CrossMark

SYSTEMS-LEVEL QUALITY IMPROVEMENT

Identification of Major Adverse Kidney Events Within the Electronic Health Record

Matthew W. Semler[1] · Todd W. Rice[1] · Andrew D. Shaw[2] · Edward D. Siew[3] ·
Wesley H. Self[4] · Avinash B. Kumar[2] · Daniel W. Byrne[5] · Jesse M. Ehrenfeld[2,6] ·
Jonathan P. Wanderer[2,6]

Abstract Acute kidney injury is common among critically ill adults and is associated with increased mortality and morbidity. The Major Adverse Kidney Events by 30 days (MAKE30) composite of death, new renal replacement therapy, or persistent renal dysfunction is recommended as a patient-centered outcome for pragmatic trials involving acute kidney injury. Accurate electronic detection of the MAKE30 endpoint using data within the electronic health record (EHR) could facilitate the use of the EHR in large-scale kidney injury research. In an observational study using prospectively collected data from 200 admissions to a single medical intensive care unit, we tested the performance of electronically-extracted data in identifying the MAKE30 composite compared to the reference standard of two-physician manual chart review. The incidence of MAKE30 on manual-review was 16 %, which included 8.5 % for in-hospital mortality, 3.5 % for new renal replacement therapy, and 8.5 % for persistent renal dysfunction. There was strong agreement between the electronic and manual assessment of MAKE30 (98.5 % agreement [95 % CI 96.5–100.0 %]; kappa 0.95 [95 % CI 0.87–1.00]; $P < 0.001$), with only three patients misclassified by electronic assessment. Performance of the electronic MAKE30 assessment was similar among patients with and without CKD and with and without a measured serum creatinine in the 12 months prior to hospital admission. In summary, accurately identifying the MAKE30 composite outcome using EHR data collected as a part of routine care appears feasible.

This article is part of the Topical Collection on *Systems-Level Quality Improvement*

✉ Matthew W. Semler
matthew.w.semler@vanderbilt.edu

[1] Division of Allergy, Pulmonary, and Critical Care Medicine, Vanderbilt University Medical Center, 1161 21st Ave S., T-2220 MCN, Nashville, TN 37232-2650, USA

[2] Department of Anesthesiology, Vanderbilt University Medical Center, Nashville, TN, USA

[3] Division of Nephrology and Hypertension, Vanderbilt Center for Kidney Disease (VCKD) and Integrated Program for AKI (VIP-AKI), Vanderbilt University Medical Center, Nashville, TN, USA

[4] Department of Emergency Medicine, Vanderbilt University Medical Center, Nashville, TN, USA

[5] Department of Biostatistics, Vanderbilt University Medical Center, Nashville, TN, USA

[6] Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, USA

Introduction

Acute kidney injury (AKI) occurs in up to 30 % of critically ill adults [1, 2] and is associated with increased mortality [3, 4] and morbidity [5, 6]. Clinical research evaluating the prevention and treatment of AKI has historically been hampered by the lack of consensus definitions for AKI and the unclear relationship between acute changes in kidney function and longer-term outcomes [7, 8]. To address these challenges, the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) workgroup on Clinical Trials in Acute Kidney Injury recently recommended use of "a composite endpoint of death, provision of dialysis, or sustained loss of kidney function" for phase III trials related to AKI [9]. Analogous to the Major Adverse Cardiovascular Events composite for coronary artery disease [10], the proposed Major

Adverse Kidney Events (MAKE) composite of death, new renal replacement therapy (RRT), or sustained loss of kidney function incorporates several clinically important outcomes and retains a reasonable event rate while shifting the focus from short-term, surrogate measures [8] to longer-term, more patient-centered endpoints [11, 12].

The proliferation of electronic health records (EHRs) and clinical information systems provides a novel opportunity to detect the development of AKI in hospitalized patients [13]. Several previous studies have successfully used EHRs to detect changes in serum creatinine for the purposes of generating provider alerts [14, 15]. Beyond clinical use, there is increasing interest in leveraging tools within the EHR to facilitate the conduct of large, pragmatic trials [16, 17]. In preparation for an upcoming clinical trial, we developed and tested an approach to identifying the MAKE composite endpoint from EHR data collected as part of routine care.

Methods

Study design and oversight

We conducted an observational study using data prospectively collected as a part of an ongoing pilot (NCT02345486). The protocol was approved by the institutional review board at Vanderbilt University with a waiver of informed consent.

Patient population

Among 466 consecutive adult (≥18 years old) admissions to the medical ICU at Vanderbilt University between February 3, 2015 and March 31, 2015, we used computer-generated simple randomization to select 200 cases for review with regard to the MAKE30 outcome (Fig. 1).

Study outcomes

The endpoint of interest was the proportion of patients meeting one or more criteria for Major Adverse Kidney Events within 30 days (MAKE30): in-hospital mortality; receipt of new RRT; or persistent renal dysfunction [1, 9] (Table 1). In-hospital mortality was defined as death from any cause prior to hospital discharge censored at 30 days after ICU admission. Receipt of new RRT was defined as receipt of any modality of RRT between ICU admission and the first of (1) hospital discharge or (2) 30 days in a patient not known to have received RRT prior to ICU admission. Persistent renal dysfunction was defined as a final serum creatinine value before hospital discharge (censored at 30 days after enrollment) that was ≥ 200 % of the baseline creatinine value. Patients who had received RRT prior to enrollment were ineligible for new RRT and persistent renal dysfunction endpoints, but could still meet
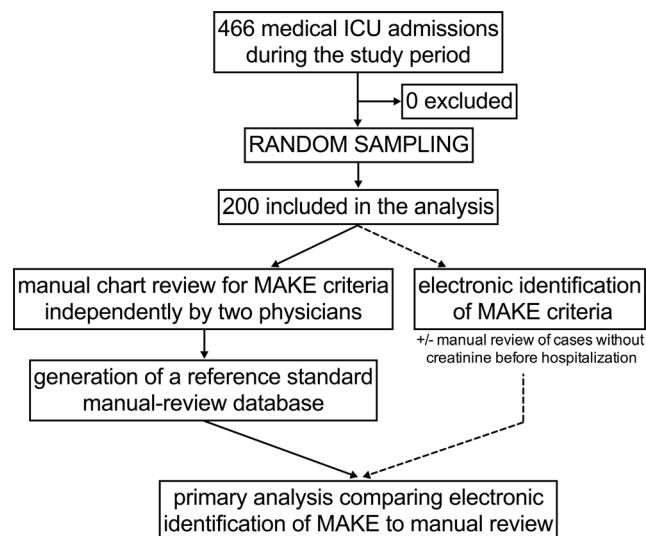

Fig. 1 Flow of patients through the study. From 466 consecutive admissions to the medical intensive care unit (ICU) between February 3, 2015 and March 31, 2015, a sample of 200 cases was selected by computer-generated simple randomization. For these 200 cases, the presence of Major Adverse Kidney Events (MAKE) was determined by (1) two-physician manual chart review and (2) electronic data extraction. Discrepancies between the two physician reviewers were resolved by a third physician to generate a reference standard manual-review dataset. Electronic identification of MAKE (with and without targeted manual review of cases missing a serum creatinine value prior to hospital admission) was compared to MAKE identified by manual review

MAKE30 criteria via the in-hospital mortality component. Secondary outcomes included death, new RRT, and persistent renal dysfunction by 90 days (MAKE90), including outcomes that occurred after hospital discharge.

Study definitions

The value for baseline serum creatinine was determined in a hierarchical approach. The lowest serum creatinine between 12 months and 24 h prior to hospital admission was used when available. If no such creatinine value was available, the lowest

Table 1   Definition of Major Adverse Kidney Events within 30 days (MAKE30)

Major Adverse Kidney Events within 30 days
One or more of the following criteria met in the 30 days after enrollment:
In-hospital mortality
Death prior to hospital discharge
New receipt of RRT
Receipt of any modality of RRT prior to hospital discharge in a patient not known to have previously received RRT.
Persistent renal dysfunction
Final serum creatinine value before hospital discharge ≥ 200 % of the baseline serum creatinine value in a patient not known to have previously received RRT.

RRT is renal replacement therapy

creatinine value between 24 h prior to hospital admission and the time of ICU admission was used. If no creatinine value was available between 12 months prior to hospital admission and the time of ICU admission, a baseline creatinine value was estimated using a previously-described three-variable formula [creatinine = 0.74 − 0.2 (if female) + 0.08 (if African American) + 0.003 × age (in years)] [18].

Chronic kidney disease (CKD) was defined as (1) a highest glomerular filtration rate (GFR) <60 mL/min/1.73 m^2 in the 12 months prior to enrollment as estimated by the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation [19] or (2) a clinical history of CKD stage 3 or greater among patients without an available serum creatinine value.

Data collection by manual chart review

Two physicians independently reviewed the institutional EHR using a structured instrument. Reviewers abstracted the following variables, which had been collected as a part of routine clinical care, into a secure online database:

(1) Demographics, diagnosis, and severity of illness;
(2) Presence of CKD and any prior receipt of RRT;
(3) Serum creatinine values, including lowest between 12 months and 24 h prior to hospital admission, lowest between 24 h prior to hospital admission and ICU admission, highest between ICU admission and hospital discharge or 30 days, and final before hospital discharge or 30 days;
(4) Vital status and receipt of RRT through hospital discharge; and
(5) Post-discharge survival, receipt of RRT, and serum creatinine values when available within the EHR (median duration of follow up among patients surviving to hospital discharge: 183 days; IQR 63–242 days).

After completion of the independent two-physician review, a third physician examined any cases with discrepancy between the initial reviewers with regard to one or more of the MAKE30 criteria. The reason for the discrepancy was recorded and a final reference standard manual-review dataset was generated for comparison to data extracted electronically.

Electronically-extracted data

Structured data from StarPanel, our enterprise EHR, is exported on a daily basis to our institution's Enterprise Data Warehouse (EDW), along with data from our patient registration system, billing system, and laboratory clinical information system. We developed a process to detect transfers and admissions to the study ICU using data extracted from our EDW on a weekly basis. The combination of patient identifiers (medical record number and encounter number) and a

timestamp for study enrollment (date and time of first ICU admission) were used to extract pre- and post-enrollment data elements, as described below.

Patients with a history of prior RRT were identified electronically using the American Medical Association's Current Procedural Terminology (CPT) codes (3066 F, 4054 F, 4055 F, 90963, 90964, 90965, 90966, 90967, 90968, 90969, 90970, 90989, 90993, G0257, G8714, G8956, G9013, G9014, G9231, 90935, 90937, 90945, 90947, 90989, 90993, 90921, 90925, 90999) as well as International Classification of Disease, Clinical Modification (ICD) codes for ICD-9 (39.95, 54.98) and ICD-10 (5A1D00Z, 5A1D60Z, 3E1M39Z) [20]. The presence of any one of these codes in our patient registration system or billing system prior to the date and time of ICU admission resulted in the patient receiving the status of "RRT received prior to enrollment". The same codes were used to determine which patients received RRT during the study period. For all patients who received RRT during the study period, a full text search of the pre-enrollment record was performed using terms related to receipt of RRT to identify patients who had received RRT prior to enrollment at an outside facility (such that CPT and ICD codes for RRT might not be available in our EHR). Search terms included "renal replacement", "RRT", "CRRT", "dialysis", "HD", "PD", "end-stage renal", and "ESRD". Patients who had not received RRT prior to enrollment and received RRT between enrollment and hospital discharge, censored at 30 days, were considered to have met the "new receipt of renal replacement therapy" component of the MAKE30 endpoint.

Using all inpatient, outpatient, and emergency department creatinine values from our institutional laboratory clinical information system, we determined (1) the lowest serum creatinine value between 12 months and 24 h prior to hospital admission, (2) the lowest creatinine value between 24 h prior to hospital admission and the time of ICU admission, and (3) an estimated baseline creatinine value using a previously-described three-variable formula [creatinine = 0.74 − 0.2 (if female) + 0.08 (if African American) + 0.003 × age (in years)] [18]. A baseline creatinine value for each patient was determined using the hierarchical approach described above. For each patient we compared the baseline creatinine value to the final creatinine value obtained between enrollment and hospital discharge, censored at 30 days. If the final creatinine value was at least twice the baseline creatinine value, the "persistent renal dysfunction" component of the MAKE30 outcome was considered present.

Mortality was determined by searching for a mortality-associated discharge disposition in our patient registration system within 30 days of the study enrollment date. Patients with a mortality-associated discharge disposition within 30 days of study enrollment were considered to have met the "mortality" component of the MAKE30 endpoint.

Patients who met any of the three components of the MAKE30 endpoint were considered to have experienced the MAKE30 composite endpoint.

Anticipating potential challenges associated with electronically identifying prior RRT receipt and baseline creatinine values among patients without previous care at the study institution, we tested the additive value of supplementing electronic data abstraction with a "targeted manual review" of the EHR for those cases without an available serum creatinine in the 12 months prior to hospital admission.

Statistical analysis

Because this study focused on comparing two approaches to measuring the same clinical endpoint, no formal power calculation was performed. Continuous variables were reported as mean ± standard deviation or median and interquartile range; categorical variables as frequencies and proportions. Bootstrapping using 1,000 sampling iterations was used to estimate 95 % confidence intervals. Between-group comparisons were made with the Mann–Whitney rank sum test for continuous variables and Chi-square or Fisher exact test for categorical variables. Inter-rater reliability was assessed using the kappa statistic. To provide an impression of how the electronically-extracted MAKE30 criteria would perform if considered a screening test for the presence of the manually-extracted MAKE30 outcome, the sensitivity and specificity were calculated. A two-sided P value < 0.05 was used to determine significance. All analyses were performed using SPSS Statistics v.22 (IBM Corp., Armonk, NY, USA) or R version 3.2.0 (R Foundation for Statistical Computing, Vienna, Austria).

Results

Baseline characteristics

Baseline characteristics of the study cohort ($n = 200$) are given in Table 2. Patients' median age was nearly 60 and almost half were men. One in five had stage 3 or greater CKD, with around 15 % having previously received RRT. Sepsis and respiratory failure were the most common admitting diagnoses, with almost 20 % of patients receiving vasopressors and nearly a third of patients on mechanical ventilation. A total of 148 (74.0 %) patients had a serum creatinine value available from the 12 months prior to hospital admission (median 0.80 mg/dL; IQR 0.62 – 1.25 mg/dL). An additional 43 (21.5 %) patients had a creatinine measurement between hospital admission and enrollment that served as the baseline creatinine value. Only 9 (4.5 %) patients had no creatinine measurements available in the 12 months prior to enrollment and required a calculated estimate of baseline creatinine.

Clinical outcomes determined by manual chart review

The two physician reviewers agreed in their assessment of the presence or absence of MAKE30 in 192 of the 200 cases (96.0 % agreement [95 % CI 93.5–98.5 %]; kappa 0.85 [95 % CI 0.73–0.94]; $P < 0.001$). Disagreement occurred in four cases because a history of RRT was missed by a reviewer and in four cases because a reviewer failed to incorporate an available pre-enrollment value into the baseline serum creatinine.

In the final reference standard manual-review dataset, 32 patients (16.0 %) experienced the MAKE30 composite outcome (Table 2). The incidence of each of the individual MAKE30 components was 8.5 % for in-hospital mortality before 30 days, 3.5 % for receipt of new RRT (2.2 % among survivors), and 8.5 % for persistent renal dysfunction (7.1 % among survivors and 6.1 % among survivors who did not require new renal replacement therapy). The incidence of MAKE by 90 days, including after hospital discharge, was 28.0 %. A total of 27 (16.1 %) patients who did not experience MAKE30 met criteria for MAKE90, 17 of whom died between hospital discharge and 90 days, 1 of whom experienced new RRT between hospital discharge and 90 days, and 9 of whom met criteria by the development of persistent renal dysfunction.

Comparison of electronically- and manually-extracted data

Correlation between electronically- and manually-extracted simple demographic data was perfect ($r^2 = 1.00$; $P < 0.001$ for age, date of hospital admission, date of ICU admission, and body mass index). Electronically- and manually-extracted baseline creatinine values are compared in Fig. 2. *Post-hoc* review of the three cases with a discrepancy greater than 0.25 mg/dL between electronically- and manually-collected values found in all cases that manual review had erroneously classified a creatinine value from shortly after ICU admission as pre-enrollment.

There was strong agreement between the electronic and manual assessment of the MAKE30 endpoint (98.5 % agreement [95 % CI 96.5–100.0 %]; kappa 0.95 [95 % CI 0.87–1.00]; $P < 0.001$) (Table 3). The electronic assessment correctly classified all patients with regard to the receipt of new renal replacement therapy. Two patients who died were misclassified by electronic identification as alive at discharge. Review of these two records revealed a programmatic error in which patients with data retrieved from our Perioperative Data Warehouse who had not experienced any operative procedure received a null value assigned to the death source, even if they died before hospital discharge. Removing this filter resolved the error and correctly classified all 200 patients. Presence of persistent renal dysfunction was correctly classified for 198 of the 200 patients. Two patients with CKD by clinical history but no serum creatinine at the study institution prior to

Table 2 Characteristics of patients with and without a Major Adverse Kidney Event in the 30 days after enrollment

Baseline Characteristics	Overall (n = 200)	MAKE30 absent (n = 168)	MAKE30 present (n = 32)	P value
Age, median [IQR], years	57.0 [44.0–69.0]	57.0 [43.0–68.0]	59.0 [53.0–74.0]	0.12
Men, No. (%)	98 (49.0 %)	75 (44.6 %)	23 (71.9 %)	0.005
Race, No. (%)				0.65
Caucasian	138 (69.0 %)	114 (67.8 %)	24 (75.0 %)	
African American	57 (28.5 %)	50 (29.8 %)	7 (21.9 %)	
Asian	1 (0.5 %)	1 (0.6 %)	0 (0.0 %)	
Other or unknown	4 (2.0 %)	3 (1.8 %)	1 (3.1 %)	
Chronic kidney disease, No. (%)	45 (22.5 %)	33 (19.6 %)	12 (37.5 %)	0.03
Prior renal replacement therapy receipt, No. (%)	28 (14.0 %)	26 (15.5 %)	2 (6.3 %)	0.16
Diagnosis prompting ICU admission, No. (%)				<0.001
Sepsis or septic shock	50 (25.0 %)	37 (22.0 %)	13 (40.6 %)	
COPD exacerbation	17 (8.5 %)	17 (10.1 %)	0 (0.0 %)	
Gastrointestinal bleeding	13 (6.5)	11 (6.5 %)	2 (6.3 %)	
Hypoxic respiratory failure	12 (6.0 %)	9 (5.4 %)	3 (9.4 %)	
Hypercarbic respiratory failure	11 (5.5 %)	11 (6.5 %)	0 (0.0 %)	
Pneumonia	11 (5.5 %)	11 (6.5 %)	0 (0.0 %)	
Diabetic ketoacidosis	10 (5.0 %)	9 (5.4 %)	1 (3.1 %)	
Altered mental status	7 (3.5 %)	7 (4.2 %)	0 (0.0 %)	
CHF or cardiogenic pulmonary edema	5 (2.5 %)	5 (3.0 %)	0 (0.0 %)	
Acute kidney injury	5 (2.5 %)	2 (1.2 %)	3 (9.4 %)	
Overdose or drug toxicity	5 (2.5 %)	5 (3.0 %)	0 (0.0 %)	
Seizure	5 (2.5 %)	3 (1.8 %)	2 (6.3 %)	
Solid tumor	5 (2.5 %)	3 (1.8 %)	2 (6.3 %)	
Cerebrovascular accident	3 (1.5 %)	2 (1.2 %)	1 (3.1 %)	
Hepatic failure	3 (1.5 %)	3 (1.8 %)	0 (0.0 %)	
Hypertensive urgency	3 (1.5 %)	2 (1.2 %)	1 (3.1 %)	
Other	35 (17.5 %)	31 (18.5 %)	4 (12.5 %)	
Body mass index, median [IQR], kg/m^2	26.6 [22.5–32.9]	26.8 [22.4–33.0]	26.4 [23.5–31.8]	0.90
Mechanical ventilation, No. (%)	60 (30.0 %)	43 (25.6 %)	17 (53.1 %)	0.002
Vasopressors, No. (%)	39 (19.5 %)	25 (14.9 %)	14 (43.8 %)	<0.001
UHC expected in-hospital mortality (%), mean ± SD	11.4 ± 18.2	8.1 ± 13.6	28.7 ± 27.6	<0.001
Serum creatinine, median [IQR], mg/dL				
Lowest in 12 months prior to hospitalization	0.80 [0.62–1.25]	0.80 [0.62–1.25]	0.75 [0.62–1.22]	0.53
No. (%) of patients	148 (74.0 %)	123 (73.2 %)	25 (78.1 %)	0.56
Between hospitalization and ICU admission	1.13 [0.77–2.48]	1.07 [0.76–2.16]	1.96 [0.87–4.59]	0.02
No. (%) of patients	173 (86.5 %)	145 (86.3 %)	28 (87.5 %)	0.86
Estimated by three-variable formula	0.82 [0.74–0.92]	0.71 [0.62–0.97]	0.84 [0.80–0.89]	0.77
No. (%) of patients	9 (4.5 %)	7 (4.2 %)	2 (6.3 %)	0.64
Study baseline	0.80 [0.64–1.11]	0.80 [0.65–1.13]	0.66 [0.61–0.93]	0.34
Renal Outcomes				
Before hospital discharge or 30 days				
Highest creatinine, median [IQR], mg/dL	1.23 [0.80–2.87]	1.08 [0.77–2.32]	2.76 [1.59–5.26]	<0.001
Final creatinine, median [IQR], mg/dL	0.91 [0.70–1.69]	0.85 [0.69–1.29]	1.92 [1.23–2.71]	<0.001
Among survivors to hospital discharge	0.89 [0.70–1.55]	0.85 [0.69–1.29]	2.26 [1.66–2.83]	<0.001
Final creatinine >200 % baseline, No. (%)	17 (8.5 %)	0 (0.0 %)	17 (53.1 %)	<0.001
Among survivors to hospital discharge	13/183 (7.1 %)	0/168 (0.0 %)	13/15 (86.7 %)	<0.001
Among survivors to hospital discharge without new RRT	11/179 (6.1 %)	0/168 (0.0 %)	11/11 (100.0 %)	<0.001
Stage II or III KDIGO by creatinine criteria, No. (%)	66 (33.0 %)	42 (25.1 %)	24 (75.0 %)	<0.001

Table 2 (continued)

Baseline Characteristics	Overall (n = 200)	MAKE30 absent (n = 168)	MAKE30 present (n = 32)	P value
Any renal replacement therapy received, No. (%)	28 (14.0 %)	20 (11.9 %)	8 (25.0 %)	0.05
New renal replacement therapy received, No. (%)	7 (3.5 %)	0 (0.0 %)	7 (21.9 %)	<0.001
Among survivors to hospital discharge	4/183 (2.2 %)	0/168 (0.0 %)	4/15 (26.7 %)	<0.001
Before 90 days, including after hospital discharge				
Final creatinine > 200 % baseline, No. (%)	30 (15.0 %)	16 (9.5 %)	14 (43.8 %)	<0.001
Among survivors to 90 days	15/162 (9.3 %)	10/151 (6.6 %)	5/11 (45.5 %)	<0.001
New renal replacement therapy received, No. (%)	8 (4.0 %)	1 (0.6 %)	7 (21.9 %)	<0.001
Clinical Outcomes				
ICU-free days, median [IQR]	25.0 [22.0–26.0]	25.0 [24.0–26.0]	0.0 [0.0–24.0]	<0.001
Ventilator-free days, median [IQR]	28.0 [26.0–28.0]	28.0 [28.0–28.0]	0.0 [0.0–27.0]	<0.001
Mortality, No. (%)				
Before ICU discharge	12 (6.0 %)	0 (0.0 %)	12 (37.5 %)	<0.001
Before hospital discharge or 30 days	17 (8.5 %)	0 (0.0 %)	17 (53.1 %)	<0.001
Before hospital discharge	19 (9.5 %)	1 (0.6 %)	18 (56.3 %)	<0.001
Before 90 days	38 (19.0 %)	17 (10.1 %)	21 (65.6 %)	<0.001
At any point during follow up	50 (25.0 %)	28 (16.7 %)	22 (68.8 %)	<0.001

Major Adverse Kidney Events within 30 days (MAKE30) is the presence of any of the following before discharge from the hospital or 30 days after enrollment: death, receipt of new renal replacement therapy, or a final serum creatinine value ≥ 200 % of baseline. Chronic kidney disease is defined as an estimated glomerular filtration rate (GFR) <60 mL/min/1.73 m^2 in the 12 months prior to enrollment or a clinical history of stage 3 or greater CKD in a patient without a creatinine value available in the 12 months prior to enrollment. University HealthSystem Consortium (UHC) expected mortality is an estimated probability of death before hospital discharge generated for each patient based on age, gender, comorbidities, admission source, race, and principal diagnosis (details at www.uhc.edu). Definitions for baseline serum creatinine are: "Lowest in 12 months prior to hospitalization" = lowest available serum creatinine between 12 months and 24 h prior to hospital admission; "Between hospitalization and ICU admission" = lowest available serum creatinine between 24 h prior to hospital admission and ICU admission; "Estimated by three-variable formula" = creatinine value calculated using a previously-described three-variable formula [creatinine (mg/dL) = 0.74 − 0.2 (if female) + 0.08 (if African American) + 0.003 × age (in years)] [18]; and "Study baseline" = lowest in 12 months prior to hospitalization if available, otherwise between hospitalization and ICU admission, using the estimated creatinine only for patients without an available creatinine in 12 months prior to enrollment. Acute kidney injury is defined according to Kidney Disease Improving Global Outcomes (KDIGO) stage II or III creatinine criteria (serum creatinine at least 2.0× baseline, increase in serum creatinine by at least 0.3 mg/dL to at least 4.0 mg/dL, or initiation of renal replacement therapy) [8]. IQR is interquartile range; CHF is congestive heart failure; ICU is intensive care unit

enrollment were inappropriately classified by the electronic assessment as experiencing new persistent renal dysfunction. Supplementing the electronic MAKE30 assessment with targeted manual review of cases without an available serum creatinine value prior to hospital admission achieved appropriate classification of these two cases. Electronic MAKE30 assessments supplemented with targeted manual review performed similarly among patients with and without CKD and among those with and without a prior serum creatinine value in the EHR (Table 3). The final electronic algorithm, supplemented by targeted manual review of cases without a pre-admission creatinine, achieved 100 % sensitivity and specificity for the MAKE30 endpoint in the current dataset.

Discussion

Establishing reliable methods for electronically collecting patient-centered outcomes is essential to leveraging the EHR

for use in pragmatic trials [16, 17, 21]. This prospective, observational study demonstrated the feasibility of electronically identifying death, new RRT receipt, or persistent renal dysfunction among critically ill adults, using EHR data collected as a part of routine care.

Although the development of stage II or III AKI by KDIGO criteria [8] currently represents the most established definition of AKI in clinical research, there is increasing recognition of the need to examine outcomes meaningful both to clinicians and patients [9]. The MAKE composite endpoint captures, at a consistent time interval, mortality (the most important outcome to many patients) as well as receipt of RRT and persistent renal dysfunction (two kidney-specific events which may be more closely associated with long-term morbidity and quality-of-life than transient changes in creatinine [5, 6]). Uncertainty remains regarding how to best define the persistent renal dysfunction component of MAKE. A sustained doubling in creatinine at 30 days represents a large reduction in GFR and may prioritize specificity at the cost of
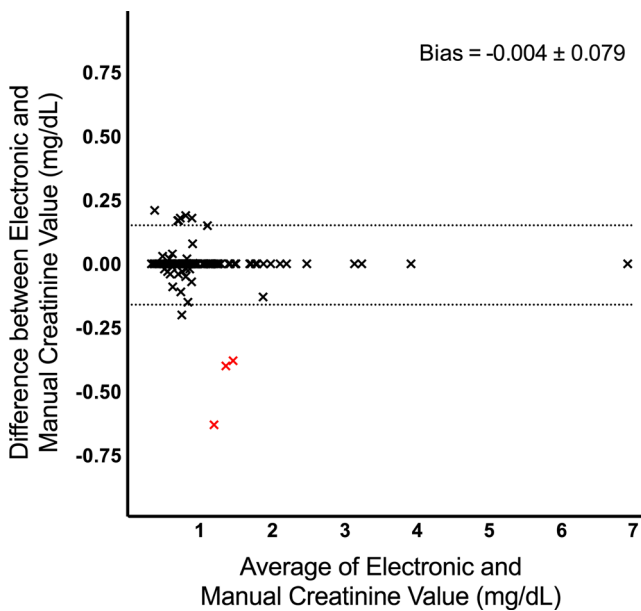
Fig. 2 Bland-Altman plot of electronically- versus manually-extracted baseline creatinine values. Among the 200 patients in the current study, 172 had never received renal replacement therapy prior to ICU admission and were eligible to experience the creatinine-based component of the MAKE30 outcome. For these 172 patients, the difference between (Y axis) and average of (X axis) electronically- and manually-extracted baseline serum creatinine values (mg/dL) are displayed. Each point represents an individual patient and dotted lines are the 95 % limits of agreement. The three cases with a discrepancy greater than 0.25 mg/dL between electronically- and manually-collected values (*red*) were found to be due to errors in the manually-collected creatinine values

decreased sensitivity. Even as work is ongoing to determine the best creatinine or estimated GFR criteria for persistent renal dysfunction [12], MAKE is increasingly being recommended [9] and used [12] as the endpoint of choice for AKI clinical trials and biomarker validation studies [1, 9, 11, 12, 22–26].

A number of prior studies have assessed the feasibility of detecting AKI with data from the EHR [13, 27], primarily using laboratory clinical information systems to identify changes in serum creatinine concentration. Although the reference standard for AKI and the performance of electronic detection have varied significantly across studies [13–15, 27–29], some methods have achieved sensitivity and specificity in excess of 90 %. Notably, many of these studies were limited to a narrow spectrum of patients, frequently excluding those with CKD.

Numerous prior studies have also evaluated the use of administrative data to identify episodes of AKI in hospitalized patients [21, 30, 31]. Most AKI studies using administrative data have applied ICD-9 codes or CPT codes to capture the diagnosis of AKI and RRT. Commonly used administrative codes for AKI (e.g. ICD-9-CM 584, ICD-10 N-17) generally demonstrate low sensitivity and higher specificity [21]. Inclusion of only patients with AKI requiring dialysis may improve diagnostic performance [30], but sensitivity in some

studies has remained as low as 40 % [31]. Use of billings codes may identify AKI with a more severe phenotype and may demonstrate better performance characteristics at higher stages of AKI.

The goal and technical approach of the current study were significantly different than these prior, related studies. We aimed to identify critically ill adults who experienced a Major Adverse Kidney Event between ICU admission and hospital discharge, using all data available within the hospital informatics systems. By merging date- and time-specific laboratory data from the inpatient and outpatient setting with administrative data on ICD-9 and CPT codes, we were able to accurately identify patients who experienced the MAKE30 composite outcome. The sensitivity and specificity of electronically extracted data for the manually-collected MAKE30 outcome were above 95 %, and increased incrementally with targeted manual review of charts known to be at higher risk for misclassification due to missing baseline creatinine data. Performance was similar among patients with and without evidence of CKD. Identifying the MAKE endpoint using all available laboratory and administrative data from an individual patient's hospitalization avoids some of the challenges associated with detecting AKI via laboratory values or administrative datasets alone. A doubling of creatinine from baseline to discharge is easier to detect than smaller, time-dependent changes in creatinine. Death and RRT may be coded more consistently in administrative data than AKI diagnoses generally. The ability to reliably detect the MAKE composite endpoint from EHR data collected during routine care suggest the MAKE endpoint is well suited for use in pragmatic AKI research.

Our study has several strengths. Manual chart review by two physicians is a well-recognized reference standard for kidney injury outcomes, and the 91 % agreement between the two initial reviewers in our study was similar to that observed in previous AKI studies [28, 31]. Including all ICU admissions allowed examination of performance characteristics over a wide spectrum of underlying illness, including patients with CKD (for whom risk of AKI is high but assessment of baseline creatinine may be challenging) and patients receiving RRT prior to admission (who remain at risk for the in-hospital mortality component of MAKE). Correction of the small number of systematic errors in electronic data extraction identified in the current study may produce even more accurate MAKE30 identification in future studies.

Our study also has limitations. We studied only 200 patients in a single ICU at a single center. Our approach to identifying MAKE in the EHR might perform differently in other populations of patients and providers or at centers that handle laboratory and administrative data differently. Replication might prove challenging in healthcare environments in which information systems differ between the outpatient and inpatient setting, or where patients are less

Table 3   Major Adverse Kidney Events by 30 days derived from electronically-extracted compared with manually-extracted data (referent)

Outcome	Kappa (95 % CI)	Concordant	Discordant	Sensitivity	Specificity
Overall cohort, electronically-extracted data (*n* = 200)					
MAKE30	0.95 (0.87–1.00)	197 (98.5 %)	3 (1.5 %)	96.9 %	98.8 %
MAKE30 In-hospital death	0.93 (0.82–1.00)	198 (99.0 %)	2 (1.0 %)	88.2 %	100.0 %
MAKE30 new RRT	1.00 (1.00–1.00)	200 (100.0 %)	0 (0.0 %)	100.0 %	100.0 %
MAKE30 Cr > 2× baseline	0.94 (0.84–1.00)	198 (99.0 %)	2 (1.0 %)	100.0 %	98.9 %
Overall cohort, electronically-extracted data with targeted manual review (*n* = 200)					
MAKE30	0.98 (0.94–1.00)	199 (99.5 %)	1 (0.5 %)	96.9 %	100.0 %
MAKE30 In-hospital death	0.93 (0.82–1.00)	198 (99.0 %)	2 (1.0 %)	88.2 %	100.0 %
MAKE30 new RRT	1.00 (1.00–1.00)	200 (100.0 %)	0 (0.0 %)	100.0 %	100.0 %
MAKE30 Cr > 2× baseline	1.00 (1.00–1.00)	200 (100.0 %)	0 (0.0 %)	100.0 %	100.0 %
Patients with an available serum creatinine in the 12 months prior to hospitalization (*n* = 148)					
MAKE30	0.98 (0.92–1.00)	147 (99.3 %)	1 (0.7 %)	96.0 %	100.0 %
Patients without an available serum creatinine in the 12 months to hospitalization (*n* = 52)					
MAKE30	1.00 (1.00–1.00)	52 (100.0 %)	0 (0.0 %)	100.0 %	100.0 %
Patients without chronic kidney disease prior to enrollment (*n* = 155)					
MAKE30	0.97 (0.90–1.00)	154 (99.4 %)	1 (0.6 %)	95.0 %	100.0 %
Patients with chronic kidney disease prior to enrollment (*n* = 45)					
MAKE30	1.00 (1.00–1.00)	45 (100.0 %)	0 (0.0 %)	100.0 %	100.0 %

Using two-physician manual chart review as the reference standard, the performance of electronically-extracted data to identify the presence of Major Adverse Kidney Events within 30 days (MAKE30) was assessed. The inter-rater reliability between electronically- and manually-extracted data for MAKE30 is given as a kappa statistic. The number and proportion of cases classified correctly (concordant) and incorrectly (discordant), are given along with the sensitivity and specificity of electronically-extracted data for identifying patients with manually-identified MAKE3. The incremental benefit of supplementing electronically-extracted data with a targeted manual review of cases without a baseline creatinine available was also assessed. Performance of the electronically-extracted MAKE30 plus targeted manual review was compared in subgroups with and without a creatinine value prior to hospitalization and with and without pre-enrollment chronic kidney disease. RRT is renal replacement therapy

consistently cared for at a single institution. Determining baseline creatinine values in studies of AKI is a recognized challenge [32] and alternative definitions might have produced different results. Censoring the primary MAKE assessment at hospital discharge or 30 days avoids biases related to differences in post-discharge follow up, but offers less information about progression to CKD than complete patient follow up to a later time-point [12]. Data from electronic medical records contain an inherent rate of noise compared to data deliberately collected by study personnel as a part of research. Reassuringly, however, the concordance between the electronically-assessed and reference standard MAKE30 outcome appeared to be similar to the concordance between the two physician reviewers for the same endpoint.

Conclusions

Accurately identifying critically ill adults who experience a Major Adverse Kidney Event using EHR data collected during routine care is feasible. Future research is needed to test the performance of the methods described here in other settings.

References

1. Kashani, K., Al-Khafaji, A., Ardiles, T., Artigas, A., Bagshaw, S. M., Bell, M., Bihorac, A., Birkhahn, R., Cely, C. M., Chawla, L. S., Davison, D. L., Feldkamp, T., Forni, L. G., Gong, M. N., Gunnerson, K. J., Haase, M., Hackett, J., Honore, P. M., Hoste, E. A. J., Joannes-Boyau, O., Joannidis, M., Kim, P., Koyner, J. L., Laskowitz, D. T., Lissauer, M. E., Marx, G., McCullough, P. A., Mullaney, S., Ostermann, M., Rimmelé, T., Shapiro, N. I., Shaw, A. D., Shi, J., Sprague, A. M., Vincent, J.-L., Vinsonneau, C., Wagner, L., Walker, M. G., Wilkerson, R. G., Zacharowski, K., and Kellum, J. A., Discovery and validation of cell cycle arrest biomarkers in human acute kidney injury. *Crit. Care Lond. Engl.* 17(1):R25, 2013. doi:10.1186/cc12503.

2. Bellomo, R., Kellum, J. A., and Ronco, C., Acute kidney injury. *Lancet Lond. Engl.* 380(9843):756–766, 2012. doi:10.1016/S0140-6736(11)61454-2.

3. Metnitz, P. G. H., Krenn, C. G., Steltzer, H., Lang, T., Ploder, J., Lenz, K., Le Gall, J.-R., and Druml, W., Effect of acute renal failure requiring renal replacement therapy on outcome in critically ill patients. *Crit. Care Med.* 30(9):2051–2058, 2002. doi:10.1097/01.CCM.0000026732.62103.DF.

4. Uchino, S., Kellum, J. A., Bellomo, R., Doig, G. S., Morimatsu, H., Morgera, S., Schetz, M., Tan, I., Bouman, C., Macedo, E., Gibney, N., Tolwani, A., Ronco, C., and Beginning and Ending Supportive Therapy for the Kidney (BEST Kidney) Investigators, Acute renal failure in critically ill patients: A multinational, multicenter study. *JAMA* 294(7):813–818, 2005. doi:10.1001/jama.294.7.813.

5. Bagshaw, S. M., Laupland, K. B., Doig, C. J., Mortis, G., Fick, G. H., Mucenski, M., Godinez-Luna, T., Svenson, L. W., and Rosenal, T., Prognosis for long-term survival and renal recovery in critically ill patients with severe acute renal failure: A population-based study. *Crit. Care Lond. Engl.* 9(6):R700–R709, 2005. doi:10.1186/cc3879.

6. Coca, S. G., Singanamala, S., and Parikh, C. R., Chronic kidney disease after acute kidney injury: A systematic review and meta-analysis. *Kidney Int.* 81(5):442–448, 2012. doi:10.1038/ki.2011.379.

7. Kellum, J. A., Levin, N., Bouman, C., and Lameire, N., Developing a consensus classification system for acute renal failure. *Curr. Opin. Crit. Care* 8(6):509–514, 2002.

8. Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group, KDIGO clinical practice guideline for acute kidney injury. *Kidney Int.* 2(Suppl):1–138, 2012.

9. Palevsky, P. M., Molitoris, B. A., Okusa, M. D., Levin, A., Waikar, S. S., Wald, R., Chertow, G. M., Murray, P. T., Parikh, C. R., Shaw, A. D., Go, A. S., Faubel, S. G., Kellum, J. A., Chinchilli, V. M., Liu, K. D., Cheung, A. K., Weisbord, S. D., Chawla, L. S., Kaufman, J.

S., Devarajan, P., Toto, R. M., Hsu, C., Greene, T., Mehta, R. L., Stokes, J. B., Thompson, A. M., Thompson, B. T., Westenfelder, C. S., Tumlin, J. A., Warnock, D. G., Shah, S. V., Xie, Y., Duggan, E. G., Kimmel, P. L., and Star, R. A., Design of clinical trials in acute kidney injury: Report from an NIDDK workshop on trial methodology. *Clin. J. Am. Soc. Nephrol. CJASN* 7(5):844–850, 2012. doi:10.2215/CJN.12791211.

10. Stone, G. W., Maehara, A., Lansky, A. J., de Bruyne, B., Cristea, E., Mintz, G. S., Mehran, R., McPherson, J., Farhat, N., Marso, S. P., Parise, H., Templin, B., White, R., Zhang, Z., Serruys, P. W., and PROSPECT Investigators, A prospective natural-history study of coronary atherosclerosis. *N. Engl. J. Med.* 364(3):226–235, 2011. doi:10.1056/NEJMoa1002358.

11. Shaw, A., Models of preventable disease: Contrast-induced nephropathy and cardiac surgery-associated acute kidney injury. *Contrib. Nephrol.* 174:156–162, 2011. doi:10.1159/000329387.

12. Weisbord, S. D., Gallagher, M., Kaufman, J., Cass, A., Parikh, C. R., Chertow, G. M., Shunk, K. A., McCullough, P. A., Fine, M. J., Mor, M. K., Lew, R. A., Huang, G. D., Conner, T. A., Brophy, M. T., Lee, J., Soliva, S., and Palevsky, P. M., Prevention of contrast-induced AKI: A review of published trials and the design of the prevention of serious adverse events following angiography (PRESERVE) trial. *Clin. J. Am. Soc. Nephrol. CJASN* 8(9):1618–1631, 2013. doi:10.2215/CJN.11161012.

13. James, M. T., Hobson, C. E., Darmon, M., Mohan, S., Hudson, D., Goldstein, S. L., Ronco, C., Kellum, J. A., Bagshaw, S. M., and Acute Dialysis Quality Initiative (ADQI) Consensus Group, Applications for detection of acute kidney injury using electronic medical records and clinical information systems: Workgroup statements from the 15(th) ADQI Consensus Conference. *Can. J. Kidney Health Dis.* 3:9, 2016. doi:10.1186/s40697-016-0100-2.

14. Colpaert, K., Hoste, E. A., Steurbaut, K., Benoit, D., Van Hoecke, S., De Turck, F., and Decruyenaere, J., Impact of real-time electronic alerting of acute kidney injury on therapeutic intervention and progression of RIFLE class. *Crit. Care Med.* 40(4):1164–1170, 2012. doi:10.1097/CCM.0b013e3182387a6b.

15. Ahmed, A., Vairavan, S., Akhoundi, A., Wilson, G., Chiofolo, C., Chbat, N., Cartin-Ceba, R., Li, G., and Kashani, K., Development and validation of electronic surveillance tool for acute kidney injury: A retrospective analysis. *J. Crit. Care* 30(5):988–993, 2015. doi:10.1016/j.jcrc.2015.05.007.

16. Angus, D. C., Fusing randomized trials with big data: The key to self-learning health care systems? *JAMA* 314(8):767–768, 2015. doi:10.1001/jama.2015.7762.

17. Semler, M. W., Rice, T. W., and Ehrenfeld, J. M., Leveraging clinical informatics in the conduct of clinical trials. *J. Med. Syst.* 39(10):112, 2015. doi:10.1007/s10916-015-0317-0.

18. Zavada, J., Hoste, E., Cartin-Ceba, R., Calzavacca, P., Gajic, O., Clermont, G., Bellomo, R., Kellum, J. A., and for the AKI6 investigators, A comparison of three methods to estimate baseline creatinine for RIFLE classification. *Nephrol. Dial. Transplant.* 25(12):3911–3918, 2010. doi:10.1093/ndt/gfp766.

19. Levey, A. S., Stevens, L. A., Schmid, C. H., Zhang, Y. L., Castro, A. F., Feldman, H. I., Kusek, J. W., Eggers, P., Van Lente, F., Greene, T., Coresh, J., and CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration), A new equation to estimate glomerular filtration rate. *Ann. Intern. Med.* 150(9):604–612, 2009.

20. Aline, F., Wanderer, J. P., and Ehrenfeld, J. M., Managing the impact of the ICD-10 transition on a data warehouse. *J. Med. Syst.* 40(3):57, 2016. doi:10.1007/s10916-015-0418-9.

21. Siew, E. D., Basu, R. K., Wunsch, H., Shaw, A. D., Goldstein, S. L., Ronco, C., Kellum, J. A., Bagshaw, S. M., and 15th ADQI Consensus Group, Optimizing administrative datasets to examine acute kidney injury in the era of big data: Workgroup statement from the 15(th) ADQI Consensus Conference. *Can. J. Kidney Health Dis.* 3:12, 2016. doi:10.1186/s40697-016-0098-5.

22. Haase, M., Bellomo, R., Albert, C., Vanpoucke, G., Thomas, G., Laroy, W., Verleysen, K., Kropf, S., Kuppe, H., Hetzer, R., and Haase-Fielitz, A., The identification of three novel biomarkers of major adverse kidney events. *Biomark. Med.* 8(10):1207–1217, 2014. doi:10.2217/bmm.14.90.

23. Mehta, R., Bihorac, A., Selby, N. M., Quan, H., Goldstein, S. L., Kellum, J. A., Ronco, C., Bagshaw, S. M., and Acute Dialysis Quality Initiative (ADQI) Consensus Group, Establishing a continuum of acute kidney injury - tracing AKI using data source linkage and long-term follow-up: Workgroup Statements from the 15th ADQI Consensus Conference. *Can. J. Kidney Health Dis.* 3:13, 2016. doi:10.1186/s40697-016-0102-0.

24. Bivet, F. G., Nonuse of RIFLE classification urine output criteria: Biases for acute kidney injury biomarkers performance assessment? *Crit. Care Med.* 40(5):1692, 2012. doi:10.1097/CCM.0b013e318246b72a. author reply 1692-1693.

25. Dewitte, A., Joannès-Boyau, O., Sidobre, C., Fleureau, C., Bats, M.-L., Derache, P., Leuillet, S., Ripoche, J., Combe, C., and Ouattara, A., Kinetic eGFR and novel AKI biomarkers to predict renal recovery. *Clin. J. Am. Soc. Nephrol. CJASN* 10(11):1900–1910, 2015. doi:10.2215/CJN.12651214.

26. Chawla, L. S., Amdur, R. L., Shaw, A. D., Faselis, C., Palant, C. E., and Kimmel, P. L., Association between AKI and long-term renal and cardiovascular outcomes in United States veterans. *Clin. J. Am. Soc. Nephrol. CJASN* 9(3):448–456, 2014. doi:10.2215/CJN.02440213.

27. Kashani, K., and Herasevich, V., Sniffing out acute kidney injury in the ICU: Do we have the tools? *Curr. Opin. Crit. Care* 19(6):531–536, 2013. doi:10.1097/MCC.0000000000000024.

28. Flynn, N., and Dawnay, A., A simple electronic alert for acute kidney injury. *Ann. Clin. Biochem.* 52(Pt 2):206–212, 2015. doi:10.1177/0004563214534832.

29. Wilson, F. P., Shashaty, M., Testani, J., Aqeel, I., Borovskiy, Y., Ellenberg, S. S., Feldman, H. I., Fernandez, H., Gitelman, Y., Lin, J., Negoianu, D., Parikh, C. R., Reese, P. P., Urbani, R., and Fuchs, B., Automated, electronic alerts for acute kidney injury: A single-blind, parallel-group, randomised controlled trial. *Lancet Lond. Engl.* 385(9981):1966–1974, 2015. doi:10.1016/S0140-6736(15)60266-5.

30. Vlasschaert, M. E. O., Bejaimal, S. A. D., Hackam, D. G., Quinn, R., Cuerden, M. S., Oliver, M. J., Iansavichus, A., Sultan, N., Mills, A., and Garg, A. X., Validity of administrative database coding for kidney disease: A systematic review. *Am. J. Kidney Dis. Off. J. Natl. Kidney Found.* 57(1):29–43, 2011. doi:10.1053/j.ajkd.2010.08.031.

31. Grams, M. E., Waikar, S. S., MacMahon, B., Whelton, S., Ballew, S. H., and Coresh, J., Performance and limitations of administrative data in the identification of AKI. *Clin. J. Am. Soc. Nephrol. CJASN* 9(4):682–689, 2014. doi:10.2215/CJN.07650713.

32. Siew, E. D., Ikizler, T. A., Matheny, M. E., Shi, Y., Schildcrout, J. S., Danciu, I., Dwyer, J. P., Srichai, M., Hung, A. M., Smith, J. P., and Peterson, J. F., Estimating baseline kidney function in hospitalized patients with impaired kidney function. *Clin. J. Am. Soc. Nephrol. CJASN* 7(5):712–719, 2012. doi:10.2215/CJN.10821011.