspondence: Tzu-Hao Chang, Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, International Center for Health Information Technology, Taipei Medical University, 9F, Education & Research Building, Shuang-Ho Campus, No. 301, Yuantong Road, Zhonghe District, New Taipei City, 235603, Taiwan, Republic of China, Tel +886-66202589 ext.10922, Email kevinchang@tmu.edu.tw

Purpose: Our objectives were to (1) employ ensemble machine learning algorithms utilizing real-world clinical data to predict 90-day prognosis, including dialysis dependence and mortality, following the first hospitalized dialysis and (2) identify the significant factors associated with overall outcomes.

Patients and Methods: We identified hospitalized patients with Acute kidney injury requiring dialysis (AKI-D) from a dataset of the Taipei Medical University Clinical Research Database (TMUCRD) from January 2008 to December 2020. The extracted data comprise demographics, comorbidities, medications, and laboratory parameters. Ensemble machine learning models were developed utilizing real-world clinical data through the Google Cloud Platform.

Results: The Study Analyzed 1080 Patients in the Dialysis-Dependent Module, Out of Which 616 Received Regular Dialysis After 90 Days. Our Ensemble Model, Consisting of 25 Feedforward Neural Network Models, Demonstrated the Best Performance with an Auroc of 0.846. We Identified the Baseline Creatinine Value, Assessed at Least 90 Days Before the Initial Dialysis, as the Most Crucial Factor. We selected 2358 patients, 984 of whom were deceased after 90 days, for the survival module. The ensemble model, comprising 15 feedforward neural network models and 10 gradient-boosted decision tree models, achieved superior performance with an AUROC of 0.865. The pre-dialysis creatinine value, tested within 90 days prior to the initial dialysis, was identified as the most significant factor.

Conclusion: Ensemble machine learning models outperform logistic regression models in predicting outcomes of AKI-D, compared to existing literature. Our study, which includes a large sample size from three different hospitals, supports the significance of the creatinine value tested before the first hospitalized dialysis in determining overall prognosis. Healthcare providers could benefit from utilizing our validated prediction model to improve clinical decision-making and enhance patient care for the high-risk population.

Keywords: AKI-D, dialysis prognosis, ensemble machine learning, prediction models, risk factors

Introduction

Acute kidney injury (AKI) is a frequently encountered complication in hospitalized patients with acute illness, and its incidence is on the rise.[1–3] Research has demonstrated that AKI occurs in approximately 20% of adults and 33% of children hospitalized due to acute illness.[2] The clinical importance of AKI is emphasized by consistent evidence linking it to long-term risks, including heightened mortality, the development of chronic kidney disease requiring renal replacement therapy (RRT), decreased health-related quality of life, and increased utilization of healthcare resources.[4]

Received: 7 December 2023
Accepted: 24 March 2024
Published: 12 April 2024

Acute kidney injury requiring dialysis (AKI-D) is linked with greater morbidity and mortality, both during and after hospitalization.[5–9] The recovery of kidney function has a poor prognosis,[8,10–14] especially for patients with comorbidities.[10,15–18] A nationwide study conducted in the US revealed a yearly spike of 10% in the occurrence of AKI-D.[19] Additionally, up to 30% of AKI-D patients might necessitate regular hemodialysis in the outpatient phase following discharge from the hospital.[17,20–24] Mortality figures for AKI-D patients have been documented at different time intervals, with rates reported at 35%, 45%, and 49% for in-hospital, 90-day, and 6-month periods, respectively.[25,26] These findings highlight the importance of increasing awareness regarding the detrimental effects and substantial economic burden that AKI-D imposes on patients, communities, and healthcare systems.

Machine learning techniques offer enhanced predictions of clinical performance and serve as valuable alternatives to traditional multivariate regression techniques in clinical medical studies.[27] Recent publications indicate an increasing use of machine learning in predicting AKI. Mohamadlou et al[28] found that a machine-learning-based AKI prediction tool could provide crucial prognostic capabilities for identifying patients at risk of AKI, enabling clinicians to intervene before renal damage becomes apparent. Similarly, Shawwa et al[29] demonstrated that machine learning can predict AKI in intensive care unit (ICU) patients using pre-admission information.

Following a comprehensive literature review, we have identified a limited number of studies that concentrate on predicting 90-day dialysis dependence in AKI-D hospitalized patients. Furthermore, we have observed that most of these studies utilize logistic regression models for outcome prediction with relatively small sample sizes, as demonstrated in Supplementary 1.

To address this issue, our study aims to employ machine learning algorithms on real-world clinical data from three different hospitals, taking into account multiple attributes, to predict the 90-day prognosis of patients who experience their first episode of AKI-D during hospitalization. Furthermore, we will identify the significant factors associated with these outcomes.

To evaluate our prediction model, we used the 5-fold cross-validation method. Firstly, we split the dataset into two subsets: a training set for development and a testing set for external validation. Secondly, the training set is split into a new training fold and a validation fold for internal validation to complete the cross-validation process.

By following these principles of objectivity, our goal is to develop a prediction model with great accuracy for the prognosis of first hospitalized AKI-D patients. And further make a valuable contribution to future clinical practice and decision-making processes that ultimately benefit both health care providers and patients in their treatment and care.

Materials and Methods

Study Design and Data Source

Study Design

For our case-control analysis, we employed the Taipei Medical University Clinical Research Database (TMUCRD), which covered the period from 2008 to 2020. The primary objective was to collect data on patients who experienced AKI-D during their first hospitalization and the related clinical information. The patients were subsequently divided into two groups based on their prognosis.

Following this, we constructed multiple prediction models with the goal of accurately predicting the recovery of renal function to dialysis dependence 90 days after the initial hospitalized dialysis. This time frame is typically considered the threshold for determining whether patients have reached end-stage renal disease (ESRD).[30] In addition, we constructed alternative prediction models to evaluate the mortality risk in the same designated patient population. After evaluating the performance of the models, we selected the one with the highest predictive accuracy. The top 10 significant features were extracted from this selected model based on the results.

Data processing for this study was produced utilizing SAS software, Version 9.4 of the SAS System for Windows. Model training and testing were conducted using Vertex AI from Google Cloud Platform.

Data Source

For this study, patient data was obtained from the TMUCRD, a comprehensive database that aggregates information from electronic medical records across three hospitals: Taipei Medical University Hospital (TMUH), Wan-Fang Hospital (WFH), and Shuang-Ho Hospital (SHH). This database contains a large amount of electronic medical record data from

2008 to 2020, with 3.8 million individual records. It includes structured data such as patient demographics, medical details, test reports, diagnosis results, treatment history, surgical records, and medication usage, as well as unstructured data like narrative progress notes, pathology reports, and medical imaging reports.

The objective was to gather data on patients who experienced AKI-D during their first hospitalization but had varying prognosis after 90 days. The index time began when the patient underwent their first hospitalized dialysis. After 90 days of the treatment regardless of the number of dialysis sessions during this period, we documented the patient's outcome based on a 30-day follow-up period. (See Supplementary 2 and Supplementary 3)

To ensure ethical compliance, the Joint Institute Review Board of Taipei Medical University in Taipei, Taiwan approved this case-control study on February 7, 2023 (approved number, N201703004). Before any analyses, all patient data were anonymized.

Patient and Feature Selection

The inclusion criteria primarily targeted patients who experienced the first episode of AKI-D during hospitalization. The patients were selected from the TMUCRD database spanning from 2008 to 2020, which includes three hospitals located in the northern part of Taiwan. Exclusion criteria were implemented to exclude those with previous outpatient dialysis records and those with underlying ESRD as suggested by relevant ICD codes (ICD-9: 585.6; ICD-10: N18.6).

To build prediction models, features were meticulously selected based on literature reviews and consultations with clinicians. The selected features were hypothesized to potentially impact dialysis dependence and mortality in first hospitalized AKI-D patients, especially those related to the renal function. We also aimed to include as many clinical features as possible to enhance the attributes of the machine learning algorithm. The selected features included:

1. Demographic information: This included age, gender, height, weight, body mass index (BMI), and admission to the ICU.
2. Comorbidities: Prior documented medical diseases before the index date of the first hospitalized dialysis were classified using the Charlson Comorbidity Index.
3. Medications: Medication records from the year preceding the index date were gathered and categorized using the anatomical therapeutic chemical classification system.
4. Laboratory tests: Laboratory parameters measured were blood urea nitrogen (BUN), creatinine, estimated glomerular filtration rate (eGFR), albumin, glucose, hemoglobin A1c, thyroid stimulating hormone, uric acid, total cholesterol, high-density lipoprotein, low-density lipoprotein, and triglycerides. C-reactive protein (CRP), D-dimer, troponin-I, creatine kinase-MB, creatine phosphokinase (CPK), N-terminal pro-brain natriuretic peptide, hemoglobin, urine microalbumin, urine protein, urine glucose, urine microalbumin-creatinine ratio, and urine protein-creatinine ratio were all measured.

Of these parameters, BUN, creatinine, and eGFR were clinically recognized as significant tests for evaluating renal function.[31] Thus, we ensured that each patient in our dataset had complete medical records of creatinine, BUN, and eGFR at all time points. Patients without these records were excluded from our analysis, as were other features with a missing rate over 30%. For the continuous variables with missing data under 30%, missing values were imputed by replacing them with the mean.

Additionally, laboratory tests were selected at three specific time points each: (1) Baseline: 90 or more days prior to the index date, (2) Pre-dialysis: within 90 days before the index date, and (3) Post-dialysis: within 2–8 weeks after the index date. Furthermore, the difference between two time points of each feature was used as an input. All three time points for each parameter were collected for the dialysis-dependent module: Baseline, Pre-dialysis, and Post-dialysis. After the initial dialysis, renal function can be unstable, making it crucial to monitor post-dialysis laboratory data in order to accurately predict the patient's long-term outcome. For the survival module, we gathered data at two time points for each parameter: Baseline and Pre-dialysis. By doing so, we can more promptly identify survival rates post-initial dialysis, particularly in critical and emergent cases.

Prediction Model Construction

To predict both dialysis dependence and mortality 90 days after the first hospitalized dialysis, we utilized Vertex AI on the Google Cloud Platform to train and deploy our machine learning models. Various models were trained, utilizing different machine learning algorithms with varying sets of hyperparameters through the Grid Search approach. Based on the area under the receiver operating characteristic curve (AUROC) value, we created an optimal ensemble model by merging the top-performing models, such as gradient boosted decision trees (GBDT) and feedforward neural networks

Gradient Boosted Decision Trees

GBDT is a highly effective machine learning algorithm utilized for both regression and classification tasks, acknowledged for its efficiency, accuracy, and interpretability. It operates as an ensemble method, combining numerous decision trees trained sequentially. In each iteration, GBDT fits the negative gradient to learn from previous iterations' errors. By consolidating the outcomes of each tree, GBDT constructs a durable and compelling model.[32]

Feedforward Neural Network

Feedforward neural network is a fundamental type of neural network architecture in which information flows in only one direction, from the input layer through one or more hidden layers to the output layers. Implementing a neural network ensemble can substantially enhance the system's generalization ability. This ensemble can capture diverse representations and different aspects of the data, resulting in better predictive capabilities. The ensemble approach utilizes multiple networks' strengths and diversity to generate more precise and robust predictions.[33]

Performance Evaluation

To evaluate our prediction model, we applied the 5-fold cross-validation method (see Supplementary 4). Firstly, we split the dataset into two subsets: a training set (80% of the patients) for development and a testing set (20% of the patients) for external validation. Secondly, we split the training set into a new training fold and a validation fold. This division served as an internal validation to complete the cross-validation process, which was repeated five times. The training fold was utilized to train the machine learning models, whereas the validation fold was employed to assess the performance of various machine learning models and optimize their hyperparameters. The model's performance on real-world data was evaluated using the testing set to provide an unbiased estimate of its predictive capabilities.

Performance of the machine learning algorithms was measured using metrics such as AUROC, accuracy, sensitivity, specificity, precision, negative predictive value, and F1-score. A combination of these metrics provides a more comprehensive understanding of the model's performance in various aspects such as discrimination ability, overall correctness, and trade-offs between precision and sensitivity: AUROC measures the ability of a model to distinguish between binary classes across various thresholds. The curve plots the true positive rate against the false positive rate, providing a comprehensive view of a model's discriminatory power. Accuracy is a general metric that measures the overall correctness of the model's predictions. Sensitivity measures the proportion of actual positive instances correctly predicted by the model. In a medical diagnosis task, high sensitivity is crucial to avoid missing positive cases. Specificity measures the proportion of actual negative instances correctly predicted by the model. It is crucial when the cost of false positives is high. Precision measures the proportion of predicted positive instances that are actually positive, helping to assess the accuracy of positive predictions. Negative Predictive Value measures the proportion of predicted negative instances that are actually negative, which is relevant in scenarios where the cost of false negatives is significant. The F1-Score is the harmonic mean of precision and recall. It is useful in situations where both false positives and false negatives are critical, as it provides a balanced measure when there is an imbalance between classes.

Furthermore, we examined the feature importance using Shapley additive explanations (SHAP)[34,35] values to understand each feature's contribution to the model's predictions. The SHAP method utilizes coalitional game theory to compute Shapley values. In this method, the feature values of a data instance act as players in a coalition. The importance of features is determined by their absolute Shapley values, with larger values indicating greater importance. To determine global importance, the absolute Shapley values per feature are averaged across the data.

Results

First, we extracted 22,382 patients with hospitalized dialysis records from the TMUCRD database. Next, we excluded patients with ESRD and previous dialysis records before the index time, leaving 10,160 patients. Then, based on the above-mentioned feature selection, we finally obtained 1080 patents and 2358 patients for the dialysis-dependent and survival analyses, respectively. The study workflow is shown in Figure 1. Table 1 and Supplementary 5 summarized the final features used for both the construction and analysis of our model. Supplementary 6 lists the number of patients with missing values for each selected feature.

Dialysis Dependent Module

Baseline Characteristics of Two Groups

The study comprised 1080 eligible first hospitalized AKI-D patients from the TMUCRD. Among them, 616 (57%) continued their dialysis after 90 days, while 464 (43%) did not. Table 1 and Supplementary 5 show the demographic characteristics, medication use, comorbidities, and laboratory results. The number of ICU admissions during hospitalization was significantly lower in patients who continued the dialysis after 90 days following initial dialysis. Significant differences were also observed for height and gender between the two groups, while there were no significant differences in age or BMI.

The Performance of the Prediction Model

For the model construction, we utilized 152 features as listed in Table 1 and Supplementary 5. The final model was generated by Vertex AI and consisted of an ensemble of 25 feedforward neural network models, which offered the most optimal prediction performance. The AUROC was 0.846 (Figure 2), and the accuracy, precision, sensitivity, specificity, and F1 score were found to be 0.788, 0.822, 0.845, 0.69, and 0.833, respectively.

Identification of the Significant Features

Based on the SHAP values, we identified the top 10 significant features that influenced our prediction model's performance. These features are illustrated in Figure 3. The 5 most important features were the baseline creatinine value, post-dialysis creatinine value, baseline eGFR value, post-dialysis eGFR value, and the BUN value difference

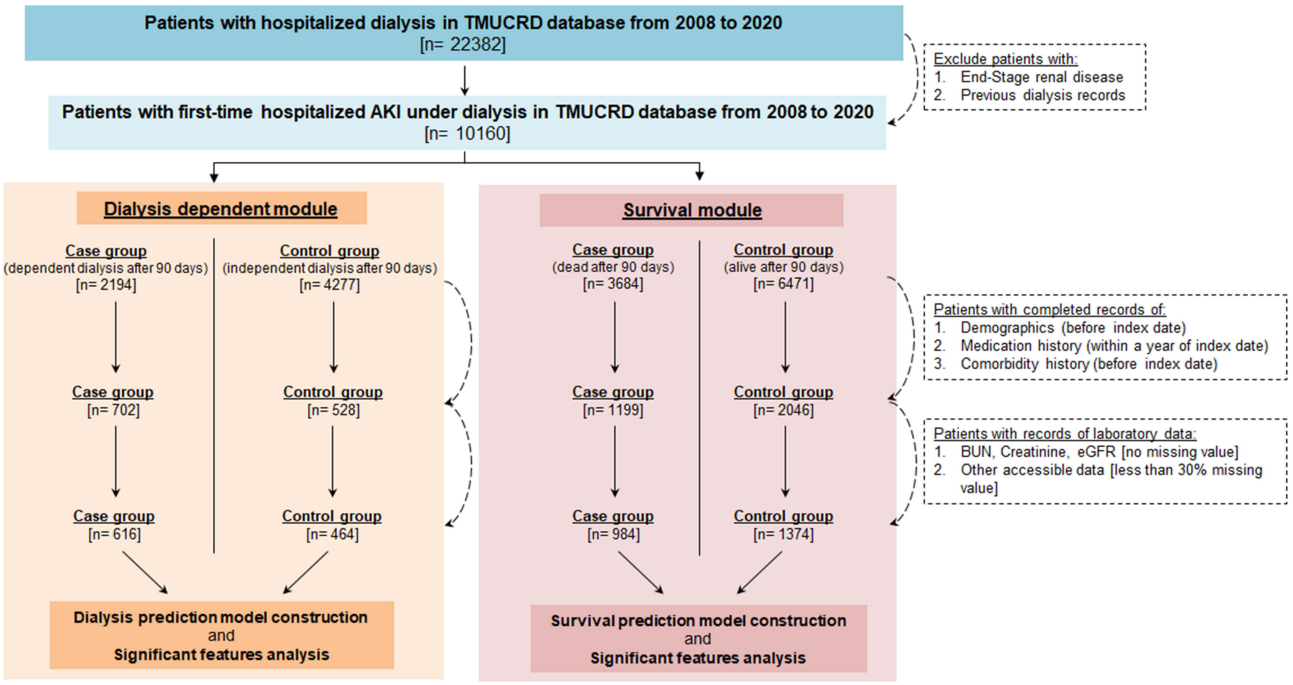

Figure 1 Workflow of the study design.

Table 1 Demographics of the Patients with AKI-Requiring Dialysis in the Study

	90-Day Dialysis Dependence					90-Day Survival				
	Case		Control		*P*-value	Case		Control		*P*-value
No. of patients	616		464			984		1374		
Gender	Male: 312	Female: 304	Male: 264	Female: 200	0.04	Male: 588	Female: 296	Male: 747	Female: 627	0.009
Age (years), mean ± SD,	69.63 ± 13.75		71.03 ± 14.54		0.06	75.21 ± 13.19		69.78 ± 14.02		*P*<0.001
Height (cm), mean ± SD	159.58 ± 8.98		160.88 ± 8.62		0.02	160.59 ± 8.80		160.16 ± 9.00		0.15
Weight (kg), mean ± SD	64.54 ± 16.17		65.6 ± 15.27		0.12	61.00 ± 13.93		65.37 ± 15.80		*P*<0.001
Body mass index (kg/m2), Mean ± SD	25.2 ± 5.20		25.26 ± 5.20		0.75	23.63 ± 4.92		25.36 ± 5.15		*P*<0.001
ICU[a] admission, n (%)	126 (20.45)		214 (46.12)		*P*<0.001	707 (71.85)		395 (28.75)		*P*<0.001
Comorbidity, n (%)										
Myocardial infarction	41 (6.66)		36 (7.76)		0.49	99 (10.06)		91 (6.62)		0.002
Congestive heart failure	200 (32.47)		141 (30.39)		0.47	327 (33.23)		432 (31.44)		0.36
Peripheral vascular disease	15 (2.44)		21 (4.53)		0.06	56 (5.69)		45 (3.28)		0.004
Cerebrovascular disease	97 (15.75)		93 (20.04)		0.07	196 (19.92)		230 (16.74)		0.048
Dementia	10 (1.62)		16 (3.45)		0.053	54 (5.49)		32 (2.33)		*P*<0.001
Chronic pulmonary disease	45 (7.31)		41 (8.84)		0.36	142 (14.43)		103 (7.50)		*P*<0.001
Connective tissue disease-rheumatic disease	9 (1.46)		8 (1.72)		0.73	21 (2.13)		17 (1.24)		0.09
Peptic ulcer disease	83 (13.47)		71 (15.3)		0.40	163 (16.57)		182 (13.25)		0.02
Mild liver disease	34 (5.52)		33 (7.11)		0.28	125 (12.70)		86 (6.26)		*P*<0.001
Diabetes without complications	282 (45.78)		197 (42.46)		0.28	343 (34.86)		611 (44.47)		*P*<0.001
Diabetes with complications	182 (29.55)		94 (20.26)		*P*<0.001	114 (11.59)		375 (27.29)		*P*<0.001
Paraplegia and hemiplegia	9 (1.46)		5 (1.08)		0.58	15 (1.52)		15 (1.09)		0.36
Renal disease	479 (77.76)		250 (53.88)		*P*<0.001	331 (33.64)		941 (68.49)		*P*<0.001
Cancer	83 (13.47)		87 (18.75)		0.02	329 (33.43)		196 (14.26)		*P*<0.001
Moderate or severe liver disease	4 (0.65)		9 (1.94)		0.054	45 (4.57)		17 (1.24)		*P*<0.001
Metastatic carcinoma	12 (1.95)		31 (6.68)		*P*<0.001	136 (13.82)		48 (3.49)		*P*<0.001
AIDS[b]/HIV[c]	0 (0)		0 (0)			1 (0.1)		0 (0)		0.42
Charlson comorbidity index, Mean ± SD	3.61 ± 2.02		3.55 ± 2.59		0.03	4.01 ± 2.97		3.53 ± 2.22		0.05

Notes: [a]ICU, intensive care unit; [b]AIDS, acquired immunodeficiency syndrome; [c]HIV, human immunodeficiency virus.

between the baseline and post-dialysis. All of the top 10 significant features indicated no statistical differences between the two groups, with the exception of the post-dialysis BUN value, as displayed in Table 2.

Survival Module

Baseline Characteristics of Two Groups

In the study, 2358 first hospitalized AKI-D patients were identified from the TMUCRD for survival analysis. Of these patients, 984 (42%) died within 90 days of initial dialysis, while 1374 (58%) did not. Demographic characteristics, medication use, comorbidities, and laboratory results are presented in Table 1 and Supplementary 5. In comparison to the dialysis dependent module, it was found that age and BMI exhibited significant differences while only height did not, between the two groups. Additionally, the patients who did not survive past 90 days from the initial dialysis had a statistically significant increase in the number of admissions to the ICU during hospitalization.

The Performance of the Prediction Model

For our model construction, we utilized 139 features detailed in Tables 1 and Supplementary 5. The final model created by Vertex AI consists of an ensemble of 15 feedforward neural network models and 10 GBDT models that deliver exceptional prediction performance. Our AUROC stands at 0.865 (Figure 4), while accuracy, precision, sensitivity, specificity, and F1 scores are recorded as 0.776, 0.729, 0.788, 0.766, and 0.757, respectively.

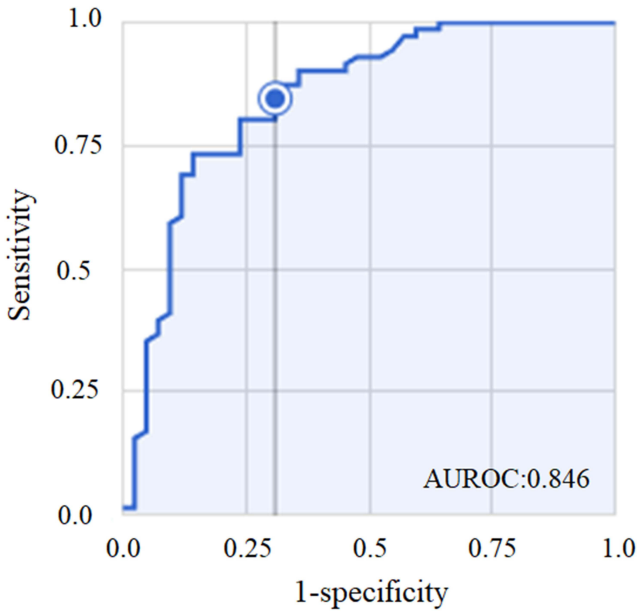

Figure 2 The AUROC for the final dialysis dependent prediction model.

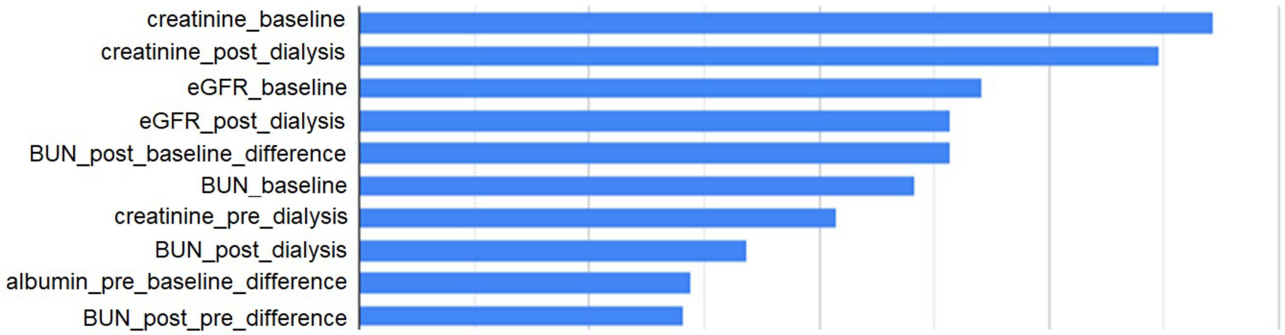

Figure 3 The top 10 significant features related to the dialysis dependent prognosis. Baseline: lab data tested beyond 90 days before index date; Pre: lab data tested within 90 days before the index date; Post: lab data tested within 2–8 weeks after the index date.

Identification of the Significant Features

We have identified the top 10 significant features that impacted the performance of our prediction model through SHAP value calculations. Figure 5 displays these features and indicates that the pre-dialysis creatinine value, baseline creatinine value, pre-dialysis eGFR value, baseline eGFR value, and age were the top 5 most significant factors. All of the top 10 significant features presented statistically significant differences between the two groups, as shown in Table 3.

Discussion

The Comparison of the Discriminative Power with Other Researches

The prediction of dialysis dependence after AKI-D has garnered considerable attention in the medical field. Accurate predictions hold great potential to enhance medical decision-making, improve quality of life, and alleviate financial burden of patients. With the recent advancements in artificial intelligence and machine learning, the application of these technologies in precision medicine has rapidly increased.

Previous studies predicting dialysis dependence or survival have mainly used single, deep, or basic machine learning algorithms (Supplementary 1). However, the implementation of ensemble learning techniques that integrate multiple innovative machine learning algorithms has been relatively restricted in these investigations.

Table 2 The Value of the Top 10 Significant Features Between Two Groups of the Dialysis Dependent Prediction Model

Features	Case	Control	*P*-value
Creatinine_baseline	5.57 ± 2.82	3.12 ± 2.22	*P*<0.001
Creatinine_post	6.87 ± 2.85	4.31 ± 2.67	*P*<0.001
eGFR[a]_baseline	14.8 ± 14.92	34.51 ± 28.18	*P*<0.001
eGFR[a]_post	9.6 ± 6.61	22.47 ± 20.76	*P*<0.001
BUN[b]_post_baseline_diff[c]	−8.32 ± 45.73	14.94 ± 39.19	*P*<0.001
BUN[b]_baseline	68.84 ± 29.64	45.46 ± 28.21	*P*<0.001
Creatinine_pre	9.01 ± 3.71	6.44 ± 3.49	*P*<0.001
BUN[b]_post	60.52 ± 34.15	60.4 ± 33.12	0.93
Albumin_pre_baseline_diff[c]	−0.26 ± 0.53	−0.43 ± 0.67	*P*<0.001
BUN[b]_post_pre_diff[c]	−48.27 ± 46.73	−33.08 ± 46.83	*P*<0.001

Notes: [a]eGFR: estimated glomerular filtration rate. [b]BUN: blood urea nitrogen. [c]Diff: difference; Baseline: lab data tested beyond 90 days before index date; Pre: lab data tested within 90 days before the index date; Post: lab data tested within 2–8 weeks after the index date.

Benjamin J Lee et al[36] developed prediction models to determine renal recovery after AKI-D within a 90-day period. The study included a total of 2214 adult AKI-D patients who underwent either acute intermittent hemodialysis or continuous RRT (CRRT). The logistic regression and classification and regression tree models had reported C-index values of 0.64 and 0.61, respectively. Dennis Emuron et al[37] built a predictive model using the Nutritional Risk Index along with other clinical parameters to predict 90-day dialysis dependence in patients with AKI requiring CRRT. The resultant logistic regression predictive model demonstrated a discriminative ability with an AUROC of 0.89. However, it is important to note that the study had only 77 patients for model training, and the final sensitivity outcome was 0.563, indicating a potential issue of the data imbalance. Vin-Cent Wu et al[38] employed a multivariate generalized additive model to predict 90-day survival and renal function recovery in AKI-D patients who initiated RRT. The study comprised a total of 258 patients. The plasma c-terminal FGF-23 predictive model exhibited the greatest ability to discriminate with an AUROC of 0.687.

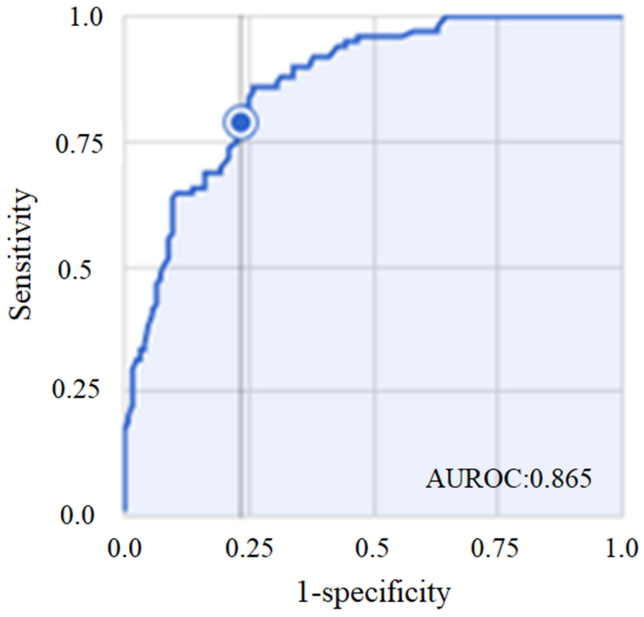

Figure 4 The AUROC for the final survival prediction model.

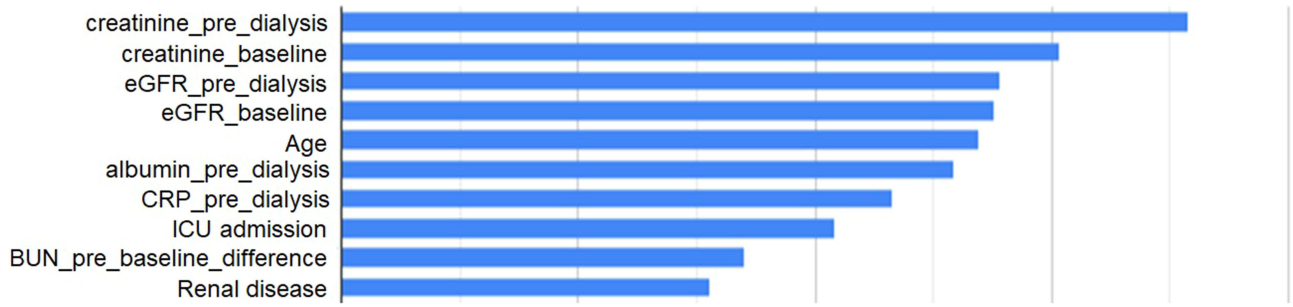

Figure 5 The top 10 significant features related to the survival prognosis. Baseline: lab data tested beyond 90 days before index date; Pre: lab data tested within 90 days before the index date; Post: lab data tested within 2–8 weeks after the index date.

Pattharawin Pattharanitima et al[39] applied multiple machine learning approaches to predict renal replacement therapy free survival (RRTFS) after discharge in 684 critically ill patients with AKI requiring CRRT, using all available biomedical data. Among the different approaches tested, the MLP + LSTM model demonstrated the highest performance, with an AUROC value of 0.70. Chang Liu et al[40] developed prediction models using routinely available clinical characteristics and risk factors to predict two outcomes: successful weaning from CRRT at the first attempt and RRTFS at hospital discharge. The study cohort involved 1135 AKI patients necessitating CRRT. The ultimate multivariable logistic regression models demonstrated discrimination with AUROC values of 0.76 and 0.78, respectively. Min Woo Kang et al[41] employed machine learning algorithms to predict patient mortality following CRRT for AKI. Biomedical data from 1571 adult patients were analyzed, and the Random Forest algorithm achieved an AUROC value of 0.768, the highest result. Daniel H Li et al[42] developed models for predicting 90-day mortality and RRT dependence in critically ill patients with severe AKI who underwent RRT. The logistic regression model's final performance, following external validation, exhibited an AUROC value of 0.61.

The study found that our predictive accuracy surpassed that of earlier review studies in terms of both 90-day dialysis dependence and 90-day mortality following first hospitalized AKI-D. We employed an ensemble model consisting of 25 feedforward neural network models, which produced the highest AUROC value of 0.846, as well as an accuracy of 0.788, precision of 0.822, sensitivity of 0.845, specificity of 0.69, and F1 score of 0.833 for predicting dialysis dependence. Regarding survival prediction, our ensemble model, consisting of 15 feedforward neural network models and 10 GBDT models achieved the highest AUROC value of 0.865, as well as an accuracy of 0.776, precision of 0.729, sensitivity of 0.788, specificity of 0.766, and F1 score of 0.757.

Table 3 The Value of the Top 10 Significant Features Between Two Groups of the Survival Prediction Model

Features	Case	Control	*P*-value
Creatinine_pre	4.27 ± 2.33	8.07 ± 3.87	*P*<0.001
Creatinine_baseline	2.01 ± 1.58	4.71 ± 2.83	*P*<0.001
eGFR[a]_pre	19.30 ± 15.56	9.34 ± 9.40	*P*<0.001
eGFR[a]_baseline	49.86 ± 29.84	21.91 ± 22.89	*P*<0.001
AGE	75.21 ± 13.19	69.78 ± 14.02	*P*<0.001
Albumin_pre	2.76 ± 0.60	3.31 ± 0.65	*P*<0.001
CRP[b]_pre	11.56 ± 9.98	6.65 ± 8.21	*P*<0.001
ICU[c]_admission	707 (71.85)	395 (28.75)	*P*<0.001
BUN[d]_pre_baseline_diff[e]	54.14 ± 42.32	42.18 ± 39.78	*P*<0.001
Renal disease	331 (33.64)	941 (68.49)	*P*<0.001

Notes: Baseline: lab data tested beyond 90 days before index date; Pre: lab data tested within 90 days before the index date; Post: lab data tested within 2–8 weeks after the index date.

Abbreviations: [a]eGFR, estimated glomerular filtration rate; [b]CRP, C-creative protein; [c]ICU, intensive care unit; [d]BUN, blood urea nitrogen; [e]Diff, difference.

We improved the performance by the implementation of ensemble learning techniques that integrate multiple innovative machine learning algorithms. With a larger sample size from three different hospitals, we integrated a significant amount of clinical data, such as demographics, comorbidities, medications, and laboratory results, to robustly train our models for improved prognosis prediction.

Identification of Potential Clinical Parameters Related to Dialysis Dependence and Mortality After 90 Days of the First Hospitalized AKI-D

In our dialysis-dependent analysis, we have identified the serum creatinine as a pivotal biomarker for predicting outcomes in the first hospitalized AKI-D. The creatinine levels at three different temporal points ranked among the top 10 significant features. It is noteworthy that patients who cannot withdraw from dialysis beyond 90 days present a markedly escalated creatinine level, as illustrated in Supplementary 5. We created a paired scatterplot comparing the value differences between two groups at different time points using the top 2 significant features: the baseline creatinine value and the creatinine value within 2–8 weeks after the AKI-D date. As shown in Figure 6, a higher positive slope was found in the case group between these two time points, indicating a poorer response to dialysis. This aligns with the existing literature, whereby an increase in serum creatinine typically indicates a decline in kidney function, possibly resulting in extended dialysis. Kashani et al[43] creatinine is the paramount functional biomarker for assessing renal function and is a key contributor to improved AKI detection and management. Similarly, Zavada et al[44] described the prominent influence of baseline creatinine on subsequent AKI detection and classification. Furthermore, the Acute Dialysis Quality Initiative Working Group released a consensus definition and classification scheme for AKI in 2004, with a focus on serum creatinine.[45]

In survival analysis, serum creatinine plays a crucial role in predicting 90-day survival after the first hospitalized AKI-D. The measurements of creatinine at two different time points ranked first and second in terms of feature significance. We presented both the top 2 features in a paired scatterplot, as shown in Figure 7. Interestingly, patients who did not survive beyond 90 days had lower creatinine levels compared to survivors. The control group also exhibited a higher positive slope, suggesting a more rapid increase in creatinine values. This may be due to more severe comorbidities in patients with relatively preserved renal function. Data in Supplementary 5 indicates that the deceased cohort was predominantly older, had elevated CRP values, reduced albumin levels, and higher rates of ICU admissions. Lobo et al[46] demonstrated that high serum CRP concentrations are associated with a greater likelihood of organ failure and death upon admission to the ICU, and are indicative of a poor outcome. Hannan et al[47] observed that low albumin levels, particularly in older patients, are linked to a higher risk of mortality. Additionally, while not among the top 10 significant features, the risk group exhibited significantly elevated levels of Troponin I and CPK. Aydin et al[48] reported that Troponin I is a commonly used biomarker for diagnosing acute myocardial infarction and CKP can estimate the size of the infarct. Despite the generally lower levels of creatinine, mortality could result from significant stressors that cause AKI-D and subsequent death. Consequently, we recommend a comprehensive approach that includes evaluating creatinine levels, age, CRP, albumin metrics, and ICU admissions for accurate survival prediction.

Finally, our study indicated the baseline creatinine value, evaluated at least 90 days before the first dialysis, was a crucial determinant of the prognosis from both prognostic modules. This suggests that not only does the laboratory data tested right before the dialysis matter, but also the baseline variability of the data plays an important role. If the change in the baseline value could be detected in advance, the high and low-risk groups could be identified earlier. If necessary, clinicians can adjust the drug dosage, medication use, and daily diet to decrease mortality rates and dialysis dependence.

Limitations

This study has limitations that should be acknowledged. The data collected were obtained from hospitals (eg TMUH, WFH, and SHH) in northern Taiwan, which may restrict the generalizability of the findings to patients in other regions. Furthermore, the study depended on retrospective data for model development and validation. The observed associations between clinical features and the outcomes of dialysis dependence or mortality may not necessarily imply a causal relationship. Additionally, the machine learning algorithms developed in this study were based on binary outcomes (eg,

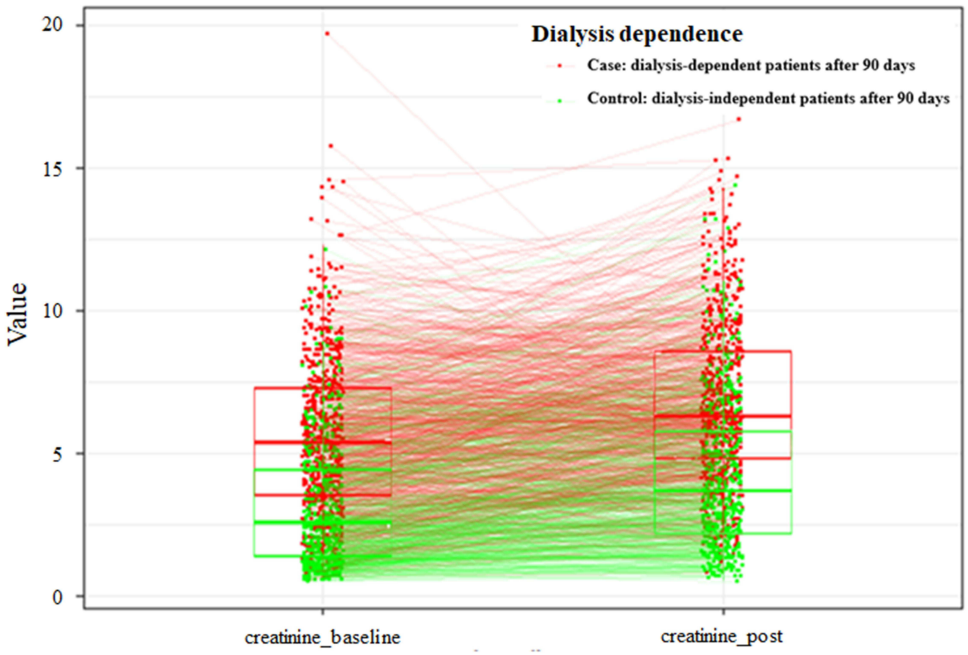

Figure 6 The paired scatterplot with box plots of dialysis dependent module. This illustrates the distribution of creatinine values at baseline and post-dialysis for the case group (depicted in red) and control group (depicted in green) within dialysis dependent module.

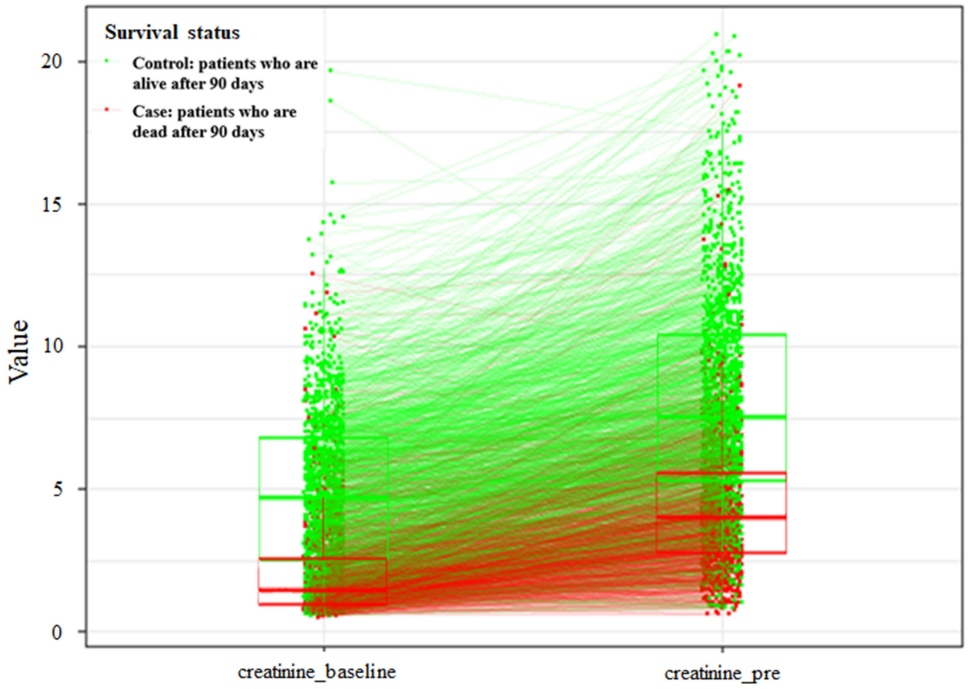

Figure 7 The paired scatterplot with box plots of survival module. The illustrates the distribution of creatinine values at baseline and pre-dialysis for the case group (depicted in red) and control group (depicted in green) within survival module.

dialysis or not), rather than continuous outcomes (eg, length of dialysis). Therefore, the models were unable to accurately predict the specific time of discontinuous dialysis but only provided predictions within a time range.

To enhance precision, future studies should validate the model using data from diverse geographical regions and a larger sample size to address the prediction of continuous outcomes. Additionally, further prospective studies are necessary to establish a more robust understanding of the causal relationships in clinical settings. Finally, we recommend

that future researchers focus on constructing a nomogram or decision tree. These methods offer clinicians a more direct and straightforward way to classify patients into different risk categories.

Conclusion

This study demonstrates the potential of ensemble machine learning algorithms to enhance prognostic accuracy for first hospitalized AKI-D patients. Our models were developed by using clinical data from three different hospitals, enabling us to successfully predict 90-day outcomes, which encompass the patient's reliance on dialysis and mortality rate. The integration of various clinical parameters significantly improved the predictive accuracy of the models. The first hospitalized dialysis outcomes were significantly influenced by the baseline creatinine value, which was evaluated at least 90 days prior. Therefore, early assessment and monitoring of creatinine levels in AKI-D patients is crucial. This allows clinicians to adjust drug dosage, medication use, and daily diet accordingly to prevent further renal damage. The use of prediction models can provide valuable insight into the overall prognosis, improving clinical decision-making, early intervention, and patient care for high-risk populations. This can lead to a better patient experience, improved quality of care, and more efficient resource utilization for healthcare providers.

Acknowledgments

The authors obtained data from the Taipei Medical University Clinical Research Database, approved by the Joint Institute Review Board of Taipei Medical University in Taipei, Taiwan (approved number, N201703004).

Funding

This study is supported by the National Science and Technology Council Grant (NSTC 111-2221-E-038 −022 -MY3). No funding bodies had any role in study design, data collection and analysis, decision to publish, or preparation of the article.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Bellomo R. The epidemiology of acute renal failure: 1975 versus 2005. *Curr Opin Crit Care*. 2006;12(6):557–560. doi:10.1097/01.ccx.0000247443.86628.68
2. Susantitaphong P, Cruz DN, Cerda J, et al. World incidence of AKI: a meta-analysis. *Clin J Am Soc Nephrol*. 2013;8(9):1482–1493. doi:10.2215/CJN.00710113
3. Uchino S. The epidemiology of acute renal failure in the world. *Curr Opin Crit Care*. 2006;12(6):538–543. doi:10.1097/01.ccx.0000247448.94252.5a
4. Lameire NH, Bagga A, Cruz D, et al. Acute kidney injury: an increasing global concern. *Lancet*. 2013;382(9887):170–179. doi:10.1016/S0140-6736(13)60647-9
5. Liangos O, Wald R, O'Bell JW, et al. Epidemiology and outcomes of acute renal failure in hospitalized patients: a national survey. *Clin J Am Soc Nephrol*. 2006;1(1):43–51. doi:10.2215/CJN.00220605
6. Lo LJ, Go AS, Chertow GM, et al. Dialysis-requiring acute renal failure increases the risk of progressive chronic kidney disease. *Kidney Int*. 2009;76(8):893–899. doi:10.1038/ki.2009.289
7. Hsu CY, Chertow GM, McCulloch CE, et al. Nonrecovery of kidney function and death after acute on chronic renal failure. *Clin J Am Soc Nephrol*. 2009;4(5):891–898. doi:10.2215/CJN.05571008
8. Cerda J. Promoting kidney function recovery in patients with AKI requiring RRT. *Clin J Am Soc Nephrol*. 2015;10(10):1859–1867. doi:10.2215/CJN.01170215
9. Gautam SC, Brooks CH, Balogun RA, et al. Predictors and outcomes of post-hospitalization dialysis dependent acute kidney injury. *Nephron*. 2015;131(3):185–190. doi:10.1159/000441607
10. Wald R, McArthur E, Adhikari NKJ, et al. Changing incidence and outcomes following dialysis-requiring acute kidney injury among critically ill adults: a population-based cohort study. *Am J Kidney Dis*. 2015;65(6):870–877. doi:10.1053/j.ajkd.2014.10.017
11. Rathore AS, Chopra T, Ma JZ, et al. Long-term outcomes and associated risk factors of post-hospitalization dialysis-dependent acute kidney injury patients. *Nephron*. 2017;137(2):105–112. doi:10.1159/000478277
12. Heung M. Outpatient dialysis for acute kidney injury: progress and pitfalls. *Am J Kidney Dis*. 2019;74(4):523–528. doi:10.1053/j.ajkd.2019.03.431
13. Negi S, Koreeda D, Kobayashi S, et al. Acute kidney injury: epidemiology, outcomes, complications, and therapeutic strategies. *Semin Dial*. 2018;31(5):519–527. doi:10.1111/sdi.12705
14. Harding JL, Li Y, Burrows NR, et al. US trends in hospitalizations for dialysis-requiring acute kidney injury in people with versus without diabetes. *Am J Kidney Dis*. 2020;75(6):897–907. doi:10.1053/j.ajkd.2019.09.012
15. Allegretti AS, Steele DJ, David-Kasdan J, et al. Continuous renal replacement therapy outcomes in acute kidney injury and end-stage renal disease: a cohort study. *Crit Care*. 2013;17(3):R109. doi:10.1186/cc12780

16. Gammelager H, Christiansen C, Johansen M, et al. Five-year risk of end-stage renal disease among intensive care patients surviving dialysis-requiring acute kidney injury: a nationwide cohort study. *Crit Care*. 2013;17(4):R145. doi:10.1186/cc12824
17. Hickson LJ, Chaudhary S, Williams AW, et al. Predictors of outpatient kidney function recovery among patients who initiate hemodialysis in the hospital. *Am J Kidney Dis*. 2015;65(4):592–602. doi:10.1053/j.ajkd.2014.10.015
18. Kao CC, Yang J-Y, Chen L, et al. Factors associated with poor outcomes of continuous renal replacement therapy. *PLoS One*. 2017;12(5):e0177759. doi:10.1371/journal.pone.0177759
19. Hsu RK, McCulloch CE, Dudley RA, et al. Temporal changes in incidence of dialysis-requiring AKI. *J Am Soc Nephrol*. 2013;24(1):37–42. doi:10.1681/ASN.2012080800
20. Bagshaw SM, Laupland KB, Doig CJ, et al. Prognosis for long-term survival and renal recovery in critically ill patients with severe acute renal failure: a population-based study. *Crit Care*. 2005;9(6):R700–9. doi:10.1186/cc3879
21. Uchino S. Acute renal failure in critically ill patients: a multinational, multicenter study. *JAMA*. 2005;294(7):813–818. doi:10.1001/jama.294.7.813
22. Bell M, Granath F, Schön S, et al. Continuous renal replacement therapy is associated with less chronic renal failure than intermittent haemodialysis after acute renal failure. *Intensive Care Med*. 2007;33(5):773–780. doi:10.1007/s00134-007-0590-6
23. Korkeila M, Ruokonen E, Takala J. Costs of care, long-term prognosis and quality of life in patients requiring renal replacement therapy during intensive care. *Intensive Care Med*. 2000;26(12):1824–1831. doi:10.1007/s001340000726
24. Silvester W, Bellomo R, Cole L. Epidemiology, management, and outcome of severe acute renal failure of critical illness in Australia. *Crit Care Med*. 2001;29(10):1910–1915. doi:10.1097/00003246-200110000-00010
25. Nisula S, Kaukonen K-M, Vaara ST, et al. Incidence, risk factors and 90-day mortality of patients with acute kidney injury in Finnish intensive care units: the FINNAKI study. *Intensive Care Med*. 2013;39(3):420–428. doi:10.1007/s00134-012-2796-5
26. Investigators RRTS. Intensity of continuous renal-replacement therapy in critically ill patients. *N Engl J Med*. 2009;361(17):1627–1638.
27. Hayward J, Alvarez SA, Ruiz C, et al. Machine learning of clinical performance in a pancreatic cancer database. *Artif Intell Med*. 2010;49 (3):187–195. doi:10.1016/j.artmed.2010.04.009
28. Mohamadlou H, Lynn-Palevsky A, Barton C, et al. Prediction of acute kidney injury with a machine learning algorithm using electronic health record data. *Can J Kidney Health Dis*. 2018;5:2054358118776326. doi:10.1177/2054358118776326
29. Shawwa K, Ghosh E, Lanius S, et al. Predicting acute kidney injury in critically ill patients using comorbid conditions utilizing machine learning. *Clin Kidney J*. 2021;14(5):1428–1435. doi:10.1093/ckj/sfaa145
30. Wang C, Hiremath S, Sikora L, et al. Outpatient kidney recovery after acute kidney injury requiring dialysis: a systematic review protocol. *Syst Rev*. 2019;8(1):214. doi:10.1186/s13643-019-1134-1
31. Goyal A. Acute Kidney Injury, in StatPearls. Treasure Island (FL) ineligible companies. Disclosure: Parnaz Daneshpajouhnejad declares no relevant financial relationships with ineligible companies. Disclosure: Muhammad Hashmi declares no relevant financial relationships with ineligible companies. Disclosure: Khalid Bashir declares no relevant financial relationships with ineligible companies; 2023.
32. Seto H, Oyama A, Kitora S, et al. Gradient boosting decision tree becomes more reliable than logistic regression in predicting probability for diabetes with big data. *Sci Rep*. 2022;12(1):15889. doi:10.1038/s41598-022-20149-z
33. Reisinho J, Coimbra M, Renna F. Deep convolutional neural network ensembles for multi-classification of skin lesions from dermoscopic and clinical images. *Annu Int Conf IEEE Eng Med Biol Soc*. 2020;2020:1940–1943. doi:10.1109/EMBC44109.2020.9176411
34. Rodriguez-Perez R, Bajorath J. Interpretation of compound activity predictions from complex machine learning models using local approximations and shapley values. *J Med Chem*. 2020;63(16):8761–8777. doi:10.1021/acs.jmedchem.9b01101
35. Chen H, Lundberg SM, Lee SI. Explaining a series of models by propagating Shapley values. *Nat Commun*. 2022;13(1):4512. doi:10.1038/s41467-022-31384-3
36. Lee BJ, Hsu C-Y, Parikh R, et al. Predicting renal recovery after dialysis-requiring acute kidney injury. *Kidney Int Rep*. 2019;4(4):571–581. doi:10.1016/j.ekir.2019.01.015
37. Emuron D, Thomas K, Mullane R. The nutritional risk index as a predictor of 90-day dialysis dependence after acute renal failure: a pilot study. *J Ren Nutr*. 2023;33(1):29–34. doi:10.1053/j.jrn.2022.03.009
38. Wu VC, Shiao -C-C, Chi N-H, et al. Outcome prediction of acute kidney injury biomarkers at initiation of dialysis in critical units. *J Clin Med*. 2018;7(8):202. doi:10.3390/jcm7080202
39. Pattharanitima P, Vaid A, Jaladanki SK, et al. Comparison of approaches for prediction of renal replacement therapy-free survival in patients with acute kidney injury. *Blood Purif*. 2021;50(4–5):621–627. doi:10.1159/000513700
40. Liu C, Peng Z, Dong Y, et al. Predicting successful continuous renal replacement therapy liberation in critically ill patients with acute kidney injury. *J Crit Care*. 2021;66:6–13. doi:10.1016/j.jcrc.2021.07.020
41. Kang MW, Kim J, Kim DK, et al. Machine learning algorithm to predict mortality in patients undergoing continuous renal replacement therapy. *Crit Care*. 2020;24(1):42. doi:10.1186/s13054-020-2752-7
42. Li DH, Wald R, Blum D, et al. Predicting mortality among critically ill patients with acute kidney injury treated with renal replacement therapy: development and validation of new prediction models. *J Crit Care*. 2020;56:113–119. doi:10.1016/j.jcrc.2019.12.015
43. Kashani K, Rosner MH, Ostermann M. Creatinine: from physiology to clinical application. *Eur J Intern Med*. 2020;72:9–14. doi:10.1016/j.ejim.2019.10.025
44. Zavada J, Hoste E, Cartin-Ceba R, et al. A comparison of three methods to estimate baseline creatinine for RIFLE classification. *Nephrol Dial Transplant*. 2010;25(12):3911–3918. doi:10.1093/ndt/gfp766
45. Bellomo R, Ronco C, Kellum JA, et al. Acute renal failure - definition, outcome measures, animal models, fluid therapy and information technology needs: the Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group. *Crit Care*. 2004;8(4):R204–12. doi:10.1186/cc2872
46. Lobo SM, Lobo FRM, Bota DP, et al. C-reactive protein levels correlate with mortality and organ failure in critically ill patients. *Chest*. 2003;123 (6):2043–2049. doi:10.1378/chest.123.6.2043
47. Hannan JL, Radwany SM, Albanese T. In-hospital mortality in patients older than 60 years with very low albumin levels. *J Pain Symptom Manage*. 2012;43(3):631–637. doi:10.1016/j.jpainsymman.2011.04.009
48. Aydin S, Ugur K, Aydin S, et al. Biomarkers in acute myocardial infarction: current perspectives. *Vasc Health Risk Manag*. 2019;15:1–10. doi:10.2147/VHRM.S166157