Joint Modeling of Clinical and Biomarker Data in Acute Kidney Injury Defines Unique Subphenotypes with Differing Outcomes

George Vasquez-Rios[1], Wonsuk Oh[2,3], Samuel Lee[4], Pavan Bhatraju[5,6], Sherry G. Mansour[7], Dennis G. Moledina[7], Faris F. Gulamali[2,3], Edward D. Siew[8], Amit X. Garg[9], Pinaki Sarder[10], Vernon M. Chinchilli[11], James S. Kaufman[12], Chi-yuan Hsu[13], Kathleen D. Liu[13], Paul L. Kimmel[14], Alan S. Go[15], Mark M. Wurfel[6], Jonathan Himmelfarb[5], Chirag R. Parikh[16], Steven G. Coca[1], and Girish N. Nadkarni[1,2,3]

Due to the number of contributing authors, the affiliations are listed at the end of this article.

Correspondence: Dr. Steven G. Coca, Barbara T. Murphy Division of Nephrology, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, Box 1243 New York, NY 10029, or Dr. Girish N. Nadkarni, Barbara T. Murphy Division of Nephrology, Division of Data-Driven and Digital Medicine (D3M), Mount Sinai Clinical Intelligence Center (MSCIC), Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, Box 1243, New York, NY 10029. Email: steven.coca@mssm.edu or girish.nadkarni@mountsinai.org

Abstract

Background AKI is a heterogeneous syndrome. Current subphenotyping approaches have only used limited laboratory data to understand a much more complex condition.

Methods We focused on patients with AKI from the Assessment, Serial Evaluation, and Subsequent Sequelae in AKI (ASSESS-AKI). We used hierarchical clustering with Ward linkage on biomarkers of inflammation, injury, and repair/health. We then evaluated clinical differences between subphenotypes and examined their associations with cardiorenal events and death using Cox proportional hazard models.

Results We included 748 patients with AKI: 543 (73%) of them had AKI stage 1, 112 (15%) had AKI stage 2, and 93 (12%) had AKI stage 3. The mean age (±SD) was 64 (13) years; 508 (68%) were men; and the median follow-up was 4.7 (Q1: 2.9, Q3: 5.7) years. Patients with AKI subphenotype 1 (N=181) had the highest kidney injury molecule (KIM-1) and troponin T levels. Subphenotype 2 (N=250) had the highest levels of uromodulin. AKI subphenotype 3 (N=159) comprised patients with markedly high pro–brain natriuretic peptide and plasma tumor necrosis factor receptor-1 and -2 and low concentrations of KIM-1 and neutrophil gelatinase–associated lipocalin. Finally, patients with subphenotype 4 (N=158) predominantly had sepsis-AKI and the highest levels of vascular/kidney inflammation (YKL-40, MCP-1) and injury (neutrophil gelatinase–associated lipocalin, KIM-1). AKI subphenotypes 3 and 4 were independently associated with a higher risk of death compared with subphenotype 2 and had adjusted hazard ratios of 2.9 (95% confidence interval, 1.8 to 4.6) and 1.6 (95% confidence interval, 1.01 to 2.6, P = 0.04), respectively. Subphenotype 3 was also independently associated with a three-fold risk of CKD and cardiovascular events.

Conclusions We discovered four AKI subphenotypes with differing clinical features and biomarker profiles that are associated with longitudinal clinical outcomes.

CJASN 18: 716–726, 2023. doi: https://doi.org/10.2215/CJN.0000000000000156

Introduction

AKI is a heterogeneous syndrome that occurs in diverse clinical scenarios.[1] Despite this heterogeneity, the classification of AKI is influenced by heuristics, comprising prerenal, intrinsic, and postrenal causes.[1,2] Importantly, multiple insults may coexist leading to biological disturbances that are common to these classifications, but that are associated with distinct outcomes. However, identification of pathological pathways, prediction of disease duration, and subsequent clinical complications are difficult to obtain using serum creatinine fluctuations, urine studies, and urine output changes.[3–6] Furthermore, kidney function after AKI can follow diverse trajectories that do not correspond with marked laboratory abnormalities, nor the ascertained etiology.[1,4,7,8] This represents a major limitation to providing individualized care during the acute and postacute stages.

Biomarkers reflective of key biological pathways have emerged showing the potential to forecast adverse outcomes in this population, including disease duration and mortality.[3–5,9–13] Furthermore, among patients with CKD, some biomarkers strongly predict kidney failure in patients with and without diabetes mellitus.[14–21] While efforts have been made to integrate individual biomarkers (or panels) into prediction models, standard statistical modeling is limited in its ability to evaluate heterogeneous data and identify clinical significance from subtypes when multiple physiologic stressors occur.[22] Therefore, the importance of incorporating big data to enrich perspectives on AKI has been acknowledged by the 15th Acute Dialysis Quality Initiative.[23]

A multidimensional approach is needed to identify meaningful AKI subphenotypes. Unsupervised learning can recognize patterns within diverse, multimodal

data. Previous reports have used AKI subphenotypes based on limited biomarker panels to assess short-term outcomes or responsiveness to vasopressors.[24,25] Nonetheless, whether machine learning algorithms incorporating a comprehensive panel of biomarkers could uncover agnostic subphenotypes that associate with longitudinal clinical events in a broader AKI population is unknown. In this study, we sought to develop unbiased clustering approaches using data obtained during the index hospitalization to unearth novel subphenotypes and assess whether these clusters were associated with long-term clinical events.

Methods

Cohort Description and Covariates

The Assessment, Serial Evaluation, and Subsequent Sequelae in AKI (ASSESS-AKI) is a multicenter, prospective, matched parallel cohort comprising 1538 hospitalized adult participants with and without AKI who survived to complete an in-person visit at 3 months after discharge. Details of the study have been published elsewhere.[26] Patients had preadmission serum creatinine measurements. The etiology of AKI was not adjudicated during the hospitalization. Exclusion criteria were broad, such as acute glomerulonephritis, hepatorenal syndrome, malignancy, urinary obstruction, severe heart failure (New York Heart Association class IV before index admission), KRT (dialysis, transplant) before hospitalization, or predicted survival ≤12 months. We focused on those patients who had AKI and who survived at least 3 months after discharge (N=769).

AKI was defined as a relative increase of ≥50% or absolute increase of ≥0.3 mg/dl in peak inpatient serum creatinine concentration above baseline. AKI stage was classified following the AKI Network guidelines. CKD status and eGFR calculation were obtained using the Chronic Kidney Disease Epidemiology Collaboration equation including the race coefficient. Participants had study visits at 3 and 12 months after discharge and annually thereafter. AKI recovery during hospitalization was defined as a decrease in serum creatinine concentration of ≥0.3 mg/dl or ≥25% relative to peak creatinine.[27] Serum creatinine at follow-up ASSESS-AKI study visits was measured using an isotope dilution mass spectrometry-traceable enzymatic assay (Roche Diagnostics, Indianapolis, IN) and a random spot urine protein-to-creatinine ratio using a turbidimetric method (Roche) in a central laboratory. Biomarker quartiles were defined within the AKI cohort using different diagnostic assays.

Variable Selection and Biomarker Measurements

A total of 53 continuous and discrete clinical variables were incorporated into our machine learning algorithms on the basis of their previously described relevance (Table 1).[3,4,13,28–30] Demographic variables, comorbidities, and kidney-specific characteristics were all included. Plasma and urine samples were collected within 96 hours before or after the diagnosis of AKI. Blood samples were collected in EDTA tubes and centrifuged to separate plasma. Samples underwent a single controlled thaw, were centrifuged at ×5000 g for 10 minutes at 4°C, separated into 1-ml aliquots, and immediately stored at −80°C until biomarkers were measured. Urine biomarkers were obtained in a protocolized fashion and measured by using a multiplex assay or by using Meso Scale Diagnostics when appropriate. Details on biomarker collection and processing have been described previously.[13]

Table 1. Clinical and biomarker variables incorporated into machine learning algorithms

Domain	Variable
Demographics	Age, sex, race, smoking status, and ICU admission
Comorbidities	Hypertension, diabetes mellitus, cardiovascular disease, CHF, chronic obstructive pulmonary disease, connective lung disease, systemic lupus erythematosus, and sepsis
Kidney-related covariates	Baseline Cr, baseline estimated glomerular filtration rate, peak Cr, peak GFR, urine albumin, urine creatinine, peak urine albumin-creatinine ratio, peak urine protein-creatinine ratio, AKI stage
Systemic inflammatory biomarkers	IL-1b, IL-2, IL-6, IL-10, IL-12, IL-13, TNF-α, and IFN-γ
Cardiac-related biomarkers	Troponin T, NT-proBNP, GAL3, and ST2
Mineral bone disease–related biomarkers	PTH, $PO4^{-3}$, and FGF23
Kidney-related biomarkers	TNFR1, TNFR2, urine IL-18 (IL-18), KIM-1, MCP-1, YKL-40, NGAL, UMOD, CysC, and uOSM

ICU, intensive care unit; CHF, congestive heart failure; Cr, creatinine; NT-proBNP, NT-pro–brain natriuretic peptide; GAL3, galectin 3; ST2, suppression of tumorigenicity 2 receptor; PTH, parathyroid hormone; $PO4^{-3}$, phosphate; FGF23, fibroblast growth factor 23; TNFR1, TNFR2, plasma tumor necrosis factor receptor-1 and -2; KIM-1, urine kidney injury molecule; MCP-1, urine MCp-1; YKL-40, urine chitinase-3–like protein 1; NGAL, urine neutrophil gelatinase–associated lipocalin; UMOD, urine uromodulin; CysC, urine cystatin C; uOSM, urine osmolarity.

Outcome Definitions

We evaluated three main outcomes: composite kidney outcome (CKD incidence among those without preexisting CKD or progression among those with preexisting CKD), cardiovascular events, and death. Among patients without underlying CKD, CKD incidence was defined as >25% reduction in the eGFR from baseline or reaching CKD 3 or worse during follow-up. Among patients with preexisting CKD at the index hospitalization (preadmission eGFR <60 ml/min per 1.73 m^2), CKD progression was defined as >50% reduction in the baseline eGFR or progression to CKD 5 or development of kidney failure on follow-up (hemodialysis or peritoneal dialysis requirement at least once/week for >12 weeks or receiving a kidney transplant and/or death while on dialysis). Cardiovascular events were a composite of myocardial infarction, heart failure, cerebrovascular accidents, peripheral artery disease, or cardiovascular interventions. Death was ascertained from proxy reports or medical records.

Data Preprocessing and Subphenotype Discovery

First, for numerical variables, outliers were adjusted using 95% Winsorization. Second, variables showing left-skewed

distribution were log-transformed and examined whether the mode leans to the median rather than the first quartile. Third, variables were normalized using a robust scaler.[31] Finally, multivariable imputation by chained equations[32] was used with five multiple imputations and 50 iterations, accounting for the statistical uncertainty in the imputations, while the chained equations allowed us to handle a wide array of variables.

Categorical and numerical variables were transformed into a lower-dimensional space using factor analysis of mixed data.[33,34] The use of a total of 53 variables can affect clustering through sparseness of variables and multicollinearity. Although several methods exist for dimensionality reduction, we used factor analysis of mixed data, a recently developed principal component analysis that simultaneously explores multivariable dependencies between various categorical and numerical variables. We selected 14 of 53 principal components (eigenvalues >1) on the basis of the Kaiser–Guttman criterion.

Unsupervised Clustering Analysis

We conducted hierarchical agglomerative clustering in which each observation starts in its own cluster and pairs of clusters are merged as one moves up the hierarchy. We used Ward linkage with Euclidean distance applied to the 14 principal components. A small number of clusters would lead to more homogeneous features among AKI subphenotypes, although not fully reflecting the complexity seen in clinical practice. Conversely, multiple clusters could dilute the influence of biomarker abnormalities and be difficult to understand. Therefore, the number of subphenotypes was determined through the visual evaluation of the dendogram (Supplemental Figure 1) and using results from 26 indexes available from the NbClust package in R, an approach that has been used to identify the optimal number of clusters.[34]

Four clusters were derived. Uniform Manifold Approximation and Projection was used for dimensionality reduction and cluster visualization revealing adequate independence (Supplemental Figure 1B). In addition, the robustness of these four clusters was established by comparing with another unsupervised algorithm (K-means clustering), visualized through the Uniform Manifold Approximation and Projection and heatmap in Supplemental Figures 2 and 3, confirming the presence of four clusters. The presence and stability of four clusters was further reproduced through consensus clustering (Supplemental Figure 4). Agreement and cluster membership between consensus and hierarchical clustering is presented in Supplemental Table 1. Finally, Shapley additive explanation (SHAP) diagrams were obtained to assess the importance of intervening features on each subphenotype by comparing the XGBoost models with and without the presence of such specific features (Figure 1).

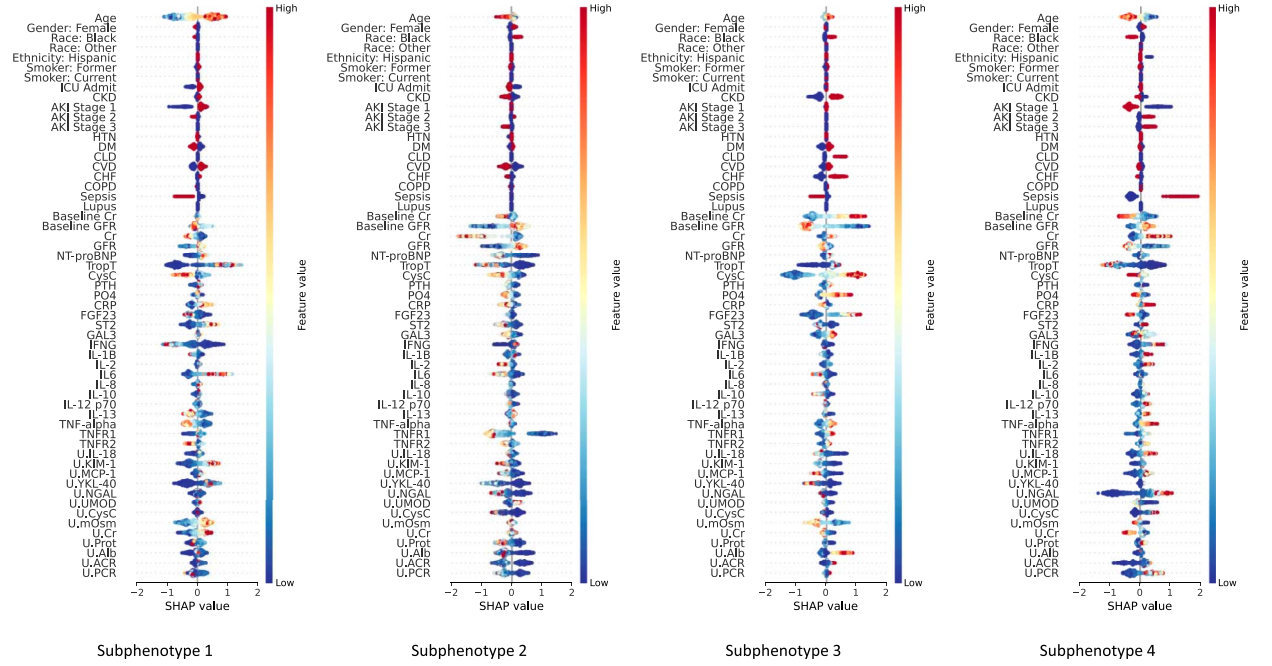

Figure 1. SHAP diagram comprising all variables included in subphenotype discovery and modeling. The SHAP diagram represents the significance of each observation toward subphenotype membership. Specifically, the color of each dot represents normalized values, and the SHAP value represents the effect of the observation on the subphenotype membership. For instance, higher TNFR1 gives a negative effect and lower TNFR1 gives a positive effect on subphenotype 1 membership. CHF, congestive heart failure; CLD, chronic liver disease; COPD, chronic obstructive pulmonary disease; Cr, creatinine; CVD, cardiovascular disease; CysC, cystatin C; DM, diabetes mellitus; FGF23, fibroblast growth factor 23; GAL3, galectin 3; HTN, hypertension; ICU, intensive care unit; IFNG, interferon gamma; NT-proBNP, NT-pro–brain natriuretic peptide; PO4, plasma phosphorus; PTH, parathyroid hormone; SHAP, Shapley additive explanation; ST2, suppression of tumorigenicity 2 receptor; TNFR, TNF receptor; TNFR1, TNFR2, plasma tumor necrosis factor receptor-1 and -2; TropT, plasma troponin T; U.ACR, urinary albumin-to-creatinine ratio; U.Alb, urine albumin; U.Cr, urine creatinine; U.IL-18, urine IL-18; U.KIM-1, urine kidney injury molecule; U.MCP-1, urine MCP-1; U.mOsm, urine osmolality; U.NGAL, urine neutrophil gelatinase–associated lipocalin; U.PCR, urinary protein-to-creatinine ratio; U.Prot, urine protein; U.UMOD, urine uromodulin; U.YKL-40, urine chitinase-3–like protein 1.

Table 2. Clinical characteristics among the four novel AKI subphenotypes

Patient Characteristics	1 (*N*=181)	2 (*N*=250)	3 (*N*=159)	4 (*N*=158)
Age, mean (SD)	71 (9)	60 (13)	67 (12)	57 (12)
AKI stages, *n* (%)				
Stage 1	165 (91)	205 (82)	119 (75)	54 (34)
Stage 2	13 (7)	37 (15)	15 (9)	47 (30)
Stage 3	3 (2)	8 (3)	25 (16)	57 (36)
Baseline creatinine, mean (SD), mg/dl	1.1 (0.3)	1 (0.3)	1.9 (0.6)	1.1 (0.4)
Baseline eGFR, mean (SD), mg/dl	68 (18)	78 (20)	39 (17)	76 (25)
Body mass index in kg/m^2, mean (SD)	29.9 (6.3)	32.2 (7.5)	31.8 (7.9)	31.4 (8.7)
Cardiovascular disease, *n* (%)	130 (72)	83 (33)	100 (63)	46 (29)
Chronic liver disease, *n* (%)	3 (2)	9 (4)	15 (9)	9 (6)
Chronic obstructive pulmonary disease, *n* (%)	32 (18)	60 (24)	50 (31)	36 (23)
Critical care, *n* (%)	170 (94)	150 (60)	95 (60)	115 (73)
Female, *n* (%)	33 (18)	88 (35)	55 (35)	64 (41)
Heart failure, *n* (%)	43 (24)	42 (17)	88 (55)	24 (15)
Hypertension, *n* (%)	139 (77)	199 (80)	140 (88)	110 (70)
Peak creatinine, median (interquartile range)	1.09 (0.9–1.2)	1 (0.9–1.2)	1.8 (1.3–2.5)	1.02 (0.8–1.2)
Peak eGFR, median (Q1, Q3), ml/min per 1.73 m^2	67 (55, 83)	77 (62, 91)	36 (27, 48)	76 (59, 95)
Race, *n* (%)				
Black	5 (3)	57 (23)	38 (24)	11 (7)
White	174 (96)	180 (72)	112 (70)	128 (81)
Other	2 (1)	13 (5)	9 (6)	19 (12)
Sepsis, *n* (%)	6 (3)	22 (9)	6 (4)	81 (51)
Smoking status, *n* (%)				
Never	59 (33)	109 (44)	69 (43)	70 (44)
Former	108 (60)	96 (38)	77 (48)	54 (34)
Current	14 (8)	45 (18)	13 (8)	34 (22)
Systemic lupus erythematosus, *n* (%)	0 (0)	0 (0.0)	0 (0)	6 (4)
Type 2 diabetes, *n* (%)	70 (39)	119 (48)	112 (70)	74 (47)
Urine albumin-creatinine ratio at AKI, median (Q1, Q3), g/g	0.48 (0.2, 0.9)	0.22 (0.1, 0.4)	0.99 (0.4, 5.8)	1.01 (0.6, 3.5)

Statistical Analyses

We reported descriptive statistics for continuous variables as mean (SD) or median (interquartile range) accordingly for each of the subphenotype. We presented categorical variables as frequencies and percentages. Clinical and biomarker concentrations among subphenotypes were evaluated by the ANOVA or Chi-square test. Cox proportional hazard regression models were used to compare time to events. These models were fully adjusted for age, sex, diabetes mellitus (yes/no), body mass index, baseline CKD (yes/no), and baseline urine albumin-creatinine ratio (UACR). In further analysis, models were adjusted for appropriate timing for biomarker collection, defined as collection within 24 hours relative from the peak creatinine date and AKI recovery. Two-tailed *P* values of < 0.05 were used. We performed all statistical analyses using R software version 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria) and Python 3.9.3. Consent was obtained from all participants at the time of enrollment.

Results

Clinical Features of Patients with AKI

Among 769 participants included in ASSESS-AKI with AKI, 748 had available data of interest and were included in the analysis. The median follow-up time for clinical events was 4.8 years. Participants were predominantly men (68%) and White (79%) and had a mean age (±SD) of 64 (13) years. The prevalence of baseline comorbidities is presented in Table 2 and Supplemental Figure 5. The overall mean (±SD) baseline eGFR and UACR were 75 (8) ml/min per 1.73 m^2 and 2.2 mg/g, respectively. Supplemental Table 2 summarizes that patients whose biomarkers were collected within 24 hours of peak creatinine obtention were largely comparable with those whose biomarkers were collected beyond this time frame.

Clinical and Biological Signatures among AKI Subphenotypes

Subphenotype 1 (*N*=181) was characterized by the oldest individuals—those with a mean (±SD) age of 71 (9) years ($P < 0.001$) and with the highest intensive care unit (ICU) admission rates and baseline cardiovascular disease (170 [94%] and 130 [72%], $P < 0.001$) among all subphenotypes. Most individuals had stage 1 AKI (91%). Subphenotype 2 (*N*=250) was characterized by individuals with a low prevalence of comorbid conditions, such as CKD (18%, $P < 0.001$). Individuals in subphenotype 3 (*N*=159) had the highest prevalence of CKD (143 [90%], $P < 0.001$) and congestive heart failure (88 [55.3%], $P < 0.001$) and were African American (38 [24%], $P < 0.001$). Patients in subphenotype 4 were the youngest, with a high incidence of sepsis and ICU admission. Notably, approximately one in every three patients had stage 3 AKI including dialysis.

Individuals with AKI in subphenotype 1 (cardiorenal injury) were characterized by the highest troponin T, kidney injury molecule (KIM-1), and urine IL-18 levels recorded among all clusters. In addition, biomarkers of cardiac congestion (pro–brain natriuretic peptide) and remodeling (suppression of tumorigenicity 2 receptor) were also higher compared with subphenotype 2. Patients in subphenotype 2 (benign cluster) were characterized by the highest levels of uromodulin, a biomarker of tubular

Table 3. Biomarker characteristics within each AKI subphenotypes (systemic inflammatory)

Subphenotype	1 (*N*=181)	2 (*N*=250)	3 (*N*=159)	4 (*N*=158)	*P*
Median (IQR)					
INF-γ	1.5 (0.9–2.6)	3.8 (1.9–8.8)	3.9 (2–8.6)	7.6 (2.5–32.9)	<0.001
IL1b	0.2 (0.1–0.3)	0.2 (0.1–0.3)	0.2 (0.1–0.3)	0.3 (0.2–0.5)	<0.001
IL2	0.3 (0.3–0.5)	0.3 (0.2–0.5)	0.4 (0.3–0.5)	0.5 (0.3–1.1)	<0.001
IL6	35.7 (15.9–66.2)	6.6 (2.7–15)	7.4 (3.7–14.7)	20.1 (9.6–56.4)	<0.001
Pro-BNP	3421 (1545–5627)	643 (217–1617)	3569 (1503–7342)	2492 (812–4039)	<0.001
Troponin T	372 (125–679)	32 (19–114)	83 (40–274)	27 (20–48)	<0.001
ST2	245 (128–463)	62 (34–117)	70 (42–126)	177 (75–518)	<0.001
Gal3	17 (13–22)	14 (11–19)	29 (22–38)	21 (15–26)	<0.001
FGF23	28 (17–47)	34 (20–64)	140 (82–241)	36 (16–78)	<0.001
TNFR1	5121 (3942–6514)	3476 (2568–4784)	7956 (5694–11,384)	7615 (5100–10,616)	<0.001
TNFR2	9916 (7295–13,423)	8181 (5723–11,098)	16,011 (11,928–20,747)	16,601 (11,274–26,406)	<0.001
Urine IL18	63 (32–116)	27 (15–51)	23 (10–47)	96 (50–184)	<0.001
KIM-1	7142 (3483–12,806)	1749 (741–3283)	1496 (811–2865)	3981 (1971–7541)	
MCP-1	701 (370–1284)	331 (154–614)	285 (124–663)	1177 (488–2162)	
YLK 40	3234 (1115–9779)	635 (266–1326)	730 (174–2980)	4459 (1649–15,150)	<0.001
NGAL	81 (43–154)	33 (15–68)	57 (22–179)	269 (112–740)	<0.001
UMOD	2370 (1722–3990)	2803 (1620–4753)	1927 (951–3035)	2003 (1142–4323)	<0.001

IQR, interquartile range; IFN-γ, interferon-gamma; NT-proBNP, plasma pro–brain natriuretic peptide, N-terminal; ST2, plasma suppression of tumorigenicity 2 receptor; GAL3, plasma galectin-3; FGF23, fibroblast growth factor 23; TNFR1, TNFR2, plasma tumor necrosis factor receptor-1 and -2; KIM-1, urine kidney injury molecule; MCP-1, urine MCP-1; YKL-40, urine chitinase-3–like protein 1; NGAL, urine neutrophil gelatinase–associated lipocalin; UMOD, urine uromodulin.

health. Patients with AKI in subphenotype 3 (AKI on CKD) were characterized by a marked cardiac congestion compared with injury. In addition, these patients had the highest concentrations of plasma tumor necrosis factor receptor-1 and -2 (TNFR1 and TNFR2, respectively) from all subphenotypes. Interestingly, biomarkers of kidney injury, such as IL-18, KIM-1, and neutrophil gelatinase–associated lipocalin (NGAL), were low and comparable with the benign cluster. Patients with AKI subphenotype 4 (sepsis AKI) were characterized by high concentrations of acute systemic (INF-γ, TNF-α), vasculature/kidney-related inflammation (YKL-40, MCP-1), and high injury pathway activation (urine NGAL, KIM-1) (Figure 2, Table 3, and Supplemental Figure 6). The data for the quartiles of the log-transformed biomarkers across all four subphenotypes can be found in Supplemental Table 3.

Association between AKI Subphenotypes and Clinical Outcomes

Over a median of 4.8 years of follow-up, there were 38 deaths per 1000 person-years, 38.2 CKD events per 1000 person-years, and 38.8 cardiovascular events per 1000 person-years (Table 4). AKI subphenotype 3 and subphenotype 4 were independently associated with death, with adjusted hazard ratios (HRs) of 2.9 (95% confidence interval [CI], 1.8 to 4.6) and 1.6 (95% CI, 1.01 to 2.6), respectively, compared with subphenotype 2 as the reference. Subphenotype 3 was independently associated with CKD events (adjusted HR, 2.6; 95% CI, 1.6 to 4.2) in fully adjusted models (Figure 3, Table 4). Furthermore, subphenotype 3 was independently associated with cardiovascular disease (adjusted HR, 2.6; 95% CI, 1.6 to 4.1). Supplemental Figures 7–9 show the K–M curves for each outcome and subphenotype. Supplemental Figures 10 and 11 show that these models remained prognostic after further adjustment for biomarker collection timing and AKI recovery during hospitalization. Furthermore, the association between subphenotypes with CKD and cardiovascular outcomes was not affected when considering death as a competing risk. The SHAP global bar plot showing a mean absolute SHAP value for the top 20 important features is presented in Supplemental Figure 12.

Discussion

We uncovered four distinct AKI subphenotypes with unsupervised machine learning algorithms. These clusters are characterized by unique clinical and biomarker signatures reflective of different pathways of disease and health. Individual biomarkers were statistically different from each other and informed on the complexity of AKI subphenotypes through an unbiased approach. Two subphenotypes were independently associated with up to three-fold higher risk of death. Another subphenotype was independently associated with longitudinal CKD outcomes and cardiovascular disease events.

AKI has been defined on the basis of the RIFLE, AKI Network, and, most recently, Kidney Disease Improving Global Outcomes criteria using serum creatinine and urine output changes.[35] While this has increased our capacity to identify AKI, molecular changes are present before clinically meaningful serum creatinine is noticed.[36,37] Tubular-interstitial markers have been demonstrated to associate with AKI duration, dialysis needs, and protracted kidney outcomes in large prospective studies.[3,4,28] Some of these pathway abnormalities may persist after discharge and may play a role in kidney events in both patients with AKI and non-AKI patients.[4,13,28,30] Therefore, understanding the fine balance between reparative and maladaptive processes in AKI could be key to identifying the risk of progression.[38] However, conceptualizing the complex nature of AKI and biological processes leading to distinct clinical trajectories is challenging for the clinician.[39] This is in part because of the relationships between patient-specific characteristics and

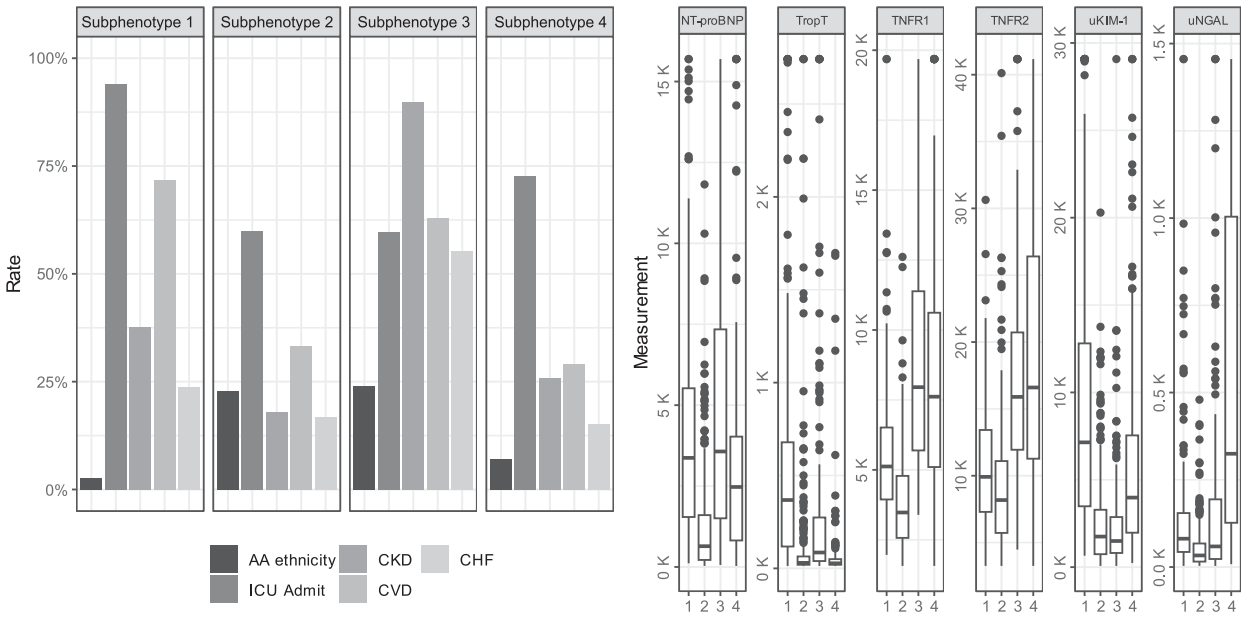

Figure 2. Key clinical characteristics (left) and selected biomarker composites (right) show statistically significant differences between each subphenotype. Selected clinical composites by subphenotypes (left). Rates of clinical comorbid conditions are reflected in the bar graph (left). Selected biomarker composites by AKI subphenotype. Biomarker concentrations were adjusted for better visualization (right). AA, African American.

clinically relevant end points, which are often nonlinear, vary over time, and are typically not well described by one data domain alone.[40–43] We demonstrated that machine learning algorithms allow us to deal with such data heterogeneity in an unbiased manner and discover clusters with unique clinical and molecular characteristics that allow us to gather deeper information on the diagnosis of AKI while still providing a prognostic value.

AKI subphenotype 1 (cardiorenal injury) was overrepresented by the eldest patients with prevalent cardiovascular comorbidities with highest mean levels of troponin T, KIM-1, and urine IL-18. Among these patients, cardiac injury seemed to be comparatively more prominent than cardiac congestion, and biomarkers of cardiac remodeling (suppression of tumorigenicity 2 receptor) were higher compared with the control group (AKI subphenotype 2). Yet, most of these patients had stage 1 AKI and low peak serum creatinine levels. These findings highlight the importance of incorporating additional tools for estimating kidney function in patients susceptible to muscle mass remodeling.[44] AKI subphenotype 1 could represent several clinical scenarios, including patients with myocardial infarction or pulmonary embolism who are susceptible to hemodynamic instability or patients with active cardiovascular disease undergoing contrast-based studies.[45]

Patients in AKI subphenotype 3 (AKI on CKD) had the highest prevalence of baseline of congestive heart failure, CKD, and diabetes mellitus. These patients had the highest concentrations of pro–brain natriuretic peptide and TNFR1 among all subphenotypes and the lowest baseline eGFR. Although these patients had the highest peak serum creatinine levels during hospitalization, they had low concentrations of urine IL-18, NGAL, and KIM-1, comparable with individuals in AKI subphenotype 2. We hypothesize that this scenario is plausible in the context of cardiorenal AKI, where creatinine fluctuations do not necessarily

Table 4. Hazard ratios (95% confidence interval) for each subphenotype and the risks of longitudinal clinical events

Subphenotype	Mortality			CKD Outcomes			CVD Events		
	Event Rate per 1000 PY	HR (95% CI)	*P*	Event Rate per 1000 PY	HR (95% CI)	*P*	Event Rate per 1000 PY	HR (95% CI)	*P*
Subphenotype 2 (ref)	26.3	1 (ref)	NA	30.8	1 (ref)		29.9	1 (ref)	NA
Subphenotype 1	39	1.4 (0.8 to 2.2)	0.16	33.3	1 (0.7 to 1.5)	0.90	41.9	1.2 (0.8 to 1.9)	0.3
Subphenotype 3	73	2.9 (1.8 to 4.6)	<0.001	61.5	2.6 (1.6 to 4.2)	<0.001	66.7	2.6 (1.6 to 4.1)	<0.002
Subphenotype 4	29.4	1.6 (1.01 to 2.6)	0.04	37.5	0.8 (0.5 to 1.2)	0.3	28.4	0.9 (0.6 to 1.6)	0.9

Models adjusted for age, sex, diabetes mellitus (yes/no), body mass index, baseline CKD (yes/no), and baseline urine albumin-creatinine ratio. CVD, cardiovascular disease; PY, patient-year; HR, hazard ratio; CI, confidence interval.

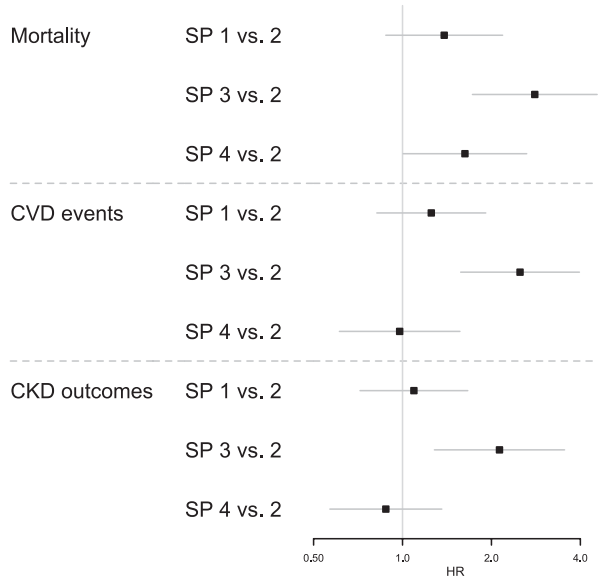

Figure 3. Forest plot analysis showing the 95% confidence interval–adjusted hazard ratios (HRs) for all outcomes among all subphenotypes. Forest plot of mortality, CVD events, and CKD outcomes for models adjusted for age, sex, diabetes mellitus (yes/no), body mass index, baseline CKD (yes/no), and baseline urine albumin-creatinine ratio. SP2 was the reference group. SP, subphenotype.

follow from marked acute injury marker elevations.[45] Surprisingly, AKI subphenotype 3 was independently associated with the highest risk of poor outcomes, including death, CKD outcomes, and cardiovascular events. While TNFR was not among the top classifiers for subphenotype 3, elevated TNFR levels have been found to be prognostic of cardiovascular disease, kidney disease progression, and death among patients in the pre- and post-AKI stages as well as in patients with stable CKD with and without diabetes mellitus.[29,30,46] These high levels could be produced in the heart and vasculature while not necessarily driving worse AKI pathophysiology, as demonstrated in the low rates of AKI stages 2 and 3 that we found. Yet, given their capability to orchestrate diverse pathological cellular changes in different organs, TNFRs could activate redundant pathways leading to death.[46] AKI subphenotype 3 could be seen in patients with marked heart and kidney disease, who present with cardiorenal syndrome and multiple hospital admissions because of volume excess. Thus, expanding prior data,[47] it is possible that patients with creatinine fluctuations in the context of ischemic or hemodynamic changes (*i.e.*, cardiorenal syndrome) without significant tubular injury activity could still present with higher morbidity and risk of clinical events.

AKI subphenotype 4 was characterized by patients with sepsis-AKI. They had the highest concentrations of biomarkers of acute systemic inflammation and markedly elevated levels of biomarkers of vascular/kidney inflammation (MCP-1, YKL-40) and injury (KIM-1, NGAL) as expected in a scenario of physiological and inflammatory stress. Although the burden of preexisting comorbidities was not apparent in this group, the severe physiological stress elicited during sepsis and the pathological pathways activated likely reflect a different pathophysiology and phenotype of AKI that requires more attention prospectively.[48]

Our findings provide insight into the clinical relevance of measuring biomarkers during the acute setting, especially when revealing two subphenotypes with the highest mortality and kidney events among an already high-risk AKI population. Biomarker composites were statistically different within each subphenotype, suggesting that different degrees of inflammation, injury, and fibrosis/repair could identify the nature of AKI and inform about future clinical events. In the advent of new therapies that modulate inflammation/injury-predominant–type kidney disease (*i.e.*, sodium-glucose cotransporter-2 inhibitors) or fibrosis (*i.e.*, mineralocorticoid receptor antagonists), identifying these pathologic domains could guide therapeutic targets in the AKI or post-AKI stages. Our analyses were robust and analyzed through different unsupervised algorithms, setting an appropriate balance between unbiased approach and interpretability of data. While these studies do not inform on the underlying pathophysiology, this study is novel by incorporating a broad set of systemic and kidney-related biomarkers involved in kidney disease progression and associated outcomes. Future studies may attempt to explore whether individuals within the same AKI subphenotype could share similar proteins and transcripts responsible for the biomarker composites, which could explain future clinical outcomes.

A previous report has demonstrated the presence of two parsimonious AKI models: one driven by congestive heart failure and a benign biomarker profile and the other characterized by prevalent CKD and an unfavorable biomarker profile after analyzing 29 variables by latent class analyses and K-mean clustering.[49] The present study was characterized by more ample clinical and biomarker data (53 variables), which could have resulted in different clustering representation. In detail, we applied several sophisticated methods to transform the raw data into different dimensionality, allowing us to use extensive mixed-type variables and tackle two typical high-dimensionality problems (sparseness and multiple chain imputations). These efforts turned into accessing granular subphenotypes where each cluster was well separated from the other (Supplemental Figure 1). However, the number and cluster characteristics largely depend on feature input and intrinsic qualities of the algorithm, explaining different cluster solutions. The goal of this study is unique and complementary to previous data because we aimed to identify potentially useful data patterns over statistically and numerally significant features that emulate real-world data abundance and assist with the enrichment of AKI diagnosis. However, we acknowledge the potential trade-off with loss of power for time-to-event analyses by deriving four subphenotypes instead of two. The heterogeneity and clinical relevance found in this study should be confirmed in additional studies with larger sample sizes.

Regarding limitations, AKI definition, inclusion criteria, and exclusion criteria were stringent and only patients who survived to the 3-month follow-up visit were included in the longitudinal data collection, inducing selection bias. Furthermore, despite the heterogeneity of our AKI

population (medical/surgical floors, ICU), a high proportion of patients with AKI were in stage 1, which could carry a lower risk of CKD and limit the ability to conduct analyses on those with stage 2 and 3 AKI. In addition, biospecimens were collected at varying time points relative to the peak of serum creatinine. Although, when our models were adjusted for timing of biomarker collection and AKI recovery, they remained prognostic for clinical events. Owing to the uniqueness of this study, it was challenging to identify an appropriate validation cohort and reproduce our findings. However, we have posted our source code on GitHub (https://github.com/Nadkarni-Lab/george_cjasn_2022) to make the subtyping machine learning pipeline publicly available. In addition, we posted an XGBoost model for subtypes on Github so that subtypes are reproducible in the future (https://github.com/Nadkarni-Lab/george_cjasn_2022/blob/main/data/xgboost.pkl). Of note, becacuse this is a data-driven approach, data pattern aggregation and cluster membership depend on the input of data and intrinsic characteristics of the analysis, leading to diverse ensemble solutions. If the input variables were to be modified, we could expect different AKI subphenotypes. Therefore, our subphenotypes may not necessarily be reproduced in other settings (*i.e.*, coronavirus disease 2019, only critical care settings, *etc.*) or accurately discriminate among subtypes that are distinct from one another in pathophysiology and treatment. Therefore, tissue analysis from biopsies would be needed. Finally, it is uncertain whether the additional information provided by the inclusion of novel biomarkers would alter treatment or outcomes over routinely available clinical and biomarker information.

We discovered four novel, clinically meaningful, and statistically different AKI subphenotypes characterized by distinct pathway abnormalities. These clinical and biomarker signatures were independently associated with adverse clinical events and death. Our study reflects a novel approach to precision medicine in AKI and unveils a new role for biomarkers when incorporated into agnostic multidimensional analyses.

Disclosures

P. Bhatraju reports research funding from Roche Diagnostics. V.M. Chinchilli reports advisory or leadership roles for Allergan, AstraZeneca, Biohaven, Janssen, Regeneron, and Sanofi. S.G. Coca reports employment with Icahn School of Medicine at Mount Sinai, which owns part of Renalytix. S.G. Coca reports consultancy for 3ive, Axon Therapeutics, Bayer, Boehringer-Ingelheim, CHF Solutions, ProKidney, Renalytix, Reprieve Cardiovascular, Takeda, and Vifor; personal income and equity and stock options from pulseData and Renalytix; research funding from ProKidney, Renalytix, RRI, and XORTX; patents or royalties from Renalytix; advisory or leadership roles for Renalytix and Reprieve Cardiovascular; serving as an Associate Editor of *Kidney360*; and serving on the Editorial Boards of *CJASN*, *JASN*, and *Kidney International*. A.X. Garg reports research funding from Astellas and Baxter; currently serving on the Editorial Boards of the *American Journal of Kidney Diseases* and *Kidney International*; and served as the Medical Lead Role to Improve Access to Kidney Transplantation and Living Kidney Donation for the Ontario Renal Network (government-funded agency located within Ontario Health)—this position ended in October 2022. A.S. Go reports employment with Kaiser Permanente Northern California and research funding from Amarin Pharmaceuticals, Bristol Meyers Squibb, CSL Behring, Janssen Research and Development, Novartis, and Pfizer. J. Himmelfarb reports consultancy for Maze Therapeutics; ownership interest in Kuleana Technology, Inc.; research funding from Aurinia Pharmaceuticals; honoraria from various academic institutions for invited lectures; patents held and patents pending owned by the University of Washington; advisory or leadership roles for *CJASN* (Editorial Board), *BMC Medicine* (Editorial Board), and *Nature Reviews Nephrology* (Advisory Board); and research grant support from Northwest Kidney Centers. J. Himmelfarb is the founder, CEO, President, and equity holder of Kuleana Technology, Inc. C.-Y. Hsu has consulted for legal cases involving acute or chronic kidney disease and consults on an *ad hoc* basis for companies regarding kidney disease (including being on the steering committee of an industry-funded trial). C.-Y. Hsu reports research funding from Satellite Healthcare and royalties from UpToDate. J.S. Kaufman reports employment with VA New York Harbor Healthcare System; consultancy for National Kidney Foundation and Otsuka Pharmaceutical; ownership interest in Amgen; advisory or leadership roles for the National Institute of Diabetes and Digestive and Kidney Diseases, paid Steering Committee Chair; and serving as an Associate Editor of the *American Journal of Kidney Disease*. P.L. Kimmel reports employment with the National Institute of Diabetes and Digestive Kidney Diseases (NIDDK); stock in CVS; royalties from Elsevier for coediting *Chronic Renal Disease* and *Psychosocial Aspects of Chronic Kidney Disease*; advisory or leadership role as an unpaid member of the Board of Directors of Academy of Medicine of Washington, DC; and royalties as Co-Editor of *Chronic Renal Disease* and Co-Editor of *Psychosocial Aspects of Chronic Kidney Disease*. K.D. Liu reports consultancy for AM Pharma, Biomerieux, BOA Medical, Neumora, and Seastar Medical; stock in Amgen; role on the Editorial Boards of the *American Journal of Respiratory and Critical Care Medicine*, the *American Journal of Kidney Disease*, and *CJASN*; advisory or leadership roles for the American Thoracic Society and the NKF Scientific Advisory Board; and other interests or relationships with UpToDate. D.G. Moledina reports ownership interest in Predict AIN, LLC; research funding from NIDDK K23DK117065, R01DK12681, R01128087, UH3DK114866, and P30DK079310; and honoraria from *British Medical Journal*, National Kidney Foundation, and Remedy Health Media. D.G. Moledina is a coinventor of the pending patent application "Methods and Systems for Diagnosis of Acute Interstitial Nephritis" and is an editorial board member for *Kidney360*. G.N. Nadkarni reports employment with Icahn School of Medicine at Mount Sinai; consultancy for Daiichi Sankyo, GLG consulting, Qiming Capital, Reata, Renalytix, Siemens Healthineers, and Variant Bio; ownership interest in Data2Wisdom LLC, Doximity, Nexus iConnect, Pensieve Health, Renalytix, and Verici; research funding from Renalytix; honoraria from Daiichi Sankyo; support by R01DK108803, U01HG007278, U01HG009610, and 1U01DK116100; personal income and equity and stock options from Renalytix and pulseData; operational funding from Goldfinch Bio; personal income from AstraZeneca, BioVie, GLG consulting, Goldfinch Bio, Reata, Siemens Healthineers, and Variant Bio, in the past 3 years; patents or royalties from Renalytix; an advisory or leadership role for Renalytix; and speakers bureau for Daiichi Sankyo. G.N. Nadkarni is a scientific cofounder of Pensieve Health and Renalytix. C.R. Parikh reports

consultancy for Genfit Biopharmaceutical Company; personal income and equity and stock options from Renaltix AI; personal income from Akebia Therapeutics and Genfit Biopharmaceutical Company; research funding from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) and the National Heart, Lung and Blood Institute (NHLBI); and advisory or leadership roles for Genfit Biopharmaceutical Company and Renalytix. P. Sarder reports research funding from NIH—NIDDK, OD, and NCI; serving as an Associate Editor of *Plos One*; and serving as an Editorial Board Member of *JASN*. E.D. Siew reports employment with Nashville Veterans Affairs; stock in Apple and Amazon; royalties as an author for UpToDate; personal income from Akebia Therapeutics, DaVita, and UpToDate; and serving as an Associate Editor of *CJASN*. M.M. Wurfel reports consultancy for and research funding from Roche Diagnostics. All remaining authors have nothing to disclose.

Funding

The ASSESS-AKI was supported by cooperative agreements from the NIDDK (U01DK082223, U01DK082185, U01DK082192, and U01DK082183). The authors also acknowledge funding support from NIH grants R01HL085757, R01DK098233, R01DK101507, R01DK114014, K23DK100468, R03DK111881, R01DK093771, K01DK120783, K24DK113381, K24DK092291, R01DK114014, P30DK079310, and P30DK114809. S.G. Coca has salary support from NIH grants R01DK115562, UO1DK106962, R01HL085757, R01DK112258, R01DK115562, R01DK126477, and UH3DK114920. W. Oh is supported by grant T32 DK007757 from NIH NIDDK. D. Moledina reports researching funding from NIH grant P30DK079310.

Acknowledgments

The authors thank all the ASSESS-AKI study participants, research coordinators, and support staff for making this study possible. The funder did not have any role in study design, data collection, analysis, reporting, or the decision to submit for publication.

George Vasquez-Rios and Wonsuk Oh contributed equally to the elaboration and intellectual content. All the ASSESS-AKI investigators participated in the design of the study, collection, and analysis of the data.

Because Dr. Edward D. Siew is an Associate Editor of *CJASN*, he was not involved in the peer-review process for this manuscript. Another editor oversaw the peer-review and decision-making process for this manuscript.

Author Contributions

Conceptualization: Steven G. Coca, Girish N. Nadkarni.
Data curation: Amit X. Garg, Alan S. Go, Jonathan Himmelfarb, Chi-yuan Hsu, Paul L. Kimmel, Kathleen D. Liu, Sherry G. Mansour, Dennis G. Moledina, Chirag R. Parikh, Edward D. Siew, Mark M. Wurfel.
Formal analysis: Faris F. Gulamali, Samuel Lee, Girish N. Nadkarni, Wonsuk Oh.
Funding acquisition: Amit X. Garg, Jonathan Himmelfarb, James S. Kaufman, Chirag R. Parikh, Edward D. Siew.
Investigation: Pavan Bhatraju, Steven G. Coca, Girish N. Nadkarni, Wonsuk Oh, Chirag R. Parikh, George Vasquez-Rios.
Methodology: Vernon M. Chinchilli, Steven G. Coca, Amit X. Garg, Alan S. Go, Chi-yuan Hsu, James S. Kaufman, Paul L. Kimmel, Samuel Lee, Kathleen D. Liu, Sherry G. Mansour, Wonsuk Oh, Pinaki Sarder, George Vasquez-Rios, Mark M. Wurfel.
Project administration: Vernon M. Chinchilli, Girish N. Nadkarni, Wonsuk Oh.
Resources: Pavan Bhatraju, Vernon M. Chinchilli, Steven G. Coca, Amit X. Garg, Alan S. Go, Jonathan Himmelfarb, Chi-yuan Hsu, James S. Kaufman, Paul L. Kimmel, Kathleen D. Liu, Sherry G. Mansour, Dennis G. Moledina, Girish N. Nadkarni, Chirag R. Parikh, Mark M. Wurfel.
Software: Faris F. Gulamali, Girish N. Nadkarni, Wonsuk Oh.
Supervision: Steven G. Coca, Girish N. Nadkarni.
Validation: Vernon M. Chinchilli, Faris F. Gulamali, Samuel Lee, Wonsuk Oh, Pinaki Sarder, Edward D. Siew.
Visualization: Faris F. Gulamali, Samuel Lee, Pinaki Sarder.
Writing – original draft: Steven G. Coca, Girish N. Nadkarni, Wonsuk Oh, George Vasquez-Rios.
Writing – review & editing: Pavan Bhatraju, Vernon M. Chinchilli, Steven G. Coca, Amit X. Garg, Alan S. Go, Faris F. Gulamali, Jonathan Himmelfarb, Chi-yuan Hsu, James S. Kaufman, Paul L. Kimmel, Samuel Lee, Kathleen D. Liu, Sherry G. Mansour, Dennis G. Moledina, Girish N. Nadkarni, Wonsuk Oh, Chirag R. Parikh, Pinaki Sarder, Edward D. Siew, George Vasquez-Rios, Mark M. Wurfel.

Data Sharing Statement

Analyzable data have been deposited to NIDDK Repository, https://repository.niddk.nih.gov/studies/assess-aki/

Supplemental Material

This article contains the following supplemental material online at http://links.lww.com/CJN/B750.

Supplemental Table 1. Contingency table: consensus clustering (Y axis) versus hierarchical clustering (X axis).

Supplemental Table 2. Clinical (above) and biomarker (below) differences between patients with and without biomarker collection within 24 hours of peak creatinine collection.

Supplemental Table 3A. Biomarker concentrations by quartiles within each AKI subphenotypes (systemic inflammatory).

Supplemental Table 3B. Biomarker concentrations within each AKI subphenotypes (cardiac injury, congestion, and repair/fibrosis).

Supplemental Table 3C. Biomarker concentrations within each AKI subphenotypes (kidney inflammation, injury, health, and metabolism).

Supplemental Figure 1. Hierarchical clustering dendogram (left). UMAP shows density-based clustering and dimensionality reduction (right).

Supplemental Figure 2. Heatmap showing the robustness of four clusters by comparison between hierarchical clustering and K-mean clustering.

Supplemental Figure 3. Hierarchical clustering (left) versus K-mean clustering (right) to demonstrate the robustness of four subphenotypes.

Supplemental Figure 4. Consensus clustering demonstrates 2–6 subphenotypes.

Supplemental Figure 5. Cluster description on the basis of categorical features.

Supplemental Figure 6. Cluster description on the basis of numeric features with robust scaler transformation.

Supplemental Figure 7. Kaplan–Meier and cumulative incidence curves for mortality per each subphenotype in adjusted models.

Supplemental Figure 8. Kaplan–Meier and cumulative incidence curves for CKD outcomes per subphenotype in adjusted models.

Supplemental Figure 9. Kaplan–Meier and cumulative incidence curves for CVD events per each subphenotype in adjusted models.

Supplemental Figure 10. Forest plot analyses when subphenotypes are adjusted for timing for biomarker collection (within 24 hours of peak creatinine).

Supplemental Figure 11. Forest plot analyses when subphenotypes are adjusted for AKI recovery.

Supplemental Figure 12. SHAP global bar plot showing the mean absolute SHAP value for the top 20 important features, ordered by feature importance.

References

1. Kellum JA, Romagnani P, Ashuntantang G, Ronco C, Zarbock A, Anders HJ. Acute kidney injury. *Nat Rev Dis Primers.* 2021;7(1):52. doi:10.1038/s41572-021-00284-z
2. Thau MR, Bhatraju PK. Sub-phenotypes of acute kidney injury: do we have progress for personalizing care? *Nephron.* 2020;144(12):677–679. doi:10.1159/000511321
3. Coca SG, Garg AX, Thiessen-Philbrook H, et al. Urinary biomarkers of AKI and mortality 3 years after cardiac surgery. *J Am Soc Nephrol.* 2014;25(5):1063–1071. doi:10.1681/ASN.2013070742
4. Koyner JL, Garg AX, Coca SG, et al. Biomarkers predict progression of acute kidney injury after cardiac surgery. *J Am Soc Nephrol.* 2012;23(5):905–914. doi:10.1681/ASN.2011090907
5. Kullmar M, Massoth C, Ostermann M, et al. Biomarker-guided implementation of the KDIGO guidelines to reduce the occurrence of acute kidney injury in patients after cardiac surgery (PrevAKI-multicentre): protocol for a multicentre, observational study followed by randomised controlled feasibility trial. *BMJ Open.* 2020;10(4):e034201. doi:10.1136/bmjopen-2019-034201
6. Ostermann M, Zarbock A, Goldstein S, et al. Recommendations on acute kidney injury biomarkers from the acute disease quality initiative consensus conference: a consensus statement. *JAMA Netw Open.* 2020;3(10):e2019209. doi:10.1001/jamanetworkopen.2020.19209
7. Ikizler TA, Parikh CR, Himmelfarb J, et al. A prospective cohort study of acute kidney injury and kidney outcomes, cardiovascular events, and death. *Kidney Int.* 2021;99(2):456–465. doi:10.1016/j.kint.2020.06.032
8. Siew ED, Abdel-Kader K, Perkins AM, et al. Timing of recovery from moderate to severe AKI and the risk for future loss of kidney function. *Am J Kidney Dis.* 2020;75(2):204–213. doi:10.1053/j.ajkd.2019.05.031
9. Coca SG, Vasquez-Rios G, Mansour SG, et al. Plasma soluble tumor necrosis factor receptor concentrations and clinical events after hospitalization: findings from ASSESS-AKI and ARID studies. *Am J Kidney Dis.* 2023;81(2):190–200. doi:10.1053/j.ajkd.2022.08.007
10. Parikh CR, Coca SG, Thiessen-Philbrook H, et al. Postoperative biomarkers predict acute kidney injury and poor outcomes after adult cardiac surgery. *J Am Soc Nephrol.* 2011;22(9):1748–1757. doi:10.1681/ASN.2010121302
11. Kashani K, Al-Khafaji A, Ardiles T, et al. Discovery and validation of cell cycle arrest biomarkers in human acute kidney injury. *Crit Care.* 2013;17(1):R25. doi:10.1186/cc12503
12. Parikh CR, Jani A, Melnikov VY, Faubel S, Edelstein CL. Urinary interleukin-18 is a marker of human acute tubular necrosis. *Am J Kidney Dis.* 2004;43(3):405–414. doi:10.1053/j.ajkd.2003.10.040
13. Puthumana J, Thiessen-Philbrook H, Xu L, et al. Biomarkers of inflammation and repair in kidney disease progression. *J Clin Invest.* 2021;131(3):e139927. doi:10.1172/jci139927
14. Sabbisetti VS, Waikar SS, Antoine DJ, et al. Blood kidney injury molecule-1 is a biomarker of acute and chronic kidney injury and predicts progression to ESRD in type I diabetes. *J Am Soc Nephrol.* 2014;25(10):2177–2186. doi:10.1681/ASN.2013070758
15. Niewczas MA, Gohda T, Skupien J, et al. Circulating TNF receptors 1 and 2 predict ESRD in type 2 diabetes. *J Am Soc Nephrol.* 2012;23(3):507–515. doi:10.1681/ASN.2011060627
16. Ye X, Luo T, Wang K, et al. Circulating TNF receptors 1 and 2 predict progression of diabetic kidney disease: a meta-analysis. *Diabetes Metab Res Rev.* 2019;35(8):e3195. doi:10.1002/dmrr.3195
17. Gohda T, Niewczas MA, Ficociello LH, et al. Circulating TNF receptors 1 and 2 predict stage 3 CKD in type 1 diabetes. *J Am Soc Nephrol.* 2012;23(3):516–524. doi:10.1681/ASN.2011060628
18. Panduru NM, Sandholm N, Forsblom C, et al. Kidney injury molecule-1 and the loss of kidney function in diabetic nephropathy: a likely causal link in patients with type 1 diabetes. *Diabetes Care.* 2015;38(6):1130–1137. doi:10.2337/dc14-2330
19. Coca SG, Nadkarni GN, Huang Y, et al. Plasma biomarkers and kidney function decline in early and established diabetic kidney disease. *J Am Soc Nephrol.* 2017;28(9):2786–2793. doi:10.1681/ASN.2016101101
20. Schulz CA, Engstrom G, Nilsson J, et al. Plasma kidney injury molecule-1 (p-KIM-1) levels and deterioration of kidney function over 16 years. *Nephrol Dial Transplant.* 2020;35(2):265–273. doi:10.1093/ndt/gfy382
21. Jotwani V, Garimella PS, Katz R, et al. Tubular biomarkers and chronic kidney disease progression in SPRINT participants. *Am J Nephrol.* 2020;51(10):797–805. doi:10.1159/000509978
22. Shou L, Huang WW, Barszczyk A, et al. Blood biomarkers predict cardiac workload using machine learning. *Biomed Res Int.* 2021;2021:6172815. doi:10.1155/2021/6172815
23. Bagshaw SM, Goldstein SL, Ronco C, Kellum JA; ADQI 15 Consensus Group. Acute kidney injury in the era of big data: the 15(th) consensus conference of the acute dialysis quality initiative (ADQI). *Can J Kidney Health Dis.* 2016;3(5). doi:10.1186/s40697-016-0103-z
24. Bhatraju PK, Zelnick LR, Herting J, et al. Identification of acute kidney injury subphenotypes with differing molecular signatures and responses to vasopressin therapy. *Am J Respir Crit Care Med.* 2019;199(7):863–872. doi:10.1164/rccm.201807-1346oc
25. Chaudhary K, Vaid A, Duffy A, et al. Utilization of deep learning for subphenotype identification in sepsis-associated acute kidney injury. *Clin J Am Soc Nephrol.* 2020;15(11):1557–1565. doi:10.2215/CJN.09330819
26. Go AS, Parikh CR, Ikizler TA, et al. The assessment, serial evaluation, and subsequent sequelae of acute kidney injury (ASSESS-AKI) study: design and methods. *BMC Nephrol.* 2010;11(1):22. doi:10.1186/1471-2369-11-22
27. Bhatraju PK, Zelnick LR, Chinchilli VM, et al. Association between early recovery of kidney function after acute kidney injury and long-term clinical outcomes. *JAMA Netw Open.* 2020;3(4):e202682. doi:10.1001/jamanetworkopen.2020.2682
28. Menez S, Moledina DG, Garg AX, et al. Results from the TRIBE-AKI Study found associations between post-operative blood biomarkers and risk of chronic kidney disease after cardiac surgery. *Kidney Int.* 2021;99(3):716–724. doi:10.1016/j.kint.2020.06.037
29. Vasquez-Rios G, Moledina DG, Jia Y, et al. Pre-operative kidney biomarkers and risks for death, cardiovascular and chronic kidney disease events after cardiac surgery: the TRIBE-AKI study. *J Cardiothorac Surg.* 2022;17(1):338. doi:10.1186/s13019-022-02066-4
30. Coca SG, Vasquez-Rios G, Mansour SG, et al. Plasma soluble tumor necrosis factor receptor concentrations and clinical events after hospitalization: findings from the ASSESS-AKI and ARID studies. *Am J Kidney Dis.* 2023;81(2):190–200. doi:10.1053/j.ajkd.2022.08.007
31. Vermeulen AF. Unsupervised learning: deep learning. In: *Industrial Machine Learning.* Apress; 2020. doi:10.1007/978-1-4842-5316-8_8
32. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med.* 2011;30(4):377–399. doi:10.1002/sim.4067
33. Nuvunga JJ, Oliveira LA, Pamplona AK, Silva CP, Lima RR, Balestre M. Factor analysis using mixed models of multi-environment trials with different levels of unbalancing. *Genet Mol Res.* 2015;14(4):14262–14278. doi:10.4238/2015.november.13.10
34. Han L, Shen P, Yan J, et al. Exploring the clinical characteristics of COVID-19 clusters identified using factor analysis of mixed data-based cluster analysis. *Front Med (Lausanne).* 2021;8:644724. doi:10.3389/fmed.2021.644724

35. Khwaja A. KDIGO clinical practice guidelines for acute kidney injury. *Nephron Clin Pract.* 2012;120(4):c179–c184. doi:10.1159/000339789
36. Zhang WR, Parikh CR. Biomarkers of acute and chronic kidney disease. *Annu Rev Physiol.* 2019;81(1):309–333. doi:10.1146/annurev-physiol-020518-114605
37. Zabetian A, Coca SG. Plasma and urine biomarkers in chronic kidney disease: closer to clinical application. *Curr Opin Nephrol Hypertens.* 2021;30(6):531–537. doi:10.1097/MNH.0000000000000735
38. Basile DP, Bonventre JV, Mehta R, et al. Progression after AKI: understanding maladaptive repair processes to predict and identify therapeutic treatments. *J Am Soc Nephrol.* 2016;27(3):687–697. doi:10.1681/ASN.2015030309
39. Koyner JL, Garg AX, Thiessen-Philbrook H, et al. Adjudication of etiology of acute kidney injury: experience from the TRIBE-AKI multi-center study. *BMC Nephrol.* 2014;15(1):105. doi:10.1186/1471-2369-15-105
40. Popescu SG, Whittington A, Gunn RN, et al. Nonlinear biomarker interactions in conversion from mild cognitive impairment to Alzheimer's disease. *Hum Brain Mapp.* 2020;41(15):4406–4418. doi:10.1002/hbm.25133
41. Frohlich H, Balling R, Beerenwinkel N, et al. From hype to reality: data science enabling personalized medicine. *BMC Med.* 2018;16(1):150. doi:10.1186/s12916-018-1122-7
42. Ben-Hamo R, Jacob Berger A, Gavert N, et al. Predicting and affecting response to cancer therapy based on pathway-level biomarkers. *Nat Commun.* 2020;11(1):3296. doi:10.1038/s41467-020-17090-y
43. Kosorok MR, Laber EB. Precision medicine. *Annu Rev Stat Appl.* 2019;6(1):263–286. doi:10.1146/annurev-statistics-030718-105251
44. Prowle JR, Kolic I, Purdell-Lewis J, Taylor R, Pearse RM, Kirwan CJ. Serum creatinine changes associated with critical illness and detection of persistent renal dysfunction after AKI. *Clin J Am Soc Nephrol.* 2014;9(6):1015–1023. doi:10.2215/CJN.1114111
45. Parikh CR, Liu C, Mor MK, et al. Kidney biomarkers of injury and repair as predictors of contrast-associated AKI: a substudy of the PRESERVE trial. *Am J Kidney Dis.* 2020;75(2):187–194. doi:10.1053/j.ajkd.2019.06.011
46. Al-Lamki RS, Mayadas TN. TNF receptors: signaling pathways and contribution to renal dysfunction. *Kidney Int.* 2015;87(2):281–296. doi:10.1038/ki.2014.285
47. Brisco MA, Zile MR, Hanberg JS, et al. Relevance of changes in serum creatinine during a heart failure trial of decongestive strategies: insights from the DOSE trial. *J Card Fail.* 2016;22(10):753–760. doi:10.1016/j.cardfail.2016.06.423
48. Prowle JR. Sepsis-associated AKI. *Clin J Am Soc Nephrol.* 2018;13(2):339–342. doi:10.2215/CJN.07310717
49. Bhatraju PK, Prince DK. *Identification of Molecularly Distinct Sub-Phenotypes in AKI and Association with Long-Term Clinical Outcomes*. American Society of Nephrology Kidney Week; 2021.

Received: September 20, 2022 **Accepted:** March 13, 2023
Published Online Ahead of Print: March 28, 2023

G.V.-R. and W.O. are co-first authors.

AFFILIATIONS

[1]Division of Nephrology, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, New York
[2]Mount Sinai Clinical Intelligence Center, Icahn School of Medicine at Mount Sinai, New York, New York
[3]Division of Data-Driven and Digital Medicine, Icahn School of Medicine at Mount Sinai, New York, New York
[4]Icahn School of Medicine at Mount Sinai, New York, New York
[5]Division of Nephrology, Department of Medicine, Kidney Research Institute, University of Washington, Seattle, Washington
[6]Division of Pulmonary, Critical Care and Sleep Medicine, Department of Medicine, University of Washington, Seattle, Washington
[7]Section of Nephrology, Yale University School of Medicine, New Haven, Connecticut
[8]Division of Nephrology and Hypertension, Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee
[9]Division of Nephrology, Department of Medicine, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada
[10]Department of Biomedical Engineering, SUNY Buffalo, Buffalo, New York
[11]Department of Public Health Sciences, Penn State College of Medicine, Hershey, Pennsylvania
[12]Division of Nephrology, Veterans Affairs New York Harbor Healthcare System and New York University School of Medicine, New York, New York
[13]Division of Nephrology, Department of Medicine, University of California, San Francisco, San Francisco, California
[14]Division of Kidney, Urologic, and Hematologic Diseases, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland
[15]Kaiser Permanente Northern California, Oakland, California
[16]Division of Nephrology, Department of Medicine, Johns Hopkins University, Baltimore, Maryland