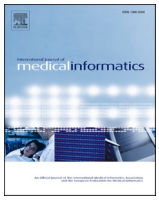
International Journal of Medical Informatics 158 (2022) 104661

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf



Development of a knowledge mining approach to uncover heterogeneous risk predictors of acute kidney injury across age groups

Lijuan Wu [a,b,*], Yong Hu [a,b,*], Xiangzhou Zhang [a,b], Jia Zhang [c], Mei Liu [d,*]

[a] Big Data Decision Institute, Jinan University, Guangzhou 510632, China
[b] Guangdong Engineering Technology Research Center for Big Data Precision Healthcare, Guangzhou 510632, China
[c] The First Affiliated Hospital of Chongqing Medical University, Chongqing 400016, China
[d] Department of Internal Medicine, Division of Medical Informatics, University of Kansas Medical Center, Kansas City 66160, USA

ARTICLE INFO

Keywords:
Knowledge mining approach
Machine learning
Risk heterogeneity, electronic medical records
Acute kidney injury

ABSTRACT

Objectives: Acute kidney injury (AKI) risk increases with age and the underlying clinical predictors may be heterogeneous across age strata. This study aims to uncover the AKI risk factor heterogeneity among general inpatients across age groups using electronic medical records (EMR).

Methods: Patient data (n = 179,370 encounters) were collected from an academic hospital between 2007 and 2016, and were stratified into four age groups: 18–35, 36–55, 56–65, and > 65. Potential risk factors extracted for the cohort included demographics, vital signs, laboratory values, past medical diagnoses, medications and admission diagnoses. We developed a data driven knowledge mining approach consisting of a machine learning algorithm to identify AKI predictors across age strata and a statistical method to quantify the impact of those factors on AKI risk. Identified predictors were evaluated for their predictability of AKI in terms of area-under-the-receiver-operating-characteristic-curve (AUC) and validated against expert knowledge.

Results: Among the final analysis cohort of 76,957 hospital admissions, AKI prediction across age groups 18–35 (16.73%), 36–55 (32.74%), 56–65 (23.52%), and > 65 years (27.01%) achieved AUC of 0.85 (95% CI, 0.80–0.88), 0.86 (95% CI, 0.83–0.89), 0.87 (95% CI, 0.86–0.90), and 0.87 (95% CI, 0.86–0.90), respectively. Compared to expert knowledge, absolute consistency rates of the top-150 identified risk factors for each group were 78.4%, 77.2%, 81.3%, and 79.5%, respectively. Impact of many predictors on AKI varied across age groups; for example, high body mass index (BMI) was found to be associated with higher AKI risk in elderly patients, but low BMI was found to be associated with higher AKI risk in younger patients.

Conclusions: We verified the effectiveness of the knowledge mining method from the perspectives of accuracy, stability and credibility, and used this approach to clarify the heterogeneity of AKI risk factors between age groups. Future decision support systems need to consider such heterogeneity to enhance personalized patient care.

1. Introduction

Acute Kidney Injury (AKI) is a sudden episode of kidney dysfuntion that develops rapidly over a few hours to a few days and is common in patients [1–3]. Thus, identifying strategies to assess patient risk and manage patients according to their susceptibilities and exposures is critical for improving patient care and outcomes [4–6]. Although the understanding of risk factors for the general AKI population may be substantial, the pathogenesis of AKI can differ between older and younger patients [7], which may lead to heterogeneity in the impact of risk factors associated with patients in different age groups.

Electronic medical record (EMR) data-driven approach to risk factor identification has the potential to accelerate hypothesis generation and knowledge discovery. Machine learning methods (e.g., gradient boosting machine [8], random forest [9], linear support vector classification [10] and deep neural network [11–14]) play an important role in data mining and disease prediction. In order to effectively improve doctors' trust in the machine learning models, many interpretable methods have been developed to explain the predictions of the black box models [15], such as SHAP (SHapley Additive exPlanations) [16] and LIME (Local

* Corresponding authors at: Jinan University, 601 Huangpu Road West, Guangzhou 510632, China. (L. Wu). University of Kansas Medical Center, 3001B Student Center, Mail Stop 3065, 3901 Rainbow Boulevard, Kansas City, KS 66160, USA (M. Liu).
E-mail addresses: wulj1989@163.com (L. Wu), henryhu200211@163.com (Y. Hu), meiliu@kumc.edu (M. Liu).

https://doi.org/10.1016/j.ijmedinf.2021.104661
Received 3 July 2021; Received in revised form 21 October 2021; Accepted 5 December 2021
Available online 9 December 2021

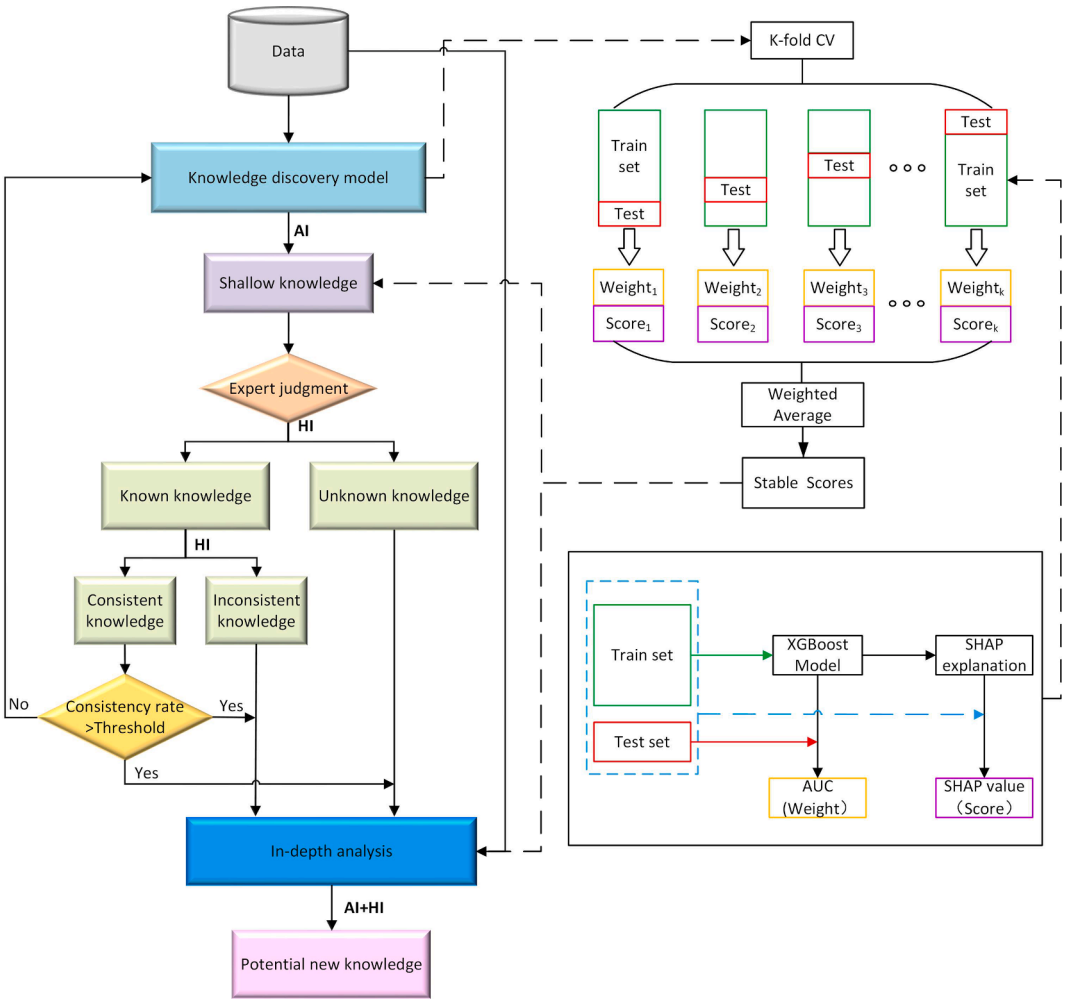


Fig. 1. Knowledge mining approach. (AI: Artificial Intelligence; HI: Human Intelligence).

Interpretable Model-Agnostic Explanation) [17].

To our best knowledge, no study has attempted to design a knowledge mining approach to uncover AKI risk factor differences among age groups using EMR data. In this study, we developed an EMR-data driven and machine-learning based knowledge mining approach to elucidate heterogeneous risk predictors of AKI across age groups and validated its effectiveness with respect to accuracy, stability and credibility.

2. Materials and methods

2.1. Study population

All adult patients (age $\geq$ 18) admitted to a tertiary academic hospital for two or more days from November 2007 to December 2016 were included. This study was conducted at the encounter level with a total of 179,370 inpatient encounters. From those de-identified encounters, we excluded cases (a) missing necessary data for outcome determination, i.e., less than two serum creatinine measurements and (b) that had evidence of moderate or severe kidney dysfunction at admission, i.e., estimated Glomerular Filtration Rate (eGFR) less than 60 mL/min/1.73 m^2 or abnormal serum creatinine (SCr) level of $>$ 1.3 mg/dL on the admission day. The final analysis cohort contained 76,957 encounters (see Fig. C.1). The de-identified data were obtained from the Healthcare Enterprise Resource for Ontological Narration (HERON), an i2b2-based clinical integrated data repository [18]. The operation of HERON as an honest broker research repository was approved by the University of Kansas Medical Center Institutional Review Board (Human Subject Committee) and is reviewed annually (HSC #12337).

2.2. AKI definition

AKI was defined per the KDIGO (Kidney Disease: Improving Global Outcomes) creatinine criteria [19] (see Table C.1). Due to the relative sparsity of urine output data on the general hospital wards, urine output criteria for AKI were not considered. To establish an admission baseline, we used the last SCr value during the 2-day window prior to admission if available; otherwise, the first SCr value after admission was used as the baseline. All SCr levels between admission and discharge were assessed on a rolling basis to determine the occurrence of any AKI.

2.3. Data collection and processing

We collected 1,888 clinical variables from EMR including basic demographics, vital signs, laboratory values, past medical diagnoses, medications and admission diagnoses (see Table C.2). SCr and eGFR were not included as predictors in the model because they determine the outcome variable, AKI vs non-AKI. Admission diagnosis was captured as the all patient refined diagnosis related group (APR-DRG). Past medical diagnoses were represented by mapping ICD-9 diagnosis codes into the Clinical Classifications Software (CCS) major diagnosis categories developed by the Agency for Healthcare Research and Quality. Medication exposure included both inpatient and outpatient medications. Inpatient medication referred to drugs dispensed during hospitalization and outpatient medication included outpatient prescriptions and drugs

Table 1
Clinical variable characteristics extracted for patients by age category.

Variable	Age 18–35 (n = 12,873)		Age 36–55 (n = 25,197)		Age 56–65 (n = 18,098)		Age > 65 (n = 20,789)	
	AKI	NONAKI	AKI	NONAKI	AKI	NONAKI	AKI	NONAKI
Demographics (Demo); 3(No.); details: age, race, gender; n (%)								
Age	983(7.29)	11,890(92.71)	2,222(8.82)	22,975(91.18)	1,906(10.53)	16,192(89.47)	2,193(10.55)	18,596(98.45)
Race								
White	660(67.14)	8,038(67.60)	1,537(69.17)	16,652(72.48)	1,444(75.76)	13,024(80.43)	1,767(80.57)	15,463(83.15)
Black	150(15.26)	1,958(16.47)	420(18.90)	3,808(16.57)	264(13.85)	1,926(11.89)	231(10.53)	1,644(8.84)
Asian	7(0.71)	147(1.24)	12(0.54)	190(0.83)	17(0.89)	112(0.69)	18(0.82)	151(0.81)
Other	121(12.31)	1,792(15.07)	253(11.39)	2325(10.12)	181(9.50)	1130(6.98)	177(8.07)	1,338(7.20)
Male	549(55.85)	6,066(51.02)	1,302(58.60)	12,337(53.70)	1,170(61.39)	9,297(57.42)	1,288(58.73)	10,150(54.58)
Vitals (Vitals); 5(No.); details: BMI, diastolic BP, systolic BP, pulse, temperature; n (%)								
BMI (kg/m^2)								
Unknown	39(4.16)	1209(10.13)	92(4.14)	1848(8.04)	57(2.99)	830(5.13)	65(2.96)	947(5.09)
<18.5	81(8.63)	563(4.72)	56(2.52)	537(2.34)	44(2.31)	479(2.96)	69(3.15)	680(3.66)
18.5–24.9	325(34.65)	4076(34.15)	460(20.70)	5133(22.34)	379(19.88)	3630(22.42)	535(24.40)	5542(29.80)
25.0–29.9	201(21.43)	2594(21.73)	567(25.52)	5879(25.59)	493(25.87)	4454(27.51)	701(31.97)	5950(32.00)
>30.0	292(31.13)	3493(29.27)	1047(47.12)	9578(41.69)	933(48.95)	6799(41.99)	823(37.53)	5477(29.45)
Lab tests (Lab); 14(No.); details: Albumin, ALT, AST, Ammonia, Calcium, BUN, Bilirubin, CK-MB, CK, Glucose, Lipase, Platelets, Troponin, WBC; n (%)								
WBC (10^9/L)								
Unknown	0(0.0)	18(0.15)	3(0.14)	27(0.12)	0(0.0)	22(0.14)	1(0.05)	22(0.12)
<3.5	58(6.18)	559(4.68)	157(7.07)	1354(5.89)	140(7.35)	1015(6.27)	79(3.60)	770(4.14)
3.5–10.5	442(47.12)	8657(72.53)	1129(50.81)	16942(73.74)	958(50.26)	11878(73.36)	1143(52.12)	14461(77.76)
>10.5	438(46.70)	2701(22.63)	933(41.99)	4652(20.25)	808(42.39)	3277(20.24)	970(44.23)	3343(17.98)
Admission diagnoses (DRG); 315(No.); details: University Health System Consortium (UHC) APR-DRG; n (%)								
Cystic fibrosis	121(12.90)	564(4.73)	26(1.17)	125(0.54)	6(0.31)	39(0.24)	0(0.00)	3(0.02)
Liver transplant	9(0.96)	9(0.08)	68(3.06)	49(0.21)	65(3.41)	69(0.43)	15(0.68)	20(0.11)
Medical History (CCS); 280(No.); details: ICD9 codes mapped to CCS major diagnoses; n (%)								
Essential hypertension	104(11.09)	886(7.42)	648(29.16)	5834(25.39)	825(43.28)	5968(36.86)	1053(48.02)	8715(6.86)
Nutritional deficiencies	193(20.58)	1040(8.71)	203(9.14)	1844(8.03)	156(8.18)	1416(8.75)	172(7.84)	1721(9.25)
Medications (MED); 1271(No.); details: All medications are mapped to RxNorm ingredient; n (%)								
Tazobactam	403(42.96)	2089(17.50)	791(35.60)	3742(16.29)	567(29.75)	2662(16.44)	515(23.48)	2892(15.55)
Vancomycin	366(39.02)	1936(16.22)	762(34.29)	4485(19.52)	620(32.53)	3656(22.58)	655(29.87)	489(22.53)
Onset time, days, median [interquartile range] of admission days								
Days	3 [2–6]	–	3[2–5]	–	3[2–6]	–	3[2–6]	—

Note: AKI = acute kidney injury, NONAKI = not acute kidney injury.

taken at home. We used the RxNorm terminology to standardize the drug names.

We preformed the following data processing: (a) most recent vital signs and lab tests were collected and discretized (see Table C.3) (e.g. "*unknown*", "*less than standard value*", "*the standard value*", or "*more than the standard value*"); (b) medical history and admission diagnoses are binary, representing the presence or absence of disease; (c) used the *cumulative-exposure-days* of medications during the 7-day window before the reference point as predictors; (d) performed *one-hot-coding* on categorical variables (e.g., vital signs and lab tests) to convert them into binary representations; and missing values were treated as a separate category to preserve information contained in the absence of measurement [20]; (e) considered the *length of stay* as a predictor; (f) removed variables whose variance is 0 (e.g., features that are all missing information); Finally, based on medical significance and age distribution of the study cohort, we stratified them into four age groups: 18–35 (16.73%), 36–55 (32.74%), 56–65 (23.52%), and > 65 years (27.01%).

2.4. Knowledge mining approach

As shown in Fig. 1, we designed a novel knowledge mining approach combining artificial intelligence (AI) and expert knowledge to systematically mine AKI risk factor differences among age groups using EMR data. There are two main machine learning methods in the framework, namely eXtreme Gradient Boosting (XGBoost) [21] and Tree SHAP method (TreeExplainer) [16] (see Methods A.1 and A.2). XGBoost is a machine learning technique that ensembles many decision trees, and it optimizes the model in a stage-wise fashion for prediction. SHAP is an approach to interpret model predictions that assigns each feature an importance value for a particular prediction. In this study, we used SHAP value to evaluate the marginal effects of the XGBoost model. Moreover, in order to reduce the influence of sample differentiation and enhance the stability and effectiveness of knowledge mining, we used a cross-validation strategy to introduce this data drift [22], then obtained a weighted average SHAP explanation values (*wSHAP*), where the weight was the area under the receiver operating characteristic curve (AUC) [23] of the XGBoost model in each fold and the risk score for each variable was derived from the SHAP interpretation using the entire dataset.

To assess the consistency between the knowledge mined by machine learning methods and expert knowledge, we first categorized features such as risk factors and protection factors according to the *wSHAP* interpretative scores (see Appendix B). For expert knowledge, namely human intelligence (HI), we ask multiple clinician annotators to help evaluate the relationship between the identified factors and the disease, i.e., whether they are risk factors, protective factors, or inability to judge. When the consistency rate of AI knowledge and HI knowledge exceeds a certain threshold (e.g., 50%), we can think that the knowledge mining method is credible and can be further used to explore and mine new potential risk knowledge.

Table 2
The accuracy, stability and credibility performances of knowledge mining approach for AKI age groups.

Evaluation	Method		Age 18–35	Age 36–55	Age 56–65	Age > 65
Accuracy	XGBoost (AUC,95% CI)	Original input samples	0.85 [0.80–0.88]	0.86 [0.83–0.89]	0.87 [0.86–0.90]	0.87 [0.86–0.90]
		Raw SHAP values	0.93 [0.90–0.96]	0.91 [0.88–0.93]	0.93 [0.91–0.95]	0.91 [0.89–0.93]
		Weighted SHAP values	**0.97 [0.95–0.99]**	**0.94 [0.92–0.95]**	**0.96 [0.95–0.97]**	**0.95 [0.94–0.96]**
Stability	Macro: Mean (SD)	Raw SHAP values	0.9571 (0.0101)	0.9800 (0.0048)	0.9808 (0.0034)	0.9879 (0.0023)
		Weighted SHAP values	**0.9977 (0.0005)**	**0.9989 (0.0002)**	**0.9989 (0.0002)**	**0.9993 (0.0001)**
	Micro: Median (IQR)	Raw SHAP values	0.45 (0.25–0.71)	0.47 (0.26–0.75)	0.42 (0.24–0.68)	0.47 (0.28–0.76)
		Weighted SHAP values	**0.75 (0.47–0.92)**	**0.64 (0.40–0.89)**	**0.67 (0.42–0.89)**	**0.70 (0.41–0.91)**
Credibility	*Y*: Yes, consistency		58	61	61	62
	N: No, inconsistency		16	18	14	16
	U: Unknown or cannot judge		76	71	75	72
	$Y/(Y+N)$: absolute consistency rate		**78.4%**	**77.2%**	**81.3%**	**79.5%**

Abbreviation: XGBoost: eXtreme Gradient Boosting; LinearSVC: linear support vector machine classification; CI: confidence interval. AUC: the area under the receiver operating characteristic curve, SD: standard deviation, IQR: interquartile range. SHAP: SHapley Additive exPlanations. The macro perspective represents the stability of feature ranking of the whole samples; the micro perspective represents the stability of the SHAP output of each feature on the entire samples.

2.5. Evaluations and statistical analysis

We verified the effectiveness of the knowledge discovery method from three perspectives: the accuracy of information classification, the stability of information interpretation, and the consistency with expert knowledge. In order to verify the accuracy and reliability of the knowledge mining method in identifying AKI, the AUC was used to compare the AKI distinguishing ability of original data and weighted SHAP explanation scores. Further, we analyze the stability [24] of the explanation knowledge on AKI predictors from both macro and micro perspectives. The macro perspective represents the stability of feature ranking of the whole samples; the micro perspective represents the stability of the SHAP output of each feature on the entire samples. Besides, to verify the effectiveness of introducing cross-validation strategy to reduce the influence of sample differentiation, we compared the stability of the weighted SHAP explanation scores and original SHAP values.

In order to prove the rationality and correctness of the weighted SHAP interpretation values from the perspective of intuitive visualization, we applied t-distributed Stochastic Neighbor Embedding (t-SNE) [25] to visualize the high-dimensional EMR data by giving each sample a location in a two-dimensional space. In terms of credibility, namely the consistency rate of the excavated knowledge against expert knowledge, in order to simplify the problem reasonably, the feature selection based on the wSHAP ranking was used to filter out the most influential and representative feature subsets. In addition, based on the existing research literature and medical guidelines, we further consulted three M.D. physicians (one general internist and two nephrologists) with more than 8 years of clinical and research experience to assess the relationship between the identified factors and AKI. If there is a conflicting assessment, we choose the opinion of the majority (greater than 50%). Two-tailed $p < 0.05$ denoted statistical significance for all comparisons. All data processing and analyses were performed using Python 3.7.

3. Results

3.1. Characteristics of the study cohort

Of the 76,957 encounters in the final analysis cohort, 38,887 (50.53%) patients over the age of 56 years, and 7259 patients (9.43%)

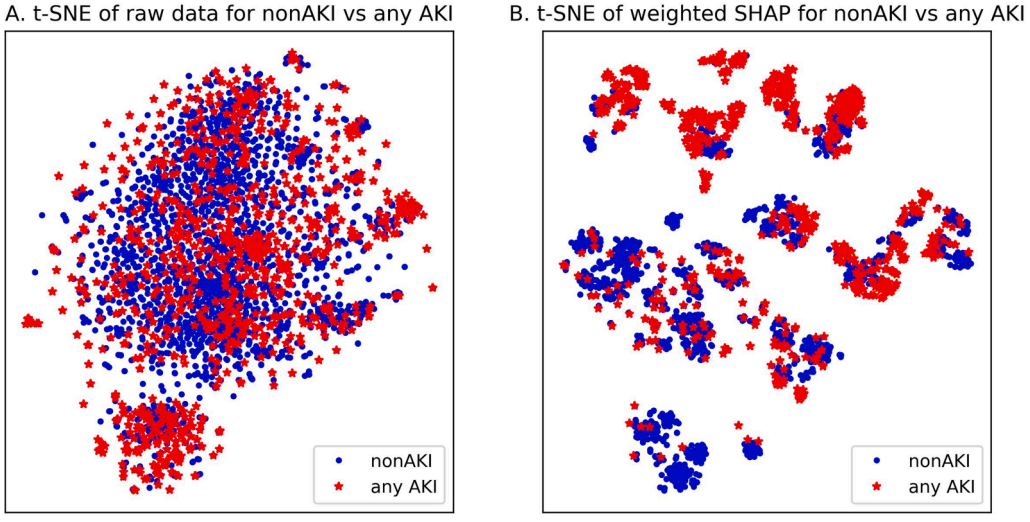


Fig. 2. The t-SNE visualization for age 18–35 group. (A. raw data features and B. the weighted SHAP features).

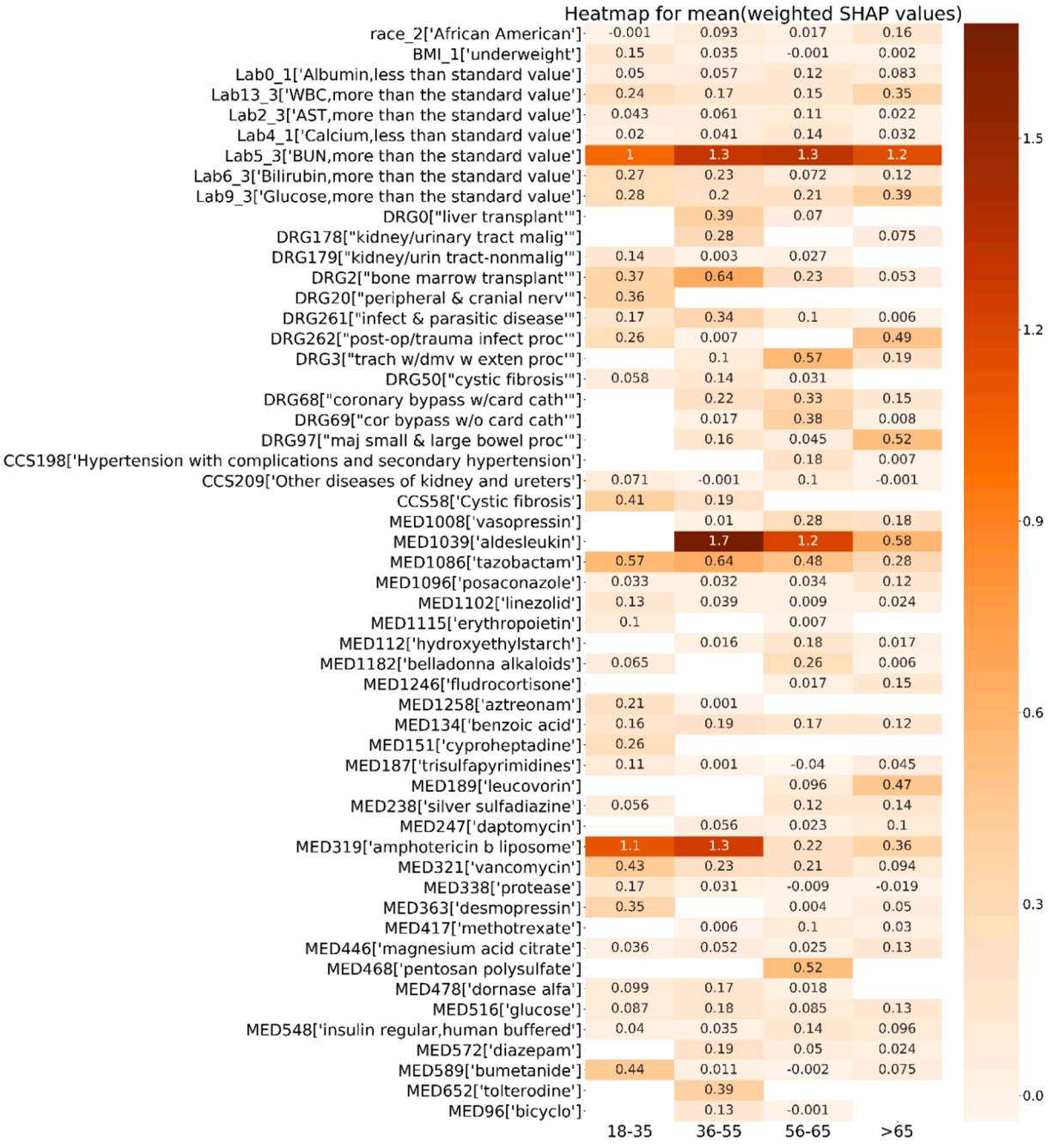


Fig. 3. Heatmap comparison of 54 risk factors in four age groups. (First, 23, 22, 24 and 19 risk factors were screened out from 4 age groups under the condition of $mean(wSHAP) > 0.1$, then 54 features were obtained by combining them. Empty cells in the figure represents a null value, which means that the impact of this factor in this age group is almost 0, i.e., the variable was filtered out during the data preprocessing stage. The higher the weighted SHAP value (darker cell color) the bigger its impact on AKI risk.)

developed some types of AKI. The characteristics of patients by age group were shown in Table 1, where the incidence of AKI increases with age, from 7.29% in the youngest group to 10.55% in the oldest group. In addition, most AKI onset time in terms of the number of days after admission occurred within a week with no significant differences between age groups. After date preprocessing, the final number of variables for 18–35, 36–55, 56–65 and > 65 age groups were 1519, 1691, 1648 and 1631, respectively.

3.2. Evaluation of the knowledge mining approach

Table 2 shows that the accuracy, stability and credibility performances of the knowledge discovery approach in four AKI age groups. For 18–35, 36–55, 56–65 and > 65 age groups, the optimal AUC obtained by the XGBoost model (parameters used are available in Table C.4, see Table C.5) were 0.85 (95% confidence interval (CI) 0.80–0.88), 0.86 (95% CI 0.83–0.89), 0.87 (95% CI 0.86–0.90), and 0.87 (95% CI 0.86–0.90), respectively. The weighted SHAP explanation values got the best performance, namely 0.97 (95% CI, 0.95–0.99), 0.94 (95% CI, 0.92–0.95), 0.96 (95% CI, 0.95–0.97), and 0.95 (95% CI, 0.94–0.96), respectively. Fig. C.2 shows that the cross-validation strategy significantly enhanced the effectiveness of knowledge discovery. Fig. 2 shows the t-SNE visualization of the patient-patient scatter plot for age 18–35 group, we can see that the weighted SHAP features can provide effective explanation information for the identification of AKI.

The weighted SHAP values obtained the higher stability than the raw

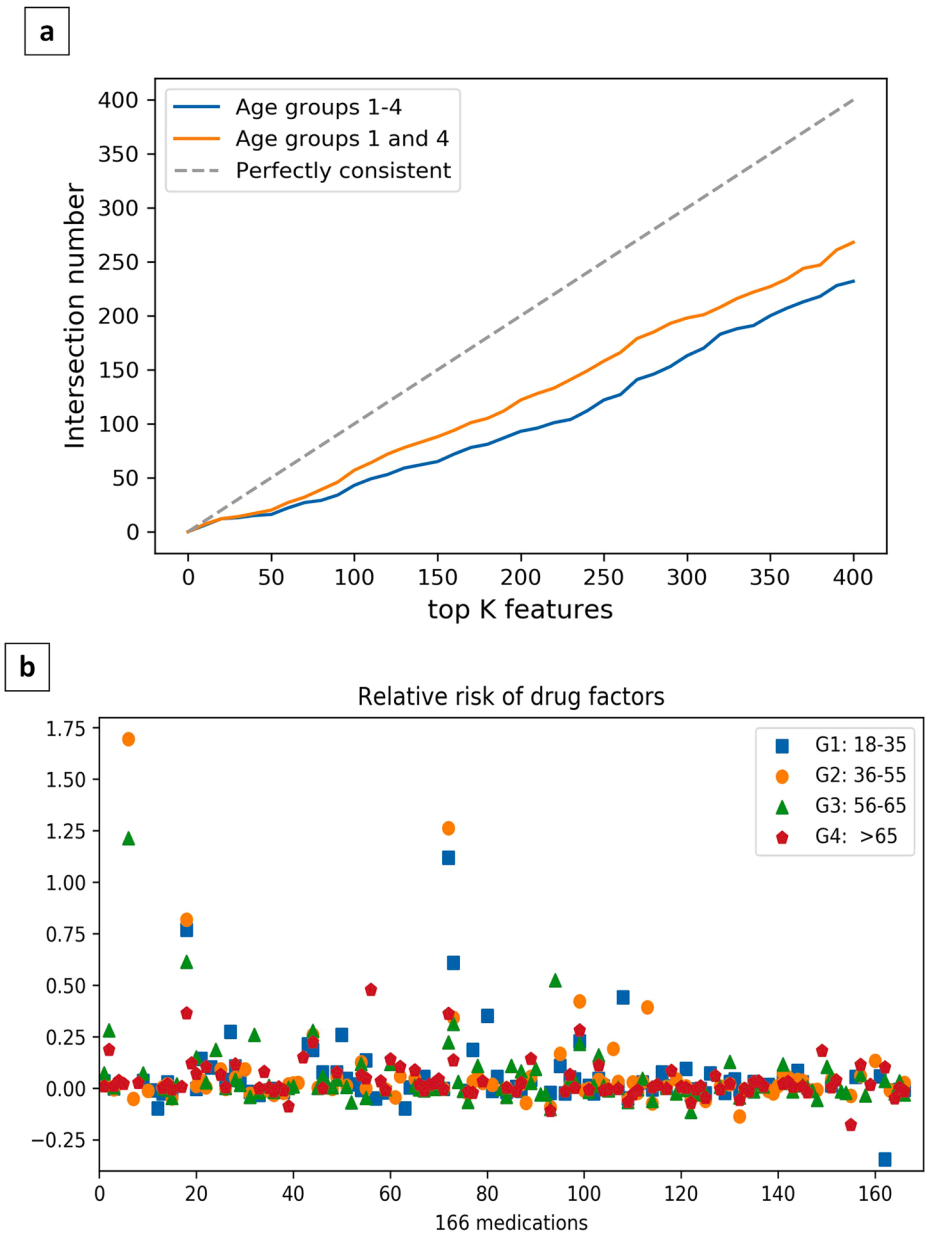


Fig. 4. Explore age group heterogeneity of AKI predictors. (a) The intersection number of top *k* features for AKI between age groups, the blue line represents the intersection number of top k features of the four age groups, the red line represents the intersection number of top k features between the young group (age 18–35) and the old group (age > 65), and the gray dotted line represents the perfectly overlapping features; (b) shows the differences in the relative risks of 166 drugs, which is the union of all drugs selected in the four age groups where their weighted SHAP > 0.01. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

SHAP values (see Table 2), for instance, the micro stability of 18–35, 36–55, 56–65, and > 65 age groups increased from 0.45 to 0.75, 0.47 to 0.64, 0.42 to 0.67, 0.47 to 0.70, respectively. Fig. C.3a-3d shows the comparison of the macro and micro stability between raw SHAP values and weight SHAP values, where the stabilities of weighted SHAP values have been significantly improved.

From Fig. C.4 we observed an exponential decline in the importance of the features, and good accuracy was obtained for the top 150 features in each age group. Therefore, we selected the top 150 features of each group to analyze the consistency rate of machine learning knowledge against expert knowledge. As shown in Table 2, the absolute consistency rate achieved for age groups 18–35, 36–55, 56–65, and > 65 was 78.4%, 77.2%, 81.3% and 79.5%, respectively. Because the consistency rate exceeds the 50% threshold, we can think that the knowledge mining approach is credible and can be further used to explore and mine new potential risk knowledge.

3.3. Heterogeneity of AKI risk predictors between age groups

To study the heterogeneity of AKI risk factors between age groups, we compared the weighted SHAP values generated by the knowledge mining method between the four age groups. Fig. 3 illustrated the comparison of 54 risk predictors with an average weighted SHAP value greater than 0.1 in at least one age group, and we observed a clear difference in the effect of each predictor on AKI risk across age groups. Fig. 4a compares the intersection number of top *k* features of AKI between age groups, suggesting that clinical predictors are heterogeneous across age strata. And Fig. 4b shows the relative risk differences between 166 drugs in four AKI age groups. In the case of medication exposures, Fig. C.5 shows the effects of *vancomycin*, *prednisone*, and *albumin* on AKI risk. Besides, age differences in some drugs may be related to the disease being treated, for example, knowledge networks constructed using correlations showed that *dornase alfa* was highly associated with *cystic fibrosis* (see Fig. 5a-b).

Table 3 shows the personalized risk factors of patient A and patient B randomly selected from the young and older groups respectively,

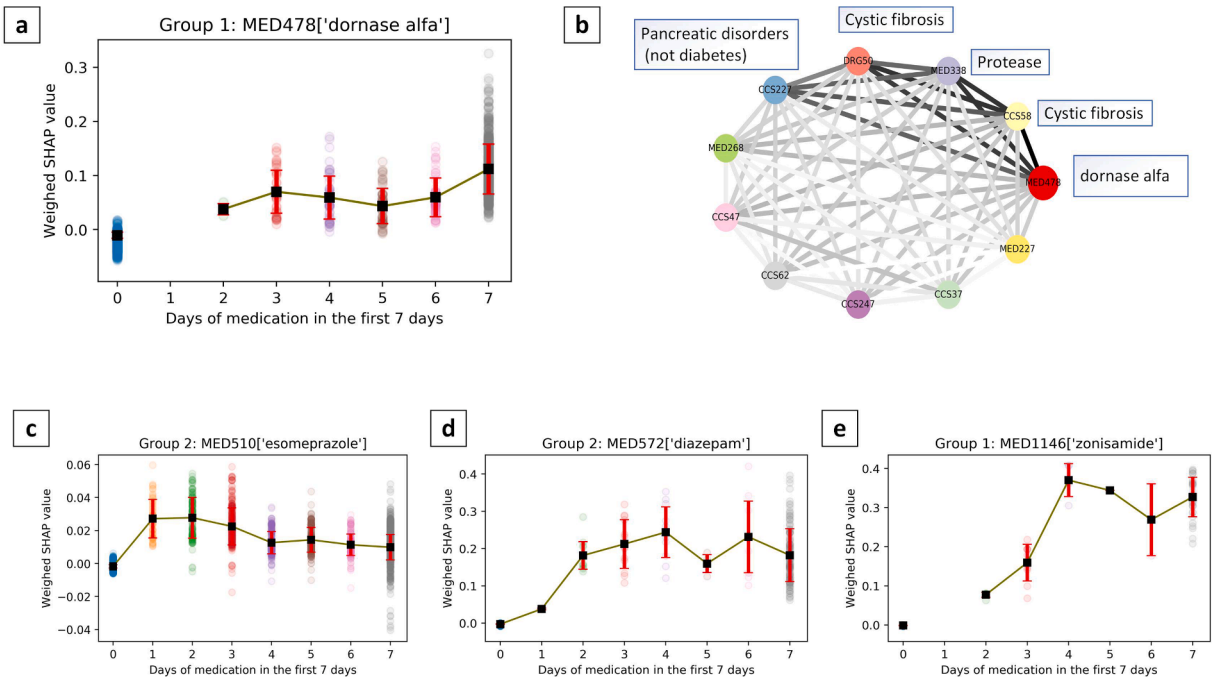


Fig. 5. Exploring new potential nephrotoxic drugs. (a-b) Variable impact and correlation network of dornase alfa (MED478) for AKI in 18–35 age group; (c-e) esomeprazole (MED510), diazepam (MED572) and zonnisamide (MED1146). Each dot is a person. The x-axis represents days of exposure to medication in the first 7 days before reference point, and the y-axis is the weighted SHAP value attributed to the variable. Higher weighted SHAP values represent higher risk of AKI due to the variable.

Table 3
Personalized interpretation of AKI risk predictors.

Rank	Name of positive effect	Value	wSHAP	Name of negative effect	Value	wSHAP
AKI patient A from group 1 (age 18–35)						
1	CCS58['Cystic fibrosis']	1	0.743	MED1086['tazobactam']	0	−0.113
2	MED516['glucose']	1	0.705	Lab5_3['BUN, more than the standard value']	0	−0.066
3	days ['Length of stay']	2	0.584	MED321['vancomycin']	0	−0.051
4	MED1102['linezolid']	7	0.256	Lab4_2['Calcium,the standard value']	1	−0.046
5	MED338['protease']	7	0.216	Lab0_1['Albumin, less than standard value']	0	−0.039
6	MED478['dornase alfa']	7	0.175	BMI_2['normal']	1	−0.035
7	MED187['trisulfapyrimidines']	7	0.144	MED1133['benzimidazole']	0	−0.032
8	MED677['polyethylene glycol 3350']	7	0.140	CCS230['Acute and unspecified renal failure']	0	−0.030
9	MED659['ursodeoxycholate']	7	0.128	Lab9_3['Glucose, more than the standard value']	0	−0.030
10	Lab5_1['BUN, less than standard value']	0	0.112	CCS168['Medical examination/evaluation']	0	−0.027
AKI patient B from group 4 (age > 65)						
1	MED1086['tazobactam']	6	0.404	Lab5_3['BUN, more than the standard value']	0	−0.428
2	Lab13_3['WBC, more than the standard value']	1	0.391	MED516['glucose']	0	−0.202
3	MED281['mercaptopurine']	6	0.200	Lab0_1['Albumin, less than standard value']	0	−0.105
4	MED1111['aminobutyrate']	5	0.101	Days ['Length of stay']	5	−0.102
5	MED691['influenza a virus']	5	0.086	MED134['benzoic acid']	0	−0.070
6	MED103['n-(hydroxyethyl) ethylenediaminetriacetic acid']	6	0.086	MED937['ropivacaine']	6	−0.061
7	Lab13_2['WBC, the standard value']	0	0.068	Lab4_1['Calcium, less than standard value']	0	−0.055
8	BMI_4['obese']	1	0.056	MED321['vancomycin']	0	−0.046
9	MED582['levofloxacin']	0	0.053	Lab9_3['Glucose, more than the standard value']	0	−0.036
10	SBP_4['stage 2 hypertension']	1	0.044	MED975['triflusal']	6	−0.028

Note: wSHAP is the weighted SHAP value obtained by the knowledge discovery model.

showing that these two patients have different risk factors. For example, *cystic fibrosis* in the young patient A has the highest risk weight (*wSHAP* = 0.743), while the use of *tazobactam* (*wSHAP* = 0.404) and the abnormality of white blood cell (*wSHAP* = 0.391) in elderly patient B suggest a higher risk of AKI. It is worth noting that the factors selected by machine learning methods have strong predictive power, but these factors are not necessarily direct causal triggers. For example, *insulin* was screened as a risk factor, but it may only be a marker of diabetes, that is, patients with diabetes or diabetic nephropathy have a higher risk of developing AKI.

4. Discussion

Data mining and knowledge discovery techniques have the potential to play important roles in clinical decision support for healthcare professionals. In this study, we developed a novel knowledge mining approach to uncover heterogeneous risk predictors of AKI across age groups, and validated its effectiveness with respect to accuracy, stability and credibility. We selected four age groups for a comprehensive analysis of risk factor heterogeneity between age groups instead of using a single model containing age variable because it would be challenging to delineate cross-effect of thousands of variables and age in one model.

The heterogeneity of AKI risk factors across age groups mainly includes two aspects: (1) risk factors only exist in a certain age group; these differences may be due to the relative age concentration of the affected population. For instance, *cystic fibrosis*, a hereditary exocrine gland disease and its patients rarely survive to adulthood, is mostly concentrated in younger age groups; in contrast, *hypertension* was identified to only present risk in older population; (2) the risk factors existed in multiple age groups, but the risk weights were significantly different. For instance, evidence from existing literature [26] suggest that African Americans are at higher risk for AKI; however, Fig. C.6 shows that this phenomenon is not prominent in the young population. Moreover, high BMI (body mass index) was found to be associated with higher AKI risk in older patients, but low BMI was found to be associated with higher AKI risk in younger patients (see Fig. C.7).

In-depth analyses of some inconsistent results revealed potential explanations. For instance, healthcare providers may be more cautious or there may already exist clinical decision support tools to guide clinicians on using nephrotoxic medications, for example NSAIDs (non-steroidal anti-inflammatory drugs) such as *ibuprofen* and *suprofen*, in patients with poor renal function, which could mislead the AI model to produce low-risk result inconsistent with known knowledge. Sometimes one approach to minimize risk of nephrotoxic drugs is to change its form or route of usage. For instance, *bacitracin* with severe nephrotoxicity is not used systemically, only for topical use and *phenol* with moderate toxicity is often used as a powerful surgical disinfectant. Since our model did not consider the route of medication administration, the different ways of usage may change the degree of risk obtained from our data, resulting in inconsistent conclusions. When inconsistent but explicable factors were included, as shown in Table C.6, the explicable consistency rates for age groups 18–35, 36–55, 56–65, and > 65 were 87.8%, 87.3%, 90.7% and 92.3%, respectively.

The data driven knowledge discovery approach combines machine derived knowledge with human intelligence to enable potential discovery of new knowledge. Recent studies have found that long-term use of proton pump inhibitor (PPI) drugs may have adverse effects on the kidney [27,28]; study identified *esomeprazole* (as one of the PPI) showing a higher relative risk of AKI (see Fig. 5c). Similarly, the incidence of AKI was higher with *diazepam* than without it in the 36–55 age group (see Fig. 5d). In addition, *zonisamide*, which was screened only in the 18–35 group, has been shown to be associated with a higher risk of AKI (see Fig. 5e). Meanwhile, two cases of *zonisamide* as an antiepileptic drug induced AKI have been reported a 29-year-old Japanese man [29] and a 33-year-old Caucasian man [30].

This study has some limitations. First, although we utilized a large cohort observed for up to a decade, they only reflect the population of one academic medical center. Second, the granularity of medication data extraction may affect the knowledge learned within machine learning models [31]. Considering the drug metabolism cycle, in this study we only considered medications taken within a week. Third, this study explored the entirety of the above mentioned EMR data types except for laboratory tests where we only selected certain lab tests based on previous literature for AKI prediction [20]. Fourth, we only invited 3 experts for knowledge evaluation. In the future, more doctors are needed to amass and codify more knowledge. Fifth, our data only contains whether there is a certain disease, but does not contain the corresponding severity score of the disease. Sixth, this study used tabular-style data lacking strong multiscale temporal structures, and the knowledge mining of temporal information and patterns will be considered in the future. Finally, this knowledge mining method would generate a list of potential new risk knowledge, but whether these new risk factors increase the risk of AKI requires rigorous demonstration from clinical experiments, and further work is necessary to investigate the nature of the association.

5. Conclusions

In this study, we leveraged a large EMR dataset to develop a novel knowledge mining approach and validated its effectiveness with respect to accuracy, stability and credibility. Using this approach, we elucidated the heterogeneity of AKI risk factors across age groups, which suggested that future clinical care need to consider risk heterogeneity. In addition, the data driven knowledge mining method combines machine derived knowledge with human intelligence to enable potential discovery of new risk knowledge.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are deeply grateful to Drs. Jia Zhang (general internist), Xi Zhang (nephrologist) and Enhao Yang (nephrologist) from the first affiliated hospital of Chongqing Medical University for their effort in assessing the risk relationship between the predictors and AKI.

Fundings

This research was partially supported by the Youth Science Fund of the National Natural Science Foundation of China [grant number 61802149], the Major Research Plan of the National Natural Science Foundation of China [Key Program, grant number 91746204], Guangzhou Basic and Applied Basic Research Project [grant number 202102020863], the Science and Technology Development in Guangdong Province [Major Projects of Advanced and Key Techniques Innovation, grant number 2017B030308008], Guangdong Engineering Technology Research Center for Big Data Precision Healthcare [grant number 603141789047], the Fundamental Research Funds for the Central Universities [grant number 21618315], the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health [grant number R01DK116986], and the National Science Foundation [grant number 2014554]. The content is solely the responsibility of the authors and does not represent the official views of the funders. The clinical dataset used for analysis described in this study was obtained from KUMC's HERON clinical data repository which is supported by institutional funding and by the KUMC CTSA grant UL1TR002366 from NIH.

Authors' contributions

LW designed the overall study, carried out the experiments, and wrote the manuscript. ML and YH critically appraised and revised the manuscript. ML performed the EMR data extraction. XZ contributed in data processing. JZ used medical literature and expert experience to make judgments on medical knowledge mined by the AI models. All authors reviewed the manuscript critically for scientific content, and all authors gave final approval of the manuscript for publication.

Data availability

The clinical data used in this study is not publicly available and restrictions apply to its use. The de-identified datasets may be available from the Healthcare Enterprise Resource for Ontological Narration (HERON), subject to local and national ethical approvals.

Code availability

We used the open-source Python libraries (e.g., xgboost and shap) to

conduct our experiments. Implementation details and code are available at https://github.com/BDAII/AKI-age to allow independent replication.

Summary points

What was already known on the topic.

- Acute kidney injury (AKI) risk increases with age and the underlying clinical predictors may be heterogeneous across age strata;
- The treatment of AKI is usually poor; therefore, efforts have been made to detect and prevent AKI early;
- There has been growing interest in harnessing electronic medical records to enable potential discovery of new risk knowledge.

What this study added to our knowledge.

- We established a knowledge mining approach and verified its effectiveness from three perspectives: accuracy, stability and consistency;
- We used this approach to clarify the heterogeneity of AKI risk factors between age groups;
- The study findings implicate that future clinical decision support systems need to consider such heterogeneity to enhance personalized care for improving patient outcomes.
- The data driven knowledge mining method combines machine derived knowledge with human intelligence to enable potential discovery of new risk knowledge.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijmedinf.2021.104661.

References

[1] E.A.J. Hoste, S.M. Bagshaw, R. Bellomo, C.M. Cely, R. Colman, D.N. Cruz, K. Edipidis, L.G. Forni, C.D. Gomersall, D. Govil, P.M. Honoré, O. Joannes-Boyau, M. Joannidis, A.-M. Korhonen, A. Lavrentieva, R.L. Mehta, P. Palevsky, E. Roessler, C. Ronco, S. Uchino, J.A. Vazquez, E. Vidal Andrade, S. Webb, J.A. Kellum, others, Epidemiology of acute kidney injury in critically ill patients: the multinational AKI-EPI study, Intensive Care Med. 41 (8) (2015) 1411–1423.

[2] H.E. Wang, P. Muntner, G.M. Chertow, D.G. Warnock, Acute kidney injury and mortality in hospitalized patients, Am. J. Nephrol. 35 (4) (2012) 349–355.

[3] S. Uchino, R. Bellomo, D. Goldsmith, S. Bates, C. Ronco, An assessment of the RIFLE criteria for acute renal failure in hospitalized patients, Crit. Care Med. 34 (7) (2006) 1913–1917.

[4] E.A.J. Hoste, J.A. Kellum, N.M. Selby, A. Zarbock, P.M. Palevsky, S.M. Bagshaw, S. L. Goldstein, J. Cerdá, L.S. Chawla, Global epidemiology and outcomes of acute kidney injury, Nat. Rev. Nephrol. 14 (10) (2018) 607–625.

[5] G.M. Chertow, E. Burdick, M. Honour, J.V. Bonventre, D.W. Bates, Acute kidney injury, mortality, length of stay, and costs in hospitalized patients, J. Am. Soc. Nephrol. Jasn. 16 (11) (2005) 3365–3370.

[6] J.A. Kellum, F.E. Sileanu, R. Murugan, N. Lucko, A.D. Shaw, G. Clermont, Classifying AKI by urine output versus serum creatinine level, J. Am. Soc. Nephrol. 26 (9) (2015) 2231–2238.

[7] H. Schiffl, The pathogenesis of acute kidney injury may differ between elderly and younger patients, Am. J. Kidney Dis. 52 (6) (2008) 1198–1199.

[8] J.L. Koyner, K.A. Carey, D.P. Edelson, M.M. Churpek, The development of a machine learning inpatient acute kidney injury prediction model, Crit. Care Med. 46 (7) (2018) 1070–1077.

[9] M. Flechet, F. Güiza, M. Schetz, P. Wouters, I. Vanhorebeek, I. Derese, J. Gunst, I. Spriet, M. Casaer, G. Van den Berghe, G. Meyfroidt, AKIpredictor, an online prognostic calculator for acute kidney injury in adult critically ill patients: development, validation and comparison to serum neutrophil gelatinase-associated lipocalin, Intensive Care Med. 43 (6) (2017) 764–773, https://doi.org/10.1007/s00134-017-4678-3.

[10] R.J. Kate, R.M. Perez, D. Mazumdar, K.S. Pasupathy, V. Nilakantan, Prediction and detection models for acute kidney injury in hospitalized older adults, BMC Med. Inf. Decis. Making 16 (2016) 39.

[11] V. Gulshan, L. Peng, M. Coram, M.C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, R. Kim, R. Raman, P.C. Nelson, J.L. Mega, D.R. Webster, Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs, JAMA 316 (22) (2016) 2402, https://doi.org/10.1001/jama.2016.17216.

[12] B. Shickel, P.J. Tighe, A. Bihorac, P. Rashidi, Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis, IEEE J. Biomed. Heal Informatics 22 (2017) 1589–1604.

[13] R. Miotto, F. Wang, S. Wang, X. Jiang, J.T. Dudley, Deep learning for healthcare: review, opportunities and challenges, Brief Bioinform. 201 (2017) 1693–1701, https://doi.org/10.1093/bib/bbx044.

[14] Y. Cheng, F. Wang, P. Zhang, J. Hu, Risk Prediction with Electronic Health Records : A Deep Learning Approach, in: Siam Int Conf Data Min, 2016: pp. 432–440.

[15] A.V. Konstantinov, L.V. Utkin, Interpretable machine learning with an ensemble of gradient boosting machines, Knowledge-Based Syst. 222 (2021) 106993, https://doi.org/10.1016/j.knosys.2021.106993.

[16] S.M. Lundberg, G. Erion, H. Chen, A. DeGrave, J.M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, S.-I. Lee, From local explanations to global understanding with explainable AI for trees, Nat. Mach. Intell. 2 (1) (2020) 56–67.

[17] M.T. Ribeiro, S. Singh, C. Guestrin, Why should I trust you?: Explaining the predictions of any classifier, in: Proc Ofthe 22nd ACM SIGKDD Int Conf Knowl Discov Data Min, 2016: pp. 1135–1144.

[18] S.N. Murphy, G. Weber, M. Mendis, V. Gainer, H.C. Chueh, S. Churchill, I. Kohane, Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2), J. Am. Med. Inform. Assoc. 17 (2) (2010) 124–130.

[19] D. Fliser, M. Laville, A. Covic, D. Fouque, R. Vanholder, L. Juillard, W. Van Biesen, A European Renal Best Practice (ERBP) position statement on the Kidney Disease Improving Global Outcomes (KDIGO) Clinical Practice Guidelines on Acute Kidney Injury: Part 1: Definitions, conservative management and contrast-induced nephropathy, Nephrol. Dial. Transplant. 27 (2012) 4263–4272.

[20] M.E. Matheny, R.A. Miller, T.A. Ikizler, L.R. Waitman, J.C. Denny, J.S. Schildcrout, R.S. Dittus, J.F. Peterson, Development of inpatient risk stratification models of acute kidney injury for use in electronic health records, Med. Decis Mak. 30 (6) (2010) 639–650.

[21] T. Chen, C. Guestrin, XGBoost: A scalable tree boosting system, in: Proc 22nd Acm Sigkdd Int Conf Knowl Discov Data Min, 2016: pp. 785–794.

[22] J.G. Moreno-Torres, J.A. Saez, F. Herrera, Study on the Impact of Partition-Induced Dataset Shift on k-Fold Cross-Validation, IEEE Trans. Neural Networks Learn. Syst. 23 (2012) 1304–1312, https://doi.org/10.1109/TNNLS.2012.2199516.

[23] A.P. Bradley, The use of the area under the ROC curve in the evaluation of machine learning algorithms, Pattern Recogn. 30 (7) (1997) 1145–1159.

[24] A. Kalousis, J. Prados, M. Hilario, Stability of feature selection algorithms: a study on high-dimensional spaces, Knowl, Inf Syst. 12 (1) (2007) 95–116.

[25] L. van der Maaten, G. Hinton, Visualizing data using t-SNE, J Mach Learn Res. 9 (2008) 2579–2605.

[26] S.L. Kane-Gill, F.E. Sileanu, R. Murugan, G.S. Trietley, S.M. Handler, J.A. Kellum, Risk Factors for Acute Kidney Injury in Older Adults With Critical Illness: A Retrospective Cohort Study, Am. J. Kidney Dis. 65 (6) (2015) 860–869.

[27] D.G. Klepser, D.S. Collier, G.L. Cochran, Proton pump inhibitors and acute kidney injury : a nested case – control study, Bmc Nephrol. 14 (2013) 150.

[28] Y. Xie, B. Bowe, T. Li, H. Xian, Y. Yan, Z. Al-Aly, Long-term kidney outcomes among users of proton pump inhibitors without intervening acute kidney injury, Kidney Int. 91 (6) (2017) 1482–1494, https://doi.org/10.1016/j.kint.2016.12.021.

[29] Y. Fujita, M. Hasegawa, K. Nabeshima, M. Tomita, K. Murakami, S. Nakai, T. Yamakita, K. Matsunaga, Acute kidney injury caused by zonisamide-induced hypersensitivity syndrome, Int. Med. 49 (5) (2010) 409–413.

[30] D. Dixit, D. Stewart, M.M. Bridgeman, A. Parikh, Zonisamide-induced acute kidney injury, Epilepsy Behav Case Reports. 3 (2015) 23–25.

[31] X. Song, L.R. Waitman, Y. Hu, A.S.L. Yu, D. Robbins, M. Liu, An exploration of ontology-based EMR data abstraction for diabetic kidney disease prediction, AMIA Summits Transl. Sci. Proc. 2019 (2019) 704.