Discovering Risk Factors Associated with Inpatient Acute Kidney Injury across Age Groups

Lijuan Wu, PhD[1], Yong Hu, PhD[1], Jia Zhang, MD[2], Xiangzhou Zhang, PhD[1], Weiqi Chen, PhD[1], Kang Liu, MS[1], Alan S.L. Yu, MD[3], John A. Kellum, MD[4], Lemuel R. Waitman, PhD[5], Mei Liu, PhD[5]

[1]Big Data Decision Institute, Jinan University, Guangzhou, China; [2]The First Affiliated Hospital of Chongqing Medical University, Chongqing, China; [3]Division of Nephrology and Hypertension and the Kidney Institute, The University of Kansas Medical Center, Kansas, USA; [4]Department of Critical Care Medicine, University of Pittsburgh School of Medicine, Pittsburgh, USA; [5]Division of Medical Informatics, The University of Kansas Medical Center, Kansas, USA

BACKGROUND & OBJECTIVE

- **Acute kidney injury (AKI)** is a challenging medical problem, affecting 10% to 15% of all hospitalized patients and >50% of patients in ICUs. Since full recovery of kidney function after AKI episodes is rare, it leads to significant short-term and long-term morbidity and mortality.
- **Risk of AKI** is determined by complex interactions of patient factors including age, pre-existing medical conditions, and exposure to nephrotoxic medications. It is well known that AKI risk increases with age, but it is unclear how other risk factors vary between younger and older patients.
- **Key Question:** What are the risk factor differences underlying AKI in older compared to younger adults?

RESEARCH DESIGN & METHOD

- **Objective:** To determine whether AKI risk factors in general inpatients differ across age groups using real-world evidence from electronic health records (EHRs).
- **Design:** A retrospective analysis of adult patients admitted to a tertiary-care academic hospital between November 2007 to December 2016 for two or more days.
- **Participants:** We excluded encounters missing serum creatinine measurements for AKI determination and patients with moderate to severe kidney dysfunction at admission. Total 76,957 encounters remained and were stratified into four age groups: 18-35, 36-55, 56-65, and >65.
- **Exposures:** We collected 1,888 clinical variables from the EHR including demographics, vital signs, medications, laboratory values, past and admission diagnoses.
- **Main Outcome:** We defined AKI according to the KDIGO (Kidney Disease: Improving Global Outcomes) serum creatinine criteria.
- **Method:** We developed an EHR-data driven knowledge discovery model to identify risk factors and their impact on AKI across age strata.
  - **XGBoost (eXtreme Gradient Boosting)** - identify predictors of AKI across age strata
  - **SHAP (Shapley Additive exPlanation)** – quantify impact of the predictors on AKI

RESULTS

Study Cohort Characteristics

	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
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)

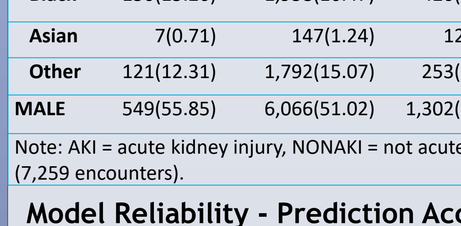
Note: AKI = acute kidney injury, NONAKI = not acute kidney injury. Values are given as number (percentage). Overall AKI rate is 9.43% (7,259 encounters).

Model Reliability - Prediction Accuracy (AUROC, 95% CI)

Models	Age 18-35	Age 36-55	Age 56-65	Age >65
XGBoost	0.85 [0.80-0.88]	0.86 [0.83-0.89]	0.87 [0.86-0.90]	0.87 [0.86-0.90]
GBM	0.85 [0.81-0.87]	0.86 [0.84-0.89]	0.87 [0.85-0.89]	0.86 [0.83-0.89]
Random Forest	0.84 [0.80-0.87]	0.85 [0.82-0.88]	0.87 [0.85-0.89]	0.86 [0.85-0.89]
LinearSVC	0.73 [0.71-0.77]	0.77 [0.76-0.79]	0.78 [0.76-0.80]	0.78 [0.76-0.80]
Logistic Regression	0.76 [0.73-0.80]	0.80 [0.78-0.83]	0.80 [0.77-0.82]	0.80 [0.77-0.82]
Naïve Bayes	0.75 [0.70-0.80]	0.76 [0.71-0.79]	0.76 [0.74-0.79]	0.76 [0.74-0.79]
Neural Network	0.74 [0.70-0.79]	0.77 [0.74-0.80]	0.79 [0.76-0.80]	0.79 [0.76-0.80]
Decision Trees	0.62 [0.59-0.67]	0.66 [0.64-0.68]	0.66 [0.64-0.68]	0.66 [0.64-0.68]

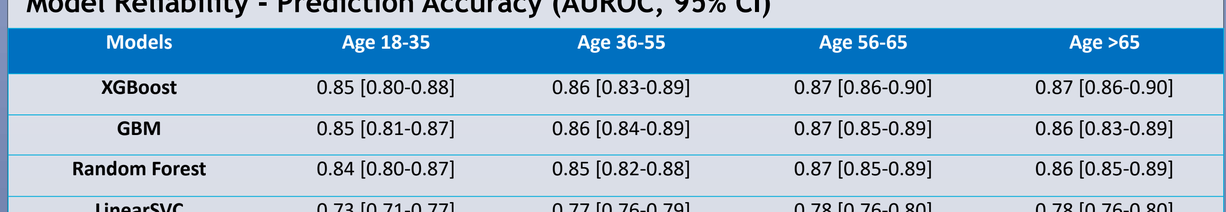
Note: XGBoost: eXtreme Gradient Boosting; GBM: Gradient Boosting Machine. CI: confidence interval. There is no significant difference in the classification results from XGBoost, GBM and Random Forest models.

Reliability of the Identified Predictors against Expert Knowledge



Age Group	Y	N0	N1	N	U	Y/(Y+N)	(Y+N1)/(Y+N)
G1: 18-35	58	9	7	16	76	78.4%	87.8%
G2: 36-55	61	10	8	18	71	77.2%	87.3%
G3: 56-65	61	7	7	14	75	81.3%	90.7%
G4: >65	62	6	10	16	72	79.5%	92.3%

Notes: Y stands for yes or consistency; N stands for no or inconsistency; N = N0 + N1; N1 stands for inconsistency but can be explained; N0 stands for inconsistency and is temporarily unexplained; U stands for unknown or cannot make judgment based on current knowledge.



HETEROGENEITY OF AKI PREDICTORS

Body Mass index (BMI)

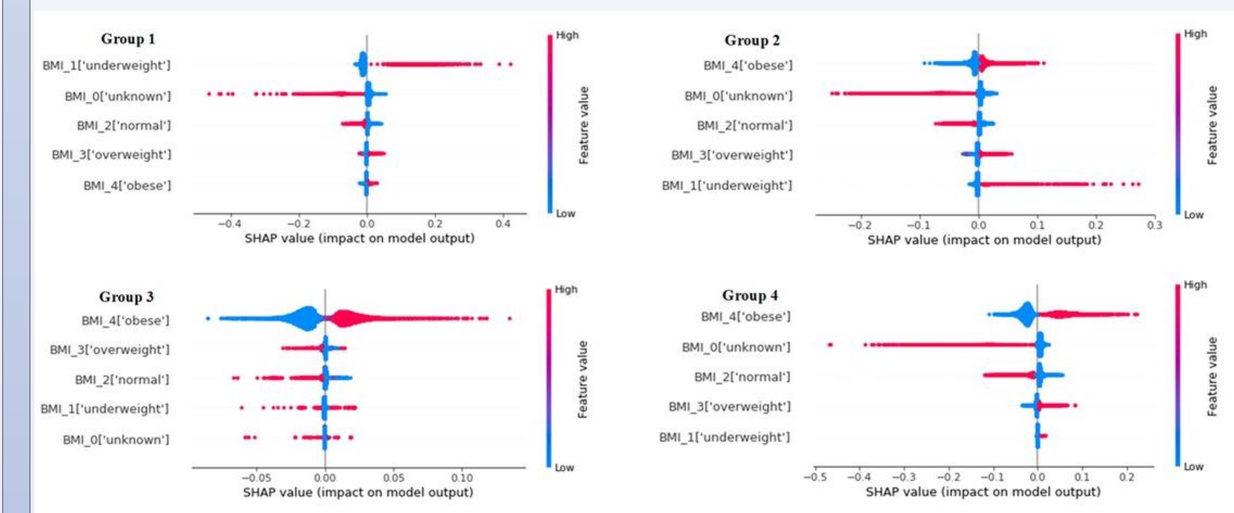
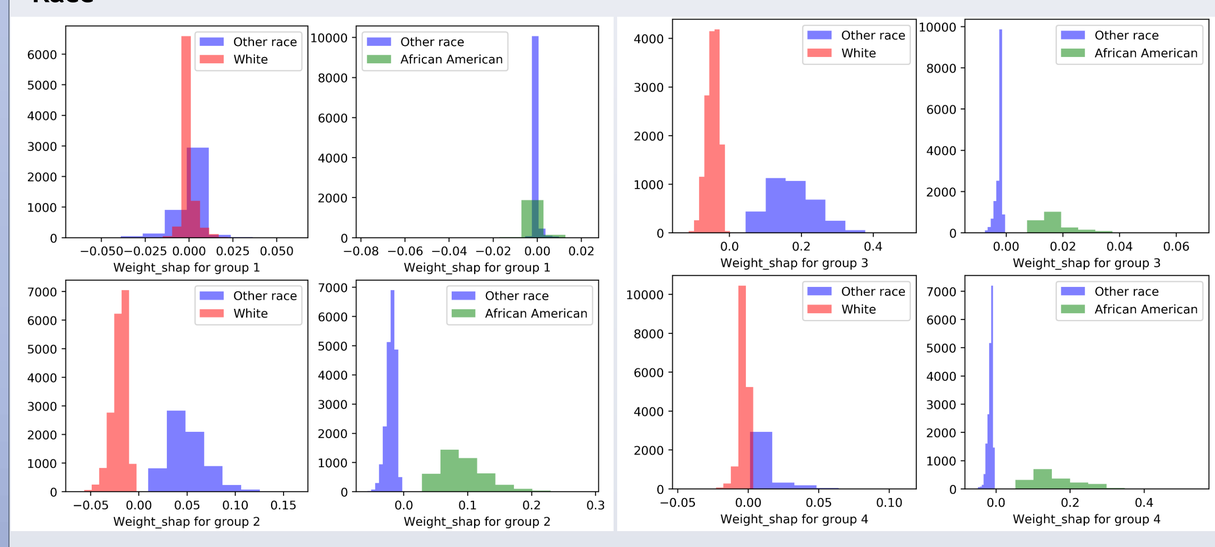


Figure. Weighted average SHAP value plot for BMI. The higher the SHAP value of a clinical factor, the higher risk of AKI due to that factor. Each dot in the plot is a person with that factor value. Dots are colored by the factor value for that person and piled up vertically to show density.

Race



Case Analysis

Rank	Positive effect	Value	wSHAP	Negative effect	Value	wSHAP
			AKI patient A from age group 1			
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 age group 4			
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.

SUMMARY

Key Findings: We identified a set of important risk factors of AKI for four age groups and observed heterogeneity in the impact of those factors on the development of AKI across groups, demonstrating age-specific risk factors differences. The results suggest that AKI prevention efforts need to be tailored to patients within specific age groups.

ACKNOWLEDGEMENT

The research was partially supported by the US NIH/NIDDK under award number R01DK116986, National Natural Science Foundation and Guangdong Province of China (Grant No. 91746204, 61802149, 2017B030308008, 603141789047, 21618315). We thank Drs. Jia Zhang (general practitioner), Xi Zhang (nephrologist), and Enhao Yang (nephrologist) for their clinical annotation effort.