



Published in final edited form as:

Stud Health Technol Inform. 2019 August 21; 264: 462–466. doi:10.3233/SHTI190264.

Stratified Mortality Prediction of Patients with Acute Kidney Injury in Critical Care

Zhenxing Xu^a, Yuan Luo^b, Prakash Adekkanattu^a, Jessica S. Ancker^a, Guoqian Jiang^c, Richard C. Kiefer^c, Jennifer A. Pacheco^b, Luke V. Rasmussen^b, Jyotishman Pathak^a, Fei Wang^a

^aWeill Cornell Medicine, Cornell University, New York, New York, USA

^bNorthwestern University Feinberg School of Medicine, Chicago, Illinois, USA

^cMayo Clinic, Rochester, Minnesota, USA

Abstract

Acute Kidney Injury (AKI) is the most common cause of organ dysfunction in critically ill adults and prior studies have shown AKI is associated with a significant increase of the mortality risk. Early prediction of the mortality risk for AKI patients can help clinical decision makers better understand the patient condition in time and take appropriate actions. However, AKI is a heterogeneous disease and its cause is complex, which makes such predictions a challenging task. In this paper, we investigate machine learning models for predicting the mortality risk of AKI patients who are stratified according to their AKI stages. With this setup we demonstrate the stratified mortality prediction performance of patients with AKI is better than the results obtained on the mixed population.

Keywords

Acute Kidney Injury; Critical Care; Forecasting

Introduction

Acute Kidney Injury (AKI) is defined as a sudden (within hours) decrease in nephritic function, including both renal structural damage and functional disorder [1]. It has been shown AKI is the most common cause of organ dysfunction in the critical care setting, and AKI is associated with a significant increase on the morbidity and mortality risk [2]. Early prediction of the mortality risk for AKI patients can help clinical decision makers better understand patient conditions in time and take appropriate actions.

There has been prior research on building mortality risk prediction models for AKI patients. For example, Luo et al. [3] created a scoring model based on multivariate logistic regression

This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).

Address for Correspondence: Fei Wang: Division of Health Informatics, Department of Healthcare Policy and Research, Weill Cornell Medical College, Cornell University, 425 E 61 St. New York City, NY 06032. few2001@med.cornell.edu.

to predict the 90-day mortality risk of AKI patients. Skarupskien et al. [4] adopted logistic regression to predict the mortality risk of AKI patients requiring renal replacement therapy after cardiac surgery. Demirjian et al. [5] used logistic regression to predict the 60-day mortality risk of the AKI patients who are enrolled in the Veterans Affairs/National Institutes of Health Acute Renal Failure Trial Network study. Ohnuma et al. [6] compared the ability of various mortality prediction models in a retrospective, multi-centric cohort of 343 Japanese patients with AKI requiring continuous renal replacement therapy in 14 ICUs captured between January and December 2010.

All of these existing research tried to build a unique predictor for the entire AKI patient cohort. However, it has been demonstrated AKI is a heterogeneous disease with various kinds of causes and clinical manifestations [7]. In this case, it would be challenging to build a single model that is capable of predicting the mortality risk over the entire AKI patient population.

With the above considerations and previous studies [8,9], we propose to perform a stratified prediction in this paper. Specifically, we will first stratify the AKI patients according to their disease stage (I, II, and III). Then, a mortality risk prediction model will be built for each patient strata. In addition to logistic regression, we will also test the performance of other machine learning models including random forest and gradient boosting tree. The Medical Information Mart for Intensive Care III (MIMIC-III) database [10] is used for empirical evaluations. Our results demonstrate the performance of stratified mortality prediction considering different AKI stages is higher than the mortality prediction on the entire mixed population. Additionally, the experiment results show different features play different roles for mortality prediction in patients with different AKI stages.

Methods

Data source

The Medical Information Mart for Intensive Care III (MIMIC-III) database is employed to extract patient data [10], which includes 58,976 admissions of patients. This database has comprehensive information (e.g., patient demographics and vital signs) regarding ICU admissions, which is open-source and freely accessible. The MIMIC-III dataset is passive and de-identified [11], which is in compliance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule and does not make a significant impact on patient privacy.

AKI stages

The criteria of AKI stages has experienced some development progress based on the studies of researchers and clinicians. In particular, there are four AKI criteria: Risk-Injury-Failure-Loss-End (RIFLE) criteria [12], pediatric RIFLE (pRIFLE) criteria [13], Acute Kidney Injury Network (AKIN) criteria [14], and Kidney Disease: Improving Global Outcomes (KDIGO) criteria [15]. KDIGO criteria unified previous criteria in 2012, which improves the sensitivity of AKI diagnostic criteria and has been widely used by researchers and

physicians [1]. In this study, we employed the KDIGO criteria to stratify patients into different AKI stages:

Stage 1: 1.5–1.9 times baseline, which is known or presumed to have occurred within the prior 7 days; or not less than 0.3 mg/dL (not less than 26.5 mol/L) absolute increase in serum creatinine (SCr); or urine volume less than 0.5 mL/kg/h for 6–12 hours.

Stage 2: SCr not less than 2.0–2.9 times baseline; or urine volume less than 0.5 mL/kg/h for not less than 12 hours.

Stage 3: SCr not less than 3.0 times from baseline; or increase in SCr not less than 4.0 mg/dL (not less than 353.6 mol/L); or initiation of renal replacement therapy; or urine volume less than 0.3 mL/kg/h for not less than 24 hours; or anuria for not less than 12 hours; or renal replacement therapy required.

Patient features

A large number of features from MIMIC-III are extracted, which is shown as follows. Additional details about features can be found at: https://github.com/xuzhenxing2018/amia/blob/master/features_name.xlsx.

1. Demographics: Gender, age, and ethnicity.
2. Medications: Medications the patient took from the patients' ICU admission until prediction time. We focused on the following categories: diuretics, non-steroidal anti-inflammatory drugs (NSAID), radiocontrast agents, and angiotensin.
3. Comorbidities: We considered all comorbidities (e.g., congestive heart failure, peripheral vascular, hypertension, diabetes, liver disease, myocardial infarction (MI), coronary artery disease (CAD), cirrhosis, and jaundice) that patients already had. There were a total of 31 comorbidities in our dataset.
4. Chart-events: Vital signs measured at the bedside. We mainly focused on diastolic blood pressure (DiasBP), glucose, heart rate, mean arterial blood pressure (MeanBP), respiration rate, blood oxygen saturation level (SpO₂), systolic blood pressure (SysBP), and temperature.
5. Lab-events: Laboratory test results. We considered bicarbonate, blood urea nitrogen (BUN), calcium, chloride, creatinine, hemoglobin, international normalized ratio (INR), platelet, potassium, prothrombin time (PT), partial thromboplastin time (PTT), and white blood count (WBC).

Data pre-processing

We mainly pre-processed two types of features: time-dependent continuous features and discrete features. For continuous features (e.g., lab and chart-events), we computed statistics including the first, last, average, minimum, maximum, slope, and the count based on observations during the observation window. The final encoding of the continuous features was represented by a vector representation with real numbers. The discrete features (e.g.,

medication and comorbidities) were encoded as zero-one multi-hot vectors. If there were missing demographics and discrete features for an ICU stay record, we deleted that ICU stay. Finally, each ICU stay was represented by a feature vector with a 224 dimension, which was indexed by `icustay_id`.

Experimental setting

In this study, we predicted whether a patient with AKI in different stages would die within the next seven days. Specifically, a predictive modeling setting with a rolling observation window design was adopted. We supposed t was the elapsed time (in hours) after a patient was admitted to the ICU, and the patient records in t were used to predict the mortality risk for the next seven days, where t took the value of 24, 48, 72, 96, 120, and 144 hours. An illustration of this rolling window design is demonstrated in Figure 1.

For each ICU stay, the AKI severity stage was identified based on whether the criteria in KDIGO stages were satisfied in an observation window, and then we put this ICU stay into corresponding cohorts (AKI Stage-I cohort, AKI Stage-II cohort, and AKI Stage-III cohort). If an ICU stay experienced multiple AKI stages during the observation window (from ICU admission to prediction point), we only chose the last AKI stage. Based on the built three cohorts, we predicted whether a patient with AKI in different stages would die within the next seven days. If an ICU stay did not meet one of the three AKI stage criteria in the observation window, we excluded the ICU stay from the prediction.

Predictive models

Several popular machine learning methods including Logistic Regression (LR) [16], L2 norm regularized Logistic Regression (Ridge) [17], Random Forest (RF) [18], and Gradient Boosting Decision Tree (GBDT) [19] were utilized to build predictive models using free and publicly available software. Specifically, for the implementations of LR, Ridge, and RF, we employed the Scikit-learn software library [20]. For the GBDT, we used the XGBoost software library [21]. For each predictive model, we used the 5-fold cross validation to assess the performance of these algorithms, and employed several popular and important metrics (AUC, recall, and precision) to evaluate the performance of these models.

Note that, the number of cases (mortality patient with AKI) and controls (alive patient with AKI) was imbalanced (e.g., an approximate 1:10 case to control ratio in 24 hours for AKI Stage-I) in our dataset. More statistics for all AKI stages are shown in Table 1. With such an imbalanced dataset, most classifiers had a potential to support the majority class (alive patient with AKI) because they were considered to maximize the overall number of correct predictions, thus resulting in poor performance in the minority class (mortality patient with AKI) prediction [22].

To address this imbalanced issue, we used case-control matching techniques [23] in this study, which matched each case with a control by considering the APACHE II score [24], Charlson comorbidity index [25], and demographic information. A matched control met two conditions: (1) has the same gender as the case and the age difference is within 5 years; (2) has the highest similarity score with the case measured by Manhattan distance on the basis

of APACHE II and Charlson comorbidity index. By this technique, a resampled balanced training set was constructed with matched cases and controls.

Results

Comparison of different methods with all patient features

We tested the performance of different predictive models with different AKI stages using all patient features with varying data observation windows. Figure 2 shows the performance of these methods in terms of AUC. In order to compare the performance of mortality prediction on subpopulations completely, we constructed a mixed cohort, which combined all the patients with different AKI stages together. This mixed cohort is represented as Stage-I-II-III in Figure 2. For the performance in terms of precision and recall, please refer to <https://github.com/xuzhenxing2019/MedInfo>.

From Figure 2, we observed that:

1. The technique of stratifying patients with AKI into different stages obtained a better performance compared to the mixed cohort. This provided some evidence for our assumption that stratifying patients with AKI into different stages has the potential to improve the performance of mortality prediction because of AKI's complex etiology and pathophysiology.
2. The technique of constructing a mortality predictor based on different cohorts has the potential to consider different AKI causes. This predictive modelling strategy can be viewed as a case for local learning [26], which first splits the data space into multiple local areas and constructs different predictors for the areas.
3. Comparing the Stages, Stage-III shows a higher mortality predictive performance. One implication is the AKI patients in Stage-III who are going to die had more distinctive clinical manifestations compared to the control patients.
4. Comparing the performance of four predictors, GBDT acquires better results.

Comparison of feature group predictability with GBDT

Five different groups of features were employed to build mortality predictors for patients with different AKI stages. In this section we explored the prediction performance of GBDT with different feature groups within a 24 hour data observation window. The mortality prediction results are demonstrated in Table 2.

We observed feature groups played different roles when patients were in different AKI stages. For example, the laboratory features acquired better performance than other feature groups when patients were diagnosed with AKI Stage-I. Chart features played a more important role in predicting mortality when patients were in Stage-II and III. One potential reason for this is the chart events had a finer time resolution, and for patients with a later stage of AKI, these events provided more in-time information of the patients' condition which progressed rapidly. In addition, there were no obvious differences between comorbidities and medications feature groups. The demographic feature group contributed the least to mortality prediction in patients with different stages.

The important features chosen from all feature groups

There were 224 features extracted in total for mortality prediction in different AKI stages. The importance of each factor in mortality prediction was further explored in our experiments. The GBDT model was employed to acquire the importance score of each feature based on 24 hour data. The range of the feature importance score for mortality prediction in AKI Stage-I, Stage-II, and Stage-III were 0.0015–0.0457, 0.0015–0.0556, and 0.0019–0.0316, respectively. The top ten features are shown in Table 3.

From the table, we can observe HeartRate, CREATININE, CHLORIDE, WBC, HEMOGLOBIN, and RespRate were more important because there were strong associations between these features and mortality in patients with AKI. These results aligned well with some previous studies [27–32]. For example, the magnitude of increase in serum creatinine levels can determine the severity of AKI [27], which was associated with worse survival rates [28]. Chloride levels are also associated with the severity of AKI [29] and Shaw et al. demonstrated an association between higher intravenous chloride loads and hospital mortality [30].

Discussion and Conclusions

Accurate prediction of the mortality risk for AKI patients is helpful for clinicians to understand the condition of the patients and take appropriate actions. AKI has a complex etiology and pathophysiology, so it is challenging to construct a single model that can accurately predict the mortality risk over the entire AKI patient cohort. This study investigated the impact of different AKI subpopulations on the performance of mortality prediction. We stratified the AKI patients on the basis of their disease stage (I, II, and III), and then a mortality risk predictor was constructed for each patient cohort. Several popular machine learning models (e.g., logistic regression, RF, and GBDT) were built based on these subpopulations for the mortality prediction in patients with AKI. GBDT showed a better performance than other methods for the mortality prediction in this study.

Prior models of risk prediction of mortality in critically ill patients with AKI typically employed mixed cohorts [5,6]. The utilization of stratification mechanisms performed in this study provides a chance to predict mortality in patients with AKI specific to each stage. This technique of constructing a mortality predictor based on different cohorts has the potential to consider different AKI causes. In addition, dynamic real-time data may allow for clinical monitoring of mortality risk beyond a solely static prediction upon ICU arrival. The utilization of dynamic clinical monitoring algorithms for mortality prediction may allow for its incorporation into the electronic medical record [33].

With the ability to identify patients at high risk of mortality by employing real-time data may allow for initiation of earlier interventions to prevent or reduce mortality [34]. It might be useful to incorporate more information (e.g., imaging and genomic biomarkers) to further improve the performance of dynamic clinical mortality risk prediction models. In the future, we will consider some advanced models (e.g., recurrent network) with stronger learning capabilities to improve the performance of prediction.

Acknowledgements

This work was supported in part by NIH Grants 2R01GM105688-06 and 1R21LM012618-01.

References

- [1]. Makris K and Spanou L, Acute kidney injury: Definition, pathophysiology and clinical phenotypes, *The Clinical Biochemist Reviews* 37 (2016) 85–98. [PubMed: 28303073]
- [2]. Doyle JF and Forni LG, Acute kidney injury: Short-term and long-term effects, *Critical Care* 20 (2016) 188. [PubMed: 27373891]
- [3]. Luo M, Yang Y, Xu J, et al. , A new scoring model for the prediction of mortality in patients with acute kidney injury, *Scientific Reports* 7 (2017) 7862. [PubMed: 28801674]
- [4]. Skarupskienė I, Adukauskienė D, Kuzminskienė J, et al. , Mortality prediction in patients with acute kidney injury requiring renal replacement therapy after cardiac surgery, *Medicina (Kaunas, Lithuania)* 53 (2017) 217–223.
- [5]. Demirjian S, Chertow GM, Zhang JH, et al. , Model to predict mortality in critically ill adults with acute kidney injury, *Clin J Am Soc Nephrol* 6 (2011) 2114–2120. [PubMed: 21896828]
- [6]. Ohnuma T, Uchino S, Toki N, et al. , External validation for acute kidney injury severity scores: A multicenter retrospective study in 14 Japanese ICUs, *American Journal of Nephrology* 42 (2015) 57–64. [PubMed: 26337793]
- [7]. Bellomo R, Kellum JA, and Ronco C, Acute kidney injury, *Lancet (London, England)* 380 (2012) 756–766.
- [8]. Miller CC, Reardon MJ, and Safi HJ, *Risk Stratification: A Practical Guide for Clinicians*, 1st ed., Cambridge University Press, Cambridge, 2001.
- [9]. Kraus VB, Blanco FJ, Englund M, et al. , Call for standardized definitions of osteoarthritis and risk stratification for clinical trials and clinical use, *Osteoarthritis and Cartilage* 23 (2015) 1233–1241. [PubMed: 25865392]
- [10]. Johnson AE, Pollard TJ, Shen L, et al. , MIMIC-III, a freely accessible critical care database, *Scientific Data* 3 (2016) 160035. [PubMed: 27219127]
- [11]. 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* 5 (2018) 2054358118776326. [PubMed: 30094049]
- [12]. 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, *Critical Care (London, England)* 8 (2004) R204–R212.
- [13]. Akcan-Arkan A, Zappitelli M, Loftis LL, et al. , Modified RIFLE criteria in critically ill children with acute kidney injury, *Kidney Int* 71 (2007) 1028–1035. [PubMed: 17396113]
- [14]. Pickering JW and Endre ZH, GFR shot by RIFLE: Errors in staging acute kidney injury, *Lancet (London, England)* 373 (2009) 1318–1319.
- [15]. Kellum JA, Lameire N, Aspelin P, et al. , Kidney disease: Improving global outcomes (KDIGO) acute kidney injury work group. KDIGO clinical practice guideline for acute kidney injury, *Kidney International Supplements* 2 (2012) 1–138.
- [16]. Hosmer DW, Lemeshow S, and Sturdivant RX, *Applied Logistic Regression*, 3rd ed., John Wiley & Sons, Inc., Hoboken, 2013.
- [17]. Le Cessie S and Van Houwelingen JC, Ridge estimators in logistic regression, *Journal of the Royal Statistical Society, Series C (Applied Statistics)* 41 (1992) 191–201.
- [18]. Breiman L, Random forests, *Machine Learning* 45 (2001) 5–32.
- [19]. Friedman JH, Stochastic gradient boosting, *Computational Statistics & Data Analysis* 38 (2002) 367–378.
- [20]. Pedregosa F, Varoquaux G, Gramfort A, et al. , Scikitlearn: Machine learning in python, *Journal of Machine Learning Research* 12 (2011) 2825–2830.

- [21]. Chen T and Guestrin C, XGBoost: A scalable tree boosting system, In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), San Francisco, CA, USA, 2016, 785–794.
- [22]. Cheng P, Waitman LR, Hu Y, et al., Predicting inpatient acute kidney injury over different time horizons: How early and accurate?, In AMIA 2017 Annual Symposium Proceedings, Washington, D.C., USA, 2017, 565–574.
- [23]. Cologne JB and Shibata Y, Optimal case-control matching in practice, *Epidemiology (Cambridge, Massachussetts)* 6 (1995) 271–275.
- [24]. Knaus WA, Draper EA, Wagner DP, et al. , APACHE II: A severity of disease classification system, *Critical Care Medicine* 13 (1985) 818–829. [PubMed: 3928249]
- [25]. Charlson ME, Pompei P, Ales KL, et al. , A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation, *Journal of Chronic Diseases* 40 (1987) 373–383. [PubMed: 3558716]
- [26]. Bottou L and Vapnik V, Local learning algorithms, *Neural Computation* 4 (1992) 888–900.
- [27]. Kellum JA, Sileanu FE, Murugan R, et al. , Classifying AKI by urine output versus serum creatinine level, *J Am Soc Nephrol* 26 (2015) 2231–2238. [PubMed: 25568178]
- [28]. Linder A, Fjell C, Levin A, et al. , Small acute increases in serum creatinine are associated with decreased long-term survival in the critically ill, *Am J Respir Crit Care Med* 189 (2014) 1075–1081. [PubMed: 24601781]
- [29]. Oh HJ, Kim S, Park JT, et al. , Baseline chloride levels are associated with the incidence of contrast-associated acute kidney injury, *Scientific Reports* 7 (2017) 17431. [PubMed: 29234129]
- [30]. Shaw AD, Raghunathan K, Peyerl FW, et al. , Association between intravenous chloride load during resuscitation and in-hospital mortality among patients with SIRS, *Intensive Care Medicine* 40 (2014) 1897–1905. [PubMed: 25293535]
- [31]. Bouchard J, Soroko SB, Chertow GM, et al. , Fluid accumulation, survival and recovery of kidney function in critically ill patients with acute kidney injury, *Kidney International* 76 (2009) 422–427. [PubMed: 19436332]
- [32]. Chen YC, Tsai FC, Chang CH, et al. , Prognosis of patients on extracorporeal membrane oxygenation: The impact of acute kidney injury on mortality, *The Annals of Thoracic Surgery* 91 (2011) 137–142. [PubMed: 21172502]
- [33]. Koyner JL, Adhikari R, Edelson DP, et al. , Development of a multicenter ward-based AKI prediction model, *Clin J Am Soc Nephrol* 11 (2016) 1935–1943. [PubMed: 27633727]
- [34]. Keim-Malpass J, Enfield KB, Calland JF, et al. , Dynamic data monitoring improves predictive analytics for failed extubation in the ICU, *Physiological Measurement* 39 (2018) 075005.

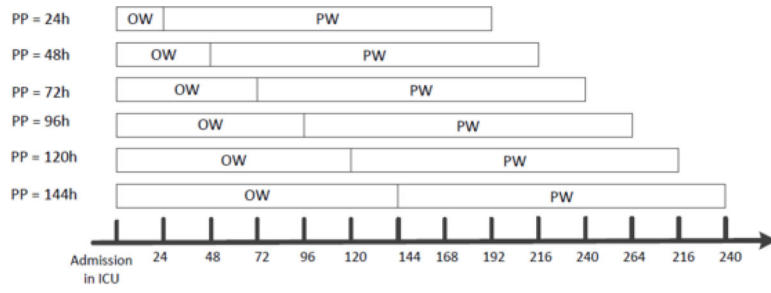


Figure 1 -
An illustration of mortality prediction for patients with AKI in different stages. PP: Prediction Point; OW: Observation Window; PW: Prediction Window

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

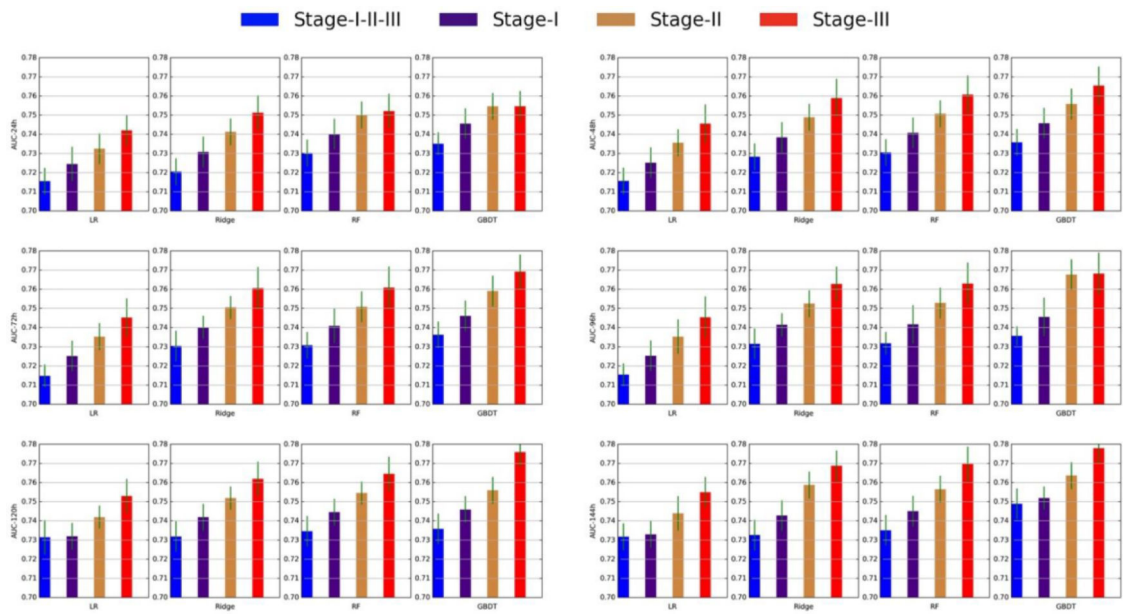


Figure 2 -
The AUC of different methods with 24–144 hours of data observation window in terms of different AKI stages

Table 1 -

The number of mortality and alive patients with AKI in different stages

	24h	48h	72h	96h	120h	144h	
Stage-I	Total	7506	8181	5801	4252	3222	2416
	Mortality	756	816	715	647	567	493
	Alive	6750	7365	5086	3605	2655	1923
	Mortality Rate	10.07%	9.97%	12.33%	15.22%	17.6%	20.41%
Stage-II	Total	3216	2895	2025	1576	1267	1007
	Mortality	514	514	421	386	347	310
	Alive	2702	2381	1604	1190	920	697
	Mortality Rate	15.98%	17.75%	20.79%	24.49%	27.39%	30.78%
Stage-III	Total	408	414	404	398	380	367
	Mortality	188	167	176	159	158	161
	Alive	220	247	228	239	222	206
	Mortality Rate	46.08%	40.34%	43.56%	39.95%	41.58%	43.87%

The performance of different group features in all AKI stages on the basis of 24 hours data

Table 2 -

	Demographics	Medications	Comorbidities	Chart-events	Lab-events
Stage-I	0.6823±0.0219	0.6971±0.0218	0.7027±0.0233	0.7031±0.0241	0.7237±0.0223
Stage-II	0.7029±0.0222	0.7128±0.0221	0.7139±0.0237	0.7339±0.0246	0.7331±0.0243
Stage-III	0.7134±0.0238	0.7236±0.0247	0.7239±0.0251	0.7424±0.0255	0.7334±0.0249

Table 3 -

The top 10 features selected from all feature groups on the basis of importance score

	Stage-I		Stage-II		Stage-III	
	Features	Importance	Features	Importance	Features	Importance
1	POTASSIUM_count	0.0457	HEMOGLOBIN_slope	0.0556	HeartRate_last	0.0316
2	metastatic_cancer	0.0365	HeartRate_last	0.0278	MeanBP_last	0.0247
3	CREATININE_avg	0.0350	CHLORIDE_slope	0.0200	HEMOGLOBIN_slope	0.0237
4	BUN_last	0.0289	WBC_min	0.0185	RespRate_last	0.0237
5	PTT_avg	0.0274	CREATININE_slope	0.0173	BICARBONATE_slope	0.0221
6	WBC_min	0.0243	CHLORIDE_max	0.0170	CREATININE_slope	0.0217
7	PLATELET_last	0.0223	SysBP_last	0.0168	BUN_slope	0.0198
8	HeartRate_max	0.0213	HeartRate_avg	0.0164	Glucose_min	0.0198
9	Temp_avg	0.0213	Temp_last	0.0156	DiasBP_last	0.0198
10	RespRate_avg	0.0182	HeartRate_max	0.0154	age	0.0178