

Predicting Inpatient Acute Kidney Injury over Different Time Horizons: How Early and Accurate?

Peng Cheng, PhD^{1,3}, Lemuel R. Waitman, PhD¹, Yong Hu, PhD², Mei Liu, PhD¹

¹University of Kansas Medical Center, Department of Internal Medicine, Division of Medical Informatics, Kansas City, KS, USA; ²Jinan University, Big Data Decision Institute, Guangzhou, China; ³Southwest University, School of Computer & Information Science, Chongqing, China

ABSTRACT

Incidence of Acute Kidney Injury (AKI) has increased dramatically over the past two decades due to rising prevalence of comorbidities and broadening repertoire of nephrotoxic medications. Hospitalized patients with AKI are at higher risk for complications and mortality, thus early recognition of AKI is crucial. Building AKI prediction models based on electronic medical records (EMRs) can enable early recognition of high-risk patients, facilitate prevention of iatrogenically induced AKI events, and improve patient outcomes. This study builds machine learning models to predict hospital-acquired AKI over different time horizons using EMR data. The study objectives are to assess (1) whether early AKI prediction is possible; (2) whether information prior to admission improves prediction; (3) what type of risk factors affect AKI prediction the most. Evaluation results showed a good cross-validated AUC of 0.765 for predicting AKI events 1-day prior and adding data prior to admission did not improve model performance.

INTRODUCTION

Acute Kidney Injury (AKI) is a common clinical event among hospitalized patients, affecting 10% to 15% of all hospitalized patients and >50% of patients in intensive care units (ICUs)¹⁻³. AKI is most easily detected on the basis of the acute sustained rise of serum creatinine (SCr). Using a large consecutive sample of 19,982 adults, Chertow et al. found that an increase in SCr of ≥ 0.5 mg/dl is associated with a 6.5-fold increase in the odds of death, a 3.5-day increase in length of stay, and nearly \$7,500 in excess hospital cost⁴. Over the past two decades, the incidence of AKI has increased significantly in North American and Europe, particularly within the United States, because of the rising prevalence of acute and chronic conditions, such as sepsis, heart failure, and diabetes⁵⁻⁹. Moreover, the role of medications in the changing epidemiology of AKI has increased from 7% to 16% over a 17-year span because of the wide availability of potential nephrotoxic drugs^{2, 10}.

Unfortunately, the current care of patients with AKI is suboptimal characterized by numerous deficiencies and systematic failings that may be avoidable¹¹. In a 2009 review of the care of patients who died in hospital with a primary diagnosis of AKI, an unacceptable delay in the recognition occurred in 43% of the patients and 20% of the cases could have been prevented with early detection¹². It is however difficult for clinicians to recognize at-risk patients prior to their AKI episodes. Once an AKI episode occurs, there is no treatment to mitigate or cure AKI¹³⁻¹⁶. According to the International Society of Nephrology¹⁷, recognizing patients at risk of developing AKI and managing these patients according to their susceptibilities and exposures is likely to result in better outcomes than merely treating the established AKI. Therefore, the ability to predict AKI in hospitalized patients and monitor them at an early stage is crucial to AKI prevention.

Safety tools based on electronic medical records (EMRs) for in-hospital AKI surveillance covering kidney injury triggers have been developed; examples include the Global Trigger Tool¹⁸, a tool that uses the Acute Kidney Injury Network (AKIN) definition to monitor AKI¹⁹, and dosing tools for improving compliance with renal-dosing of medications^{20, 21}. EMR-based AKI monitoring can expedite interventions and lead to a high percentage of patients retaining their baseline kidney function^{22, 23}. However, as the major limitation of these tools, physicians can only react after observing signs of damage. By contrast, risk prediction would recognize high-risk patients for tailored early management. Most existing AKI risk prediction models focus on predicting adverse outcomes following AKI²⁴ or predicting AKI after specific surgeries and interventions^{1, 3, 25, 26}. There exists some predictive modeling work performed on the critical care populations²⁷⁻²⁹, but much less work on the general inpatient population³⁰. Matheny et al³¹ proposed the first study to predict general in-hospital AKI using logistic regression models. Recently, Kate et al³² built machine learning models to examine the difference in prediction vs. detection of AKI in hospitalized older adults.

This study aims to build machine learning models to predict the development of AKI among general patient hospitalizations at daily intervals prior to the event. The primary objective is to assess how early and accurately the development of inpatient AKI can be predicted. The secondary objective is to assess whether clinical data prior to admission enhance the predictive models.

METHODS

Data Collection

A retrospective cohort of 60,534 patients was collected by including adult admissions (age at visit between 18 and 64) to a tertiary care, academic hospital (the University of Kansas Medical Center – KUMC) from November 2007 to March 2016 with a length of stay of at least 2 days. Given that a patient may have multiple admissions (encounters) of at least 2 days and develop AKI during one but not another, this study is conducted at the encounter level with the initial cohort of total 109,319 encounters. For each encounter, we queried the KUMC de-identified clinical data repository HERON (Health Enterprise Repository for Ontological Narration)^{33, 34} that integrated electronic health records, billing, clinical registries, and national data sources to obtain structured data on admission and discharge dates, patient demographics, medications, laboratory values, vitals, comorbidities and admission diagnosis.

Cohort Inclusion/Exclusion Criteria

From the initial cohort of 109,319 encounters, we selected an analysis cohort of encounters by excluding those (a) missing necessary data for outcome determination – less than two serum creatinine measurements and (b) had evidence of moderate or severe kidney dysfunction – estimated Glomerular Filtration Rate (eGFR) less than 60 mL/min/1.73 m² or abnormal serum creatinine (SCr) level of >1.3 mg/dL within 24 hours of hospital admission. The final analysis cohort consisted of 48,955 encounters (33,703 patients).

AKI Definition

AKI was defined using the Kidney Disease Improving Global Outcomes (KDIGO)-based modifications of the AKIN and Risk, Injury, Failure, Loss, and End-Stage (RIFLE) Kidney classification criteria³⁵. According to KDIGO, adults who demonstrate any of the following are undergoing an AKI episode:

- Increase in SCr by ≥ 0.3 mg/dL (≥ 26.4 micromol/L) within 48 hours
- Increase in SCr by ≥ 1.5 times the baseline within the previous 7 days

The baseline creatinine level was set as either the last measurement within 2-day time window prior to admission or the first available measurement during the stay. All creatinine measurements between admission and discharge were evaluated to determine the occurrence of in-hospital AKI. Based on the above AKI definition, this study classifies each encounter as ‘with AKI’ (positive) or ‘without AKI’ (negative). Out of total 48,955 encounters in the final analysis cohort, patients acquired AKI during 4,405 (8.99%) encounters.

AKI Risk Factors

A list of clinical variables used in building the AKI prediction models is described in Table 1. We referred to Matheny et al³¹ to select laboratory tests that may represent potential presence of a comorbidity that is correlated with in-hospital AKI. For example, an elevated white blood cell count (WBC) is associated with bacterial infection that may cause AKI. Serum creatinine was not included as a variable as it was used to determine the positive and negative samples. For laboratory tests and vitals, only the last recorded value before a prediction point was used and their values were categorized. Laboratory values were categorized as either “present and normal”, “present and abnormal”, or “unknown” according to standard reference ranges. Vitals were categorized into groups as shown in Table 2. Missing values in vitals and lab tests were captured as “unknowns” because information may be contained in the choice to not perform the measurement.

Medication variable included inpatient (i.e., dispensed during stay) and outpatient medications (i.e., historical meds). All medication names were normalized by mapping to RxNorm ingredient. Comorbidity and admission diagnosis, i.e., all patient refined diagnosis related group (APR-DRG) variables were collected from the University Healthsystem Consortium (UHC)³⁶ data source in HERON. Comorbidity, medication, and admission diagnosis variables took either “yes” or “no” values.

Table 1. Clinical variables considered in building predictive models for hospital-acquired AKI

Feature Category	# of Variables	Details
Demographics	3	Age, gender, race
Vitals	5	BMI, diastolic BP, systolic BP, pulse, temperature
Lab tests	14	Albumin, ALT, AST, Ammonia, Blood Bilirubin, BUN, Ca, CK-MB, CK, Glucose, Lipase, Platelets, Troponin, WBC
Comorbidities	29	UHC comorbidity
Admission diagnosis	315	UHC APR-DRG
Medications	1682	All medications are mapped to RxNorm ingredient

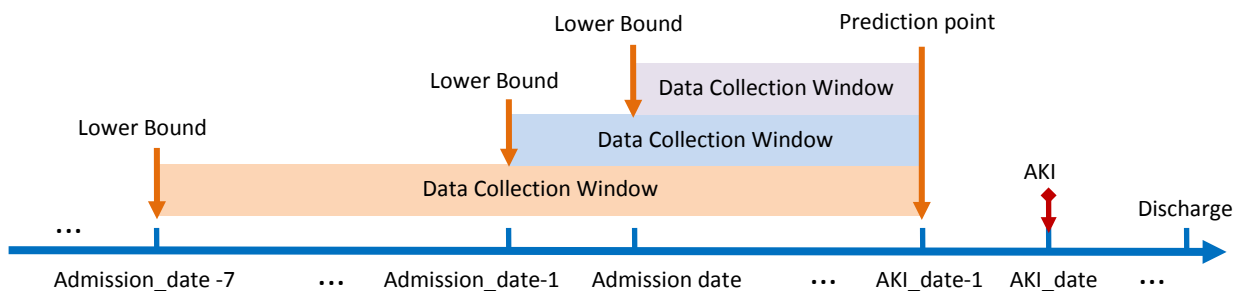
In the final dataset, vitals, lab test, and medication variables were time-stamped (with resolution to the hour and minute) relative to the admission date, referred here as time-dependent variables. Comorbidities, admission diagnosis, and demographics were presumed to be available at admission and not time-dependent.

Table 2. Categories for vital variable categories

Vitals	Categories
BMI	<18.5, [18.5–24.9], [25.0–29.9], >30.0, Unknown
Diastolic BP	<80, [80–89], [90–99], >100, Unknown
Systolic BP	<120, [120–139], [140–159], >160, Unknown
Pulse	<50, [50–65], [66–80], [81–100], >100, Unknown
Temperature	<95.0, [95.0–97.6], [97.7–99.5], [99.5–104.0], >104.0, Unknown

Evaluation Design

Model evaluation was designed to answer three specific questions: (1) Will data prior to admission improve predictive models' performance; (2) How early and accurately can AKI be forecasted; (3) How strong each type of risk factors affects the model performance. In this study, we introduced a data collection window, denoted as [*lower bound*, *upper bound*]. For the first question, we assessed model performance by varying the window's lower bound with a fixed max upper bound set at 1-day prior to AKI event, i.e., making AKI prediction 1-day prior using different amount of clinical data (Figure 1). The initial data collection window's lower bound is set at admission and we increased its width with resolution to the day.

**Figure 1.** An illustration of adjusting the lower bound of data collection window before the admission date

For the second question, we assessed model performance by lengthening the interval between the prediction time and AKI event (Figure 2). For AKI encounters, AKI event date was used as an anchor for moving the prediction point. For non-AKI encounters, discharge date was used as the anchor point.

To illustrate with an example, patient *X* was admitted to the hospital on 2015-06-05 (*Admission_date*) and developed AKI on 2015-06-07 (*AKI_date*), then the initial data collection window for prediction 1-day prior to AKI occurrence would be [*Admission_date*, *AKI_date* - 1] which is [2015-06-05, 2015-06-06]. For a non-AKI example, patient *Y*

stayed at the hospital between 2014-02-15 (*Admission_date*) and 2014-02-18 (*Discharge_date*), then the data collection window for prediction 1-day prior would be [2014-02-15, 2014-02-17]. To assess whether adding data 1-day prior to admission would improve predictive performance, the data collection window becomes [*Admission_date* - 1, *AKI_date* - 1]; thus for patient *X*, it becomes [2015-06-04, 2015-06-06]. To assess whether an accurate prediction can be made 2-days prior to AKI, the window upper bound will be *AKI_date* - 2. In conclusion, there are two parameters to adjust for evaluation: *Admission_date* - *m* is the *m*th day before the hospital admission date and *AKI_date* - *n* is the *n*th day before AKI occurrence date.

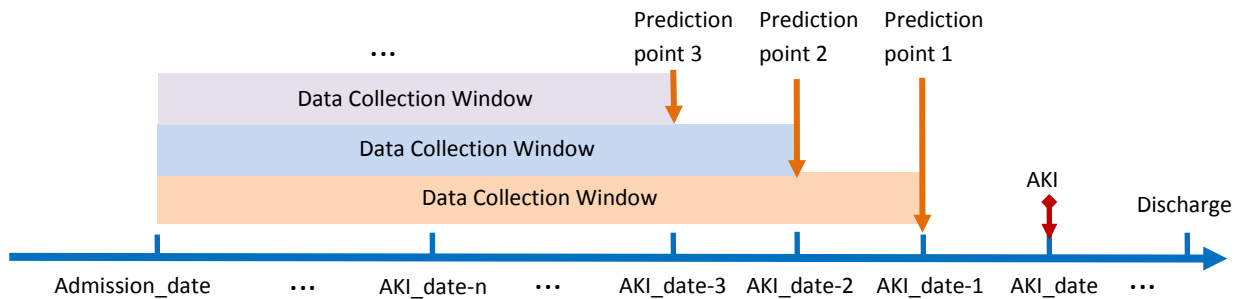


Figure 2. An illustration of adjusting the prediction point (i.e., upper bound of data collection window) before the AKI occurrence date

For the third question, we adopted two approaches to evaluate the strength of each group of risk factors (i.e., demographics, labs, vitals, medications, and comorbidities) used in the model. In the first approach, we removed one group of attributes and trained the predictive models on the remaining four groups of attributes, resulting in five transformed datasets. In the second approach, we trained the predictive models with only one group at a time. Thus, another five transformed datasets were obtained, each of which only contains one group of attributes.

Experimental Methodology

Three different machine learning methods – Logistic Regression³⁷, Random Forest³⁸, and AdaboostM1³⁹ were used for building predictive models using the Weka software library⁴⁰. The number of decision trees in Random Forest model was 100. For AdaBoostM1, the iteration number was 500, using DecisionStump as the base learner. All models were evaluated using the standard 10-fold cross-validation.

The area under the receiver operating characteristic curve (AUC), precision and recall are used to report and compare performance of the models. AUC provides a single measurement of the performance of an ROC curve, which is a graphical plot of the sensitivity or true positive rate against the false positive rate (1 - specificity). Sensitivity is the proportion of actual positives that are correctly identified as such (i.e. $SN = TP/(TP+FN)$) and specificity measures the proportion of actual negatives that are correctly predicted as such (i.e. $SP = TN/(TN+FP)$). Precision is the proportion of true positives against all predicted positive results (i.e. $P = TP/(TP+FP)$). Recall is the same as the true positive rate or sensitivity.

Our datasets are highly imbalanced with an approximate 1:10 positive (AKI) to negative (non-AKI) ratio. With such an imbalanced dataset, most classifiers will favor the majority class (non-AKI) because they are designed to maximize the overall number of correct predictions, thus resulting in poor accuracy in the minority class (AKI) prediction. Current state-of-art correction techniques to account for class imbalance are generally data-based and algorithm-based approaches⁴¹. The data-based approach uses sampling technique by either under-sampling the majority class or over-sampling the minority class. The algorithm-based approach modifies the classification algorithm such as through adjusting decision threshold. In this study, we under-sampled the majority class by randomly selecting a subset of non-AKI samples such that AKI to non-AKI sample ratio is 1:1. Hence a model performing better than random classifier must achieve an AUC larger than 0.5.

RESULTS

Data Characteristics

Distribution of patient demographic variables in AKI and non-AKI encounters is listed in Table 3. In our cohort, the odds ratio between AKI and non-AKI is not significant for all age groups; however, it does increase with age which is also observed in Matheny et al³¹. Odds ratios for gender and race are also within similar range as reported³¹.

Table 3. Distribution of patient demographic variables in AKI and non-AKI encounters

Demographics	AKI (n = 4405)	Non-AKI (n = 44550)	Odds Ratio (95% CI)
Age, n(%)			
18-25	291 (6.6)	3889 (8.7)	0.74 (0.65 – 0.84)
26-35	519 (11.8)	6612 (14.8)	0.77 (0.70 – 0.84)
36-45	734 (16.7)	7736 (17.4)	0.95 (0.88 – 1.03)
46-55	1296 (29.4)	12863 (28.9)	1.03 (0.96 – 1.10)
56-64	1565 (35.5)	13450 (30.2)	1.27 (1.19 – 1.36)
Gender, n (%)			
Female	1779 (40.4)	20533 (46.1)	0.79 (0.74 – 0.84)
Male	2626 (59.6)	24017 (53.9)	1.26 (1.18 – 1.34)
Race, n (%)			
White	3133 (71.1)	32680 (73.4)	0.89 (0.84 – 0.96)
African American	739 (16.8)	6771 (15.2)	1.12 (1.04 – 1.22)
Asian	31 (0.7)	396 (0.9)	0.79 (0.55 – 1.14)
Other	502 (11.4)	4703 (10.5)	1.09 (0.99 – 1.20)

Number of encounters in which AKI occurred in different number of days from the time of admission is shown in Table 4. The largest proportion (23.8%) of AKI encounters occurred on the 1st day after hospitalization.

Table 4. Number of encounters in which AKI occurred within different intervals from time of admission

Days after Admission	Number of AKI events (%)
1	1047 (23.8)
2	959 (21.8)
3	554 (12.6)
4	405 (9.2)
5	296 (6.7)
>6	1144 (25.9)

Results of AKI Prediction

Table 5 and Table 6 show model performance in terms of AUC values and precision and recall respectively over data collected both after and before hospital admissions. Table 7 and Table 8 show the results for evaluating early prediction.

Table 5. AUC values of prediction models on data collected before and after hospital admission

Data Collection Window	Classification Models		
	Random Forest	AdaBoostM1	Logistic
[Admission_date, AKI_date-1]	0.765	0.751	0.763
[Admission_date-1, AKI_date-1]	0.747	0.738	0.732
[Admission_date-7, AKI_date-1]	0.747	0.739	0.733
[Admission_date-15, AKI_date-1]	0.742	0.742	0.733
[Admission_date-30, AKI_date-1]	0.747	0.742	0.732

Table 6. Precision and recall of prediction models on data collected before and after hospital admission

Data Collection Window	Classification Models		
	Random Forest (Precision/Recall)	AdaBoostM1 (Precision/Recall)	Logistic (Precision/Recall)
[Admission_date, AKI_date-1]	0.692 / 0.711	0.662 / 0.736	0.704 / 0.711
[Admission_date-1, AKI_date-1]	0.675 / 0.695	0.661 / 0.698	0.690 / 0.686
[Admission_date-7, AKI_date-1]	0.679 / 0.688	0.660 / 0.694	0.689 / 0.691
[Admission_date-15, AKI_date-1]	0.671 / 0.685	0.666 / 0.693	0.690 / 0.692
[Admission_date-30, AKI_date-1]	0.675 / 0.689	0.664 / 0.694	0.690 / 0.692

The time window $[Admission_date-m, AKI_date-n]$ means that data are collected based on the following rule: the time dependent variables have at least one value available after $Admission_date - m$ and before $AKI_date - n$. $Admission_date - m$ is the m^{th} day before the hospital admission and $AKI_date - n$ is the n^{th} day before AKI occurrence date.

Table 7. AUC of prediction models when adjusting the prediction points/upper bound of data collection window

Data Collection Window	Classification Models		
	Random Forest	AdaBoostM1	Logistic
$[Admission_date, AKI_date-1]$	0.765	0.751	0.763
$[Admission_date, AKI_date-2]$	0.733	0.727	0.731
$[Admission_date, AKI_date-3]$	0.709	0.705	0.691
$[Admission_date, AKI_date-4]$	0.688	0.690	0.651
$[Admission_date, AKI_date-5]$	0.670	0.678	0.633

Table 8. Precision and recall when adjusting the prediction points/ upper bounds of data collection window

Data Collection Window	Classification Models		
	Random Forest (Precision/Recall)	AdaBoostM1 (Precision/Recall)	Logistic (Precision/Recall)
$[Admission_date, AKI_date-1]$	0.692 / 0.711	0.662 / 0.736	0.704 / 0.711
$[Admission_date, AKI_date-2]$	0.675 / 0.661	0.643 / 0.714	0.678 / 0.675
$[Admission_date, AKI_date-3]$	0.650 / 0.650	0.625 / 0.687	0.651 / 0.646
$[Admission_date, AKI_date-4]$	0.634 / 0.637	0.628 / 0.656	0.620 / 0.623
$[Admission_date, AKI_date-5]$	0.623 / 0.610	0.616 / 0.674	0.608 / 0.605

Table 9. The performance of Random Forest model on cohort data by removing one attribute group (Data Collection Window: $[Admission_date, AKI_date-1]$)

Removed attributes group	Precision	Recall	AUC
Demographics	0.692	0.701	0.762
Vitals	0.685	0.699	0.756
Labs	0.678	0.701	0.755
Admission DRGs and Comorbidities	0.667	0.681	0.723
Medications	0.629	0.622	0.679

Table 10. The performance of Random Forest model on cohort data containing only one attribute group (Data Collection Window: $[Admission_date, AKI_date-1]$)

Reserved attributes group	Precision	Recall	AUC
Demographics	0.538	0.618	0.551
Vitals	0.550	0.542	0.558
Labs	0.587	0.211	0.546
Admission DRGs and Comorbidities	0.612	0.619	0.657
Medications	0.659	0.659	0.712

Table 9 and Table 10 show the experimental results of the RandomForest model on processed data by removing one group of attributes and by reserving one group of attributes respectively. These two series of experiments were performed to assess the range of the affect size of each risk factor group on prediction.

DISCUSSION

In this study, we built machine learning based AKI prediction models using structured EMR data for patients admitted to a hospital. Experimental results showed a good cross-validated discrimination performance, best AUC of 0.76 achieved by Random Forest for predicting AKI event 1-day prior, which was similar to other AKI risk stratification models that have been created for general inpatients (0.75)³¹, elderly (0.66)³², more specific clinical scenarios (0.74 to 0.77), and patients who have undergone coronary artery bypass grafting (0.72 to 0.81)^{1, 3, 25, 26}. To ensure that the classifier output is not affected by the different splits generated by the 10-fold cross-validation, we

examined variance of the AUC, precision, and recall measures by applying the Random Forest model on dataset with a collection window of $[Admission_date, AKI_date - 1]$. The standard deviations of AUC, precision and recall are 0.01, 0.01, and 0.02 respectively, which shows that there was no great variance in the 10-fold cross-validation.

For the primary objective, we assessed how early and accurately general inpatient AKI can be predicted. Results in Tables 7 and 8 showed that model performance indeed degrades as the time window between the prediction time and AKI event time lengthened, from Random Forest’s AUC of 0.76 at 1-day prior to 0.67 at 5-days prior. While comparing different machine learning algorithms, results in Table 7 showed that Random Forest achieved the best cross-validated AUC of 0.76 for predicting AKI 1-day prior but when the time to event horizon is lengthened to 4 or 5-days prior, AdaBoostM1 had a slightly better AUC than Random Forest. However, in terms of precision and recall in Table 8, Logistic Regression actually had better precision with the same or better recall compared to Random Forest for 1 and 2-day prior, respectively.

In order for a predictive model to be clinically useful, the 15th Acute Dialysis Quality Initiative (ADQI) consensus conference³⁰ in 2016 recommended forecasting AKI events with a horizon of 48 to 72 hours. Although it would be advantageous to predict AKI events as early as possible, lengthening the prediction time to event horizon will reduce accuracy and the ADQI consensus group believes that 2 to 3 days would give physicians adequate time to modify practice, optimize hemodynamics, and mitigate potential injury without sacrificing too much in predictive power³⁰. Our study showed that the best performing model with Random Forest can forecast AKI 2-days and 3-days prior with AUC of 0.73 and 0.70, respectively.

For the secondary study objective, we assessed whether adding data prior to admission would improve model performance where the data collection window was extended to 1, 7, 15, and 30 days before admission. One would intuitively think more data is better, but the contrary was observed in results (Table 5 and 6). As the study cohort contains encounters with various length of stay and the number of days AKI occurs relative to admission in the positive samples varies greatly from 1 day to 359 days, we suspect additional data prior to admission may impact model performance differently for patients who develop AKI on the day after admission vs. five days after because amount of data available during stay is dramatically different. Therefore, we conducted an analysis on encounters in which AKI occurred 1 day after admission, comparing Random Forest’s performance with (i.e., 1 and 7 days) and without data prior to admission. Interestingly, analysis results on this sub-cohort exhibited the same trend as observed in the complete cohort, where AUC for only using data after admission is 0.84 vs. AUC for adding data from 1-day before admission is 0.81 and 7-days prior is 0.80. This implies adding data prior to admission does not improve AKI prediction performance in the general inpatient AKI population. This prompts us to further analyze the impact of using data from the entire encounter vs. only previous day on prediction performance.

We further screened data with the time window $[AKI_date-1, AKI_date-1]$, i.e. using data on the day of prediction and compared the performance with using data collected within $[Admission_date, AKI_date-1]$. Based on results in Table 11, it seems to imply that prediction using the most recent one-day data can improve AUC and precision at the expense of degrading recall ratio. F-score decreased from 0.701 to 0.687 in contrast to the AUC increase from 0.765 to 0.783. However, this may also suggest the models were overfitting the temporal clinical variables relative to the static demographics, admission diagnoses, and comorbidities. Evaluating the contribution of each clinical variable and refining their representations are future directions for this research.

Table 11. Comparing performance (AUC/Precision/Recall) on entire hospitalization data vs. recent one-day data

Data Collection Window	Random Forest (AUC/Precision/Recall)	AdaBoostM1 (AUC/Precision/Recall)	Logistic (AUC/Precision/Recall)
$[Admission_date, AKI_date-1]$	0.765 / 0.692 / 0.711	0.751 / 0.662 / 0.736	0.763 / 0.704 / 0.711
$[AKI_date-1, AKI_date-1]$	0.783 / 0.721 / 0.655	0.768 / 0.674 / 0.720	0.768 / 0.709 / 0.709

For the third objective, we assessed the effect size of each risk factor type in AKI prediction by applying the Random Forest model on datasets with a collection window of $[Admission_date, AKI_date - 1]$. Based on results in Table 9 and Table 10, medications play the biggest role in the 1-day prior AKI prediction performance followed by the combination of admission DRG and comorbidity. This is promising as medications are modifiable and clinicians may consider alternative therapies. The demographics variables had the least effect on prediction performance, which may be due to the fact that we limited our cohort to a younger cohort of 18 to 64 year olds and utilized encounter-level rather than patient-level data for prediction. A patient may have multiple hospital encounters that satisfy the cohort inclusion criteria and he/she may experience AKI in one encounter and not in another. Thus, the

same patient can belong to both the AKI and non-AKI class, making it difficult for algorithms to distinguish the two classes based on the patient-level demographic information.

Limitations

There are several limitations in the interpretation of results in this study. First of all, the predictive models were based off a younger cohort (18 to 64 years old at admission), which may not be generalizable to an older cohort. Elderly is known to be at increased risk for AKI due to longer exposure to chronic diseases and nephrotoxins⁴², thus our future studies will conduct independent subpopulation analysis for the elderly. Second, we limited the analysis to patients who were admitted to the hospital with a minimum eGFR of 60 mL/min/1.73m² and must have normal serum creatinine on the day of admission. Although patients with reduced eGFR are at increased risk for AKI, it is difficult to determine which of these patients had hospital-acquired vs. community-acquired AKI. Third, comorbidity data utilized in the predictive models were obtained from UHC, which is widely known to be well adjudicated, but not immediately available. This study treated UHC comorbidities as non-time dependent as if the clinical team would know all comorbidities at admission, which may misrepresent comorbidities that developed during the admission. Future studies will evaluate the performance of diagnosis codes from the EMR problem list relative to comorbidity information derived from billing systems. This may change the results we observed in this study that data prior to admission does not improve performance and also more accurately represent the information available to the clinical team to support adverse event surveillance. Last but not least, the study did not use urine output to define AKI nor include it as a risk variable. Although urine output is one of the diagnostic criteria of AKI, many members of the Acute Kidney Injury Network (AKIN) concerned that urine output is not specific enough for the designation of AKI because it can be influenced by factors other than renal health.

CONCLUSIONS

Predicting AKI early and accurately allows clinicians to take timely preventative or therapeutic measures. This study investigates the impact of data completeness and prediction time points on the performance of forecasting in-hospital AKI in a general inpatient population. Three machine learning algorithms, Random Forest, AdaBoostM1, and Logistic Regression, were built on clinical datasets screened with different data collection windows. The Random Forest classifier outperformed Logistic Regression and AdaBoostM1 with AUC values for 1, 2, and 3-days prior being 0.765, 0.733, and 0.709, respectively. In this study, the data prior to hospital admission did not improve prediction performance and medication played the biggest role in prediction performance.

ACKNOWLEDGEMENT

The HERON data repository described in the paper is supported by institutional funding from the University of Kansas Medical Center and CTSA grant UL1TR000001 from NCRR/NIH. The work is partially supported by PCORI CDRN-15-26643.

REFERENCES

1. Fortescue EB, Bates DW, Chertow GM. Predicting acute renal failure after coronary bypass surgery: cross-validation of two risk-stratification algorithms. *Kidney international*. 2000 Jun;**57**(6):2594-602.
2. Hou SH, Bushinsky DA, Wish JB, Cohen JJ, Harrington JT. Hospital-acquired renal insufficiency: a prospective study. *The American journal of medicine*. 1983 Feb;**74**(2):243-8.
3. Wijeyesundera DN, Karkouti K, Dupuis JY, et al. Derivation and validation of a simplified predictive index for renal replacement therapy after cardiac surgery. *JAMA*. 2007 Apr 25;**297**(16):1801-9.
4. Chertow GM, Burdick E, Honour M, Bonventre JV, Bates DW. Acute kidney injury, mortality, length of stay, and costs in hospitalized patients. *Journal of the American Society of Nephrology : JASN*. 2005 Nov;**16**(11):3365-70.
5. Liano F, Junco E, Pascual J, Madero R, Verde E. The spectrum of acute renal failure in the intensive care unit compared with that seen in other settings. The Madrid Acute Renal Failure Study Group. *Kidney international Supplement*. 1998 May;**66**:S16-24.
6. Metnitz PG, Krenn CG, Steltzer H, et al. Effect of acute renal failure requiring renal replacement therapy on outcome in critically ill patients. *Crit Care Med*. 2002 Sep;**30**(9):2051-8.
7. Turney JH, Marshall DH, Brownjohn AM, Ellis CM, Parsons FM. The evolution of acute renal failure, 1956-1988. *The Quarterly journal of medicine*. 1990 Jan;**74**(273):83-104.
8. Ympa YP, Sakr Y, Reinhart K, Vincent JL. Has mortality from acute renal failure decreased? A systematic review of the literature. *The American journal of medicine*. 2005 Aug;**118**(8):827-32.

9. Siew ED, Davenport A. The growth of acute kidney injury: a rising tide or just closer attention to detail? *Kidney international*. 2015 Jan;**87**(1):46-61.
10. Nash K, Hafeez A, Hou S. Hospital-acquired renal insufficiency. *American journal of kidney diseases : the official journal of the National Kidney Foundation*. 2002 May;**39**(5):930-6.
11. Aitken E, Carruthers C, Gall L, Kerr L, Geddes C, Kingsmore D. Acute kidney injury: outcomes and quality of care. *QJM : monthly journal of the Association of Physicians*. 2013 Apr;**106**(4):323-32.
12. Stewart J, Findlay G, Smith N, Kelly K, Mason M. Adding Insult to Injury: A review of the care of patients who died in hospital with a primary diagnosis of acute kidney injury (acute renal failure). London, UK: National Confidential Enquiry into Patient Outcome and Death; 2009.
13. Sutherland SM, Ji J, Sheikhi FH, et al. AKI in hospitalized children: epidemiology and clinical associations in a national cohort. *Clinical journal of the American Society of Nephrology : CJASN*. 2013 Oct;**8**(10):1661-9.
14. Sutherland SM, Byrnes JJ, Kothari M, et al. AKI in hospitalized children: comparing the pRIFLE, AKIN, and KDIGO definitions. *Clinical journal of the American Society of Nephrology : CJASN*. 2015 Apr 7;**10**(4):554-61.
15. Mammen C, Al Abbas A, Skippen P, et al. Long-term risk of CKD in children surviving episodes of acute kidney injury in the intensive care unit: a prospective cohort study. *American journal of kidney diseases : the official journal of the National Kidney Foundation*. 2012 Apr;**59**(4):523-30.
16. Chawla LS, Eggers PW, Star RA, Kimmel PL. Acute kidney injury and chronic kidney disease as interconnected syndromes. *The New England journal of medicine*. 2014 Jul 3;**371**(1):58-66.
17. KDIGO. KDIGO Clinical Practice Guideline for Acute Kidney Injury. *Kidney international Supplement*. 2012;**2**(1):1-141.
18. Resar RK, Rozich JD, Simmonds T, Haraden CR. A trigger tool to identify adverse events in the intensive care unit. *Joint Commission journal on quality and patient safety / Joint Commission Resources*. 2006 Oct;**32**(10):585-90.
19. Selby NM, Crowley L, Fluck RJ, et al. Use of electronic results reporting to diagnose and monitor AKI in hospitalized patients. *Clinical journal of the American Society of Nephrology : CJASN*. 2012 Apr;**7**(4):533-40.
20. Chertow GM, Lee J, Kuperman GJ, et al. Guided medication dosing for inpatients with renal insufficiency. *JAMA*. 2001 Dec 12;**286**(22):2839-44.
21. McCoy AB, Waitman LR, Gadd CS, et al. A computerized provider order entry intervention for medication safety during acute kidney injury: a quality improvement report. *American journal of kidney diseases : the official journal of the National Kidney Foundation*. 2010 Nov;**56**(5):832-41.
22. Colpaert K, Hoste E, Van Hoecke S, et al. Implementation of a real-time electronic alert based on the RIFLE criteria for acute kidney injury in ICU patients. *Acta clinica Belgica Supplementum*. 2007(2):322-5.
23. Colpaert K, Hoste EA, Steurbaut K, et al. Impact of real-time electronic alerting of acute kidney injury on therapeutic intervention and progression of RIFLE class. *Crit Care Med*. 2012 Apr;**40**(4):1164-70.
24. McCullough PA, Adam A, Becker CR, et al. Risk prediction of contrast-induced nephropathy. *The American journal of cardiology*. 2006 Sep 18;**98**(6A):27K-36K.
25. Chertow GM, Lazarus JM, Christiansen CL, et al. Preoperative renal risk stratification. *Circulation*. 1997 Feb 18;**95**(4):878-84.
26. Aronson S, Fontes ML, Miao Y, et al. Risk index for perioperative renal dysfunction/failure: critical dependence on pulse pressure hypertension. *Circulation*. 2007 Feb 13;**115**(6):733-42.
27. Uchino S, Kellum JA, Bellomo R, et al. Acute renal failure in critically ill patients: a multinational, multicenter study. *JAMA*. 2005 Aug 17;**294**(7):813-8.
28. Chawla LS, Davison DL, Brasha-Mitchell E, et al. Development and standardization of a furosemide stress test to predict the severity of acute kidney injury. *Critical care*. 2013 Sep 20;**17**(5):R207.
29. Kane-Gill SL, Sileanu FE, Murugan R, Trietley GS, Handler SM, Kellum JA. Risk factors for acute kidney injury in older adults with critical illness: a retrospective cohort study. *American journal of kidney diseases : the official journal of the National Kidney Foundation*. 2015 Jun;**65**(6):860-9.
30. Sutherland SM, Chawla LS, Kane-Gill SL, et al. Utilizing electronic health records to predict acute kidney injury risk and outcomes: workgroup statements from the 15(th) ADQI Consensus Conference. *Canadian journal of kidney health and disease*. 2016;**3**:11.
31. Matheny ME, Miller RA, Ikizler TA, et al. Development of inpatient risk stratification models of acute kidney injury for use in electronic health records. *Medical decision making : an international journal of the Society for Medical Decision Making*. 2010 Nov-Dec;**30**(6):639-50.
32. Kate RJ, Perez RM, Mazumdar D, Pasupathy KS, Nilakantan V. Prediction and detection models for acute kidney injury in hospitalized older adults. *BMC medical informatics and decision making*. 2016 Mar 29;**16**:39.

33. Waitman LR, Warren JJ, Manos EL, Connolly DW. Expressing observations from electronic medical record flowsheets in an i2b2 based clinical data repository to support research and quality improvement. *AMIA Annu Symp Proc.* 2011;**2011**:1454-63.
34. Murphy SN, Weber G, Mendis M, et al. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). *J Am Med Inform Assoc.* 2010 Mar-Apr;**17**(2):124-30.
35. Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group. KDIGO Clinical Practice Guideline for Acute Kidney Injury. *Kidney international Supplement.* **2012**(2):1-138.
36. University Healthsystem Consortium. [cited March 9, 2017]; Available from: <https://www.vizientinc.com>
37. Le Cessie S, Van Houwelingen JC. Ridge Estimators in Logistic Regression *Journal of the Royal Statistical Society Series C (Applied Statistics).* 1992;**41**(1):191-201.
38. Breiman L. Random Forests. *Machine Learning.* 2001;**45**(1):5-32.
39. Freund Y, Schapire RE. Experiments with a new boosting algorithm. *Thirteenth International Conference on Machine Learning.* San Francisco: Morgan Kaufmann; 1996. p. 148-56.
40. Hall M, Frank E, Holmes G, Pfahringer B, Reitemann P, Witten I. The WEKA data mining software: an update. *SIGKDD Explorations.* 2009;**11**(1).
41. Lin WJ, Chen JJ. Class-imbalanced classifiers for high-dimensional data. *Briefings in bioinformatics.* 2013 Jan;**14**(1):13-26.
42. Abdel-Kader K, Palevsky PM. Acute kidney injury in the elderly. *Clinics in geriatric medicine.* 2009 Aug;**25**(3):331-58.