ORIGINAL ARTICLE

Development and internal validation of a prediction model for hospital-acquired acute kidney injury

Catalina Martin-Cleary[1,2,3,*], Luis Miguel Molinero-Casares[4,*],
Alberto Ortiz[1,2,3,**] and Jose Miguel Arce-Obieta[5,**]

[1]Department of Nephrology and Hypertension, Investigación Sanitaria-Fundación Jimenez Diaz, Universidad Autónoma de Madrid, Madrid, Spain, [2]REDINREN, Madrid, Spain, [3]Fundación Renal Iñigo Alvarez de Toledo-IRSIN, Madrid, Spain, [4]Alce Ingeniería, Madrid, Spain and [5]Department of Health Information Management, University Hospital Fundación Jiménez Díaz, Madrid, Spain

Correspondence to: Catalina Martin-Cleary; E-mail: CMartinC@fjd.es
*The first two authors contributed equally to this work.
**The last two authors contributed equally to this work.

ABSTRACT

Background. Predictive models and clinical risk scores for hospital-acquired acute kidney injury (AKI) are mainly focused on critical and surgical patients. We have used the electronic clinical records from a tertiary care general hospital to develop a risk score for new-onset AKI in general inpatients that can be estimated automatically from clinical records.

Methods. A total of 47 466 patients met inclusion criteria within a 2-year period. Of these, 2385 (5.0%) developed hospital-acquired AKI. Step-wise regression modelling and Bayesian model averaging were used to develop the Madrid Acute Kidney Injury Prediction Score (MAKIPS), which contains 23 variables, all obtainable automatically from electronic clinical records at admission. Bootstrap resampling was employed for internal validation. To optimize calibration, a penalized logistic regression model was estimated by the least absolute shrinkage and selection operator (lasso) method of coefficient shrinkage after estimation.

Results. The area under the curve of the receiver operating characteristic curve of the MAKIPS score to predict hospital-acquired AKI at admission was 0.811. Among individual variables, the highest odds ratios, all >2.5, for hospital-acquired AKI were conferred by abdominal, cardiovascular or urological surgery followed by congestive heart failure. An online tool (http://www.bioestadistica.net/MAKIPS.aspx) will facilitate validation in other hospital environments.

Conclusions. MAKIPS is a new risk score to predict the risk of hospital-acquired AKI, based on variables present at admission in the electronic clinical records. This may help to identify patients who require specific monitoring because of a high risk of AKI.

Keywords: acute kidney injury, cardiovascular surgery, digestive surgery, heart failure, hospital-acquired, prediction, risk score, surgery

INTRODUCTION

Predictive models and clinical risk scores for hospital-acquired acute kidney injury (AKI) are mainly focused on critical and surgical patients, which are high-risk groups for severe AKI requiring dialysis. A PubMed search with the terms 'clinical predictive models' and 'acute kidney injury' from 2017 to the present date yielded 41 publications of clinical models for AKI in the intensive care unit (ICU) and cardiac surgery settings, and 6 in nonsurgical populations, mainly for prediction of contrast nephropathy. Prediction in the critical care setting emphasizes moderate and severe AKI. However, there is a need for clinical predictive models that predict AKI in general hospital-acquired settings and for milder forms of AKI [1–4]. Evidence of increased in-hospital and long-term mortality and risk of progression to chronic kidney disease (CKD) even after mild forms of AKI, recognition of the epidemiology of community-acquired AKI (CA-AKI) and AKI outside ICU settings, and the identification of systematic deficits in the diagnosis and management of AKI are some of the premises to create valid tools to identify at-risk patients for hospital-acquired AKI [5–9]. As electronic medical records and big data become more accessible, reliable and user-friendly clinical prediction models may become a feasible option for AKI prediction. We have now employed a large database of clinical and analytical electronic records to develop a tool that uses these records to predict the risk of hospital-acquired AKI at admission.

MATERIALS AND METHODS

We followed the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement for reporting multivariable prediction model development and validation (Supplementary data, Table S1). The hospital is a tertiary care referral hospital, affiliated with the Universidad Autónoma de Madrid. In Spain, access to primary and specialized care and hospitalization is free at the point-of-care. Primary and specialized cares are integrated and allow access to each others' clinical records.

We used the Fundación Jiménez Díaz Hospital electronic medical records of hospitalized patients who had been discharged from 1 January 2015 to 31 December 2016. Patient comorbidities, diagnosis and procedures were classified according to the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM). Blood and urine analysis data from inpatient and outpatient settings were available for cohort patients for up to 730 days prior to the hospitalization date. The study complied with the Declaration of Helsinki and Spanish law, and was approved by the Investigación Sanitaria-Fundación Jiménez Díaz ethics Committee, which waived the need for informed consent, given the nature of the study.

Study population

We identified all patients ≥18 years of age who had been discharged during the study period. We excluded patients on chronic dialysis, admitted for a renal transplant or with hospital stay <24 h. Patients who had AKI within the first 48 h of hospital admission were excluded from the model estimation as they were considered to have CA-AKI. Exceptions to this rule were patients admitted for elective surgery in whom the blood sample diagnostic of AKI was obtained post-surgery.

Baseline kidney function

Baseline kidney function was defined as the most recent serum creatinine between 1 and 365 days prior to the hospitalization date. The Modification of Diet in Renal Disease (MDRD-4) equation [10, 11] was used to estimate the glomerular filtration rate. If there was no serum creatinine within 365 days prior to hospitalization, the baseline was the lowest serum creatinine during hospitalization. Serum creatinines obtained during renal replacement therapy were excluded. There were 9116 admissions that lacked baseline serum creatinine (16.8%).

Definition of AKI

Following Kidney Disease: Improving Global Outcomes (KDIGO), hospital-acquired AKI was defined as an increase in serum creatinine during hospitalization of ≥0.3 mg/dL or >50% over the baseline that occurred after the first 48 h of hospital admission [12] or as the requirement of renal replacement therapy. Of note, the KDIGO definition may be used in general wards and differs from the prior RIFLE (Risk, Injury, Failure, Loss and End-stage renal disease) criteria, in that there is no requirement for the increase in serum creatinine to be sustained. This is different from RIFLE criteria, developed for ICUs, in which daily availability of labs is routine and in which the increase in serum creatinine should be sustained (>24 h). Severity was categorized according to KDIGO. For elective surgery, if AKI was present within 24 h of admission and sampling diagnostic of AKI was post-surgery, the patient was considered to have hospital-acquired AKI and included in the study, since these patients usually had a baseline a few days before admission, and were otherwise stable until the surgical procedure.

Study outcome

The outcome was development of hospital-acquired AKI.

Statistical analysis

R software version 3.3.1 was used.

Candidate predictor variables. We identified demographic, comorbid and laboratory candidate predictor variables for inclusion in our model based on a review of the literature and on availability in the electronic clinical records [7, 13–24]. Comorbidities that compose the Charlson Index [25, 26] present at admission were included individually, to identify those with more weight as predictor variables. For laboratory variables, we selected the values closest to the hospitalization date from the first 24 h of the admission up to 730 days before admission. Admission type and nature of surgery were also included as potential predictor variables (Supplementary data, Table S2). Code diagnoses representing each comorbidity are presented in Supplementary data, Table S3. For continuous variables with a possible non-linear relationship with the logit response, a restricted cubic spline with three knots was evaluated. Finally, only one quadratic equation was employed in the final model (potassium). Of note, the laboratory value estimated glomerular filtration rate (eGFR) was not included in the final multivariable analysis, as we were concerned that including eGFR would overestimate the model. Instead, we included the ICD-9 diagnosis of renal disease as a broader way of including CKD as a risk factor for AKI. The renal disease diagnosis was more frequent in patients with AKI {odds ratios (ORs) 1.94 [95% confidence interval (CI) 1.66–2.27]}.

Multivariable discovery. We performed backward step-wise regression modelling using Akaike information criterion and then applied Bayesian model averaging (BMA) to optimize model performance by reducing the number of variables [27]. For internal validation, bootstrap resampling was used [28, 29]. To estimate the model, numerical variables were escalated by subtracting the median and dividing by median absolute deviation. This is reflected in the coefficients and makes them comparable.

The first multivariate model to predict hospital-acquired AKI was developed using all available variables. The best possible predictive model with the lowest number of variables possible was obtained through an heuristic procedure. The model including all variables [Supplementary data, Table S4, area under the curve (AUC) = 0.81] was refined by step-wise selection. To further reduce the number of variables, the BMA selection algorithm was applied, yielding a final model with 23 variables, 7 of them laboratory values, which had the good predictive ability [estimated AUC: 0.811; AUC after bootstrap resampling ($n = 1000$ samples): 0.810]. No interactions between variables were considered in modelling. The variance inflation factor did not show evidence of collinearity between variables in the selected model [30–32].

Model calibration. To optimize the calibration of the model, a penalized logistic regression model was estimated by the least absolute shrinkage and selection operator (lasso) method for the variables in the final model. Lasso shrinks data values towards a central point or mean and adds a penalty to the absolute value of the magnitude of the coefficients [33, 34]. Rheumatic disease and cerebrovascular disease are not included in the penalized model. The AUC receiver operating characteristic curve (ROC) for the penalized logistic regression model was 0.809 (95% CI 0.801–0.816).

Random forest. Five hundred trees were generated and three variables were tried at each split. Out of bag estimate of error rate was 30.8%.

Predicting model performance. Model performance was assessed by the AUC of the ROC. pROC package was used to plot ROC and 95% CI [35].

Missing data. Missing data were handled with a simple imputation method by which missing data were replaced by the median of the cohort patients. Missing data for laboratory results were more common in the group that did not develop hospital-acquired AKI. There were no missing data in the diagnosis and procedural information.

RESULTS

Within a 2-year period, 61 705 patients were discharged and 54 095 were eligible for analysis (Figure 1). Of these, 6629 (12.3%, 95% CI 12.0–12.5%) had CA-AKI and were excluded.

A total of 47 466 patients were analysed and included in the discovery and internal validation cohorts. Of these, 2385 (5.0%, 95% CI 4.8–5.2%) developed hospital-acquired AKI (1864 KDIGO Stage 1, 378 Stage 2 and 143 Stage 3). The incidence of hospital-acquired AKI was 5.2% (1217/23 481) in 2015 and 4.9% (1168/23 985) in 2016. The length of stay (LOS) was longer for the AKI group [median AKI versus non-AKI: 10.0 (6.0–17.0) versus 4.0 (3.0–7.0) days, P < 0.0001; mean AKI versus non-AKI: 15.5 ± 22.2 versus 6.1 ± 7.5 days, P < 0.0001]. When adjusted for age and comorbidities, the mean LOS was 8.7 days longer for AKI
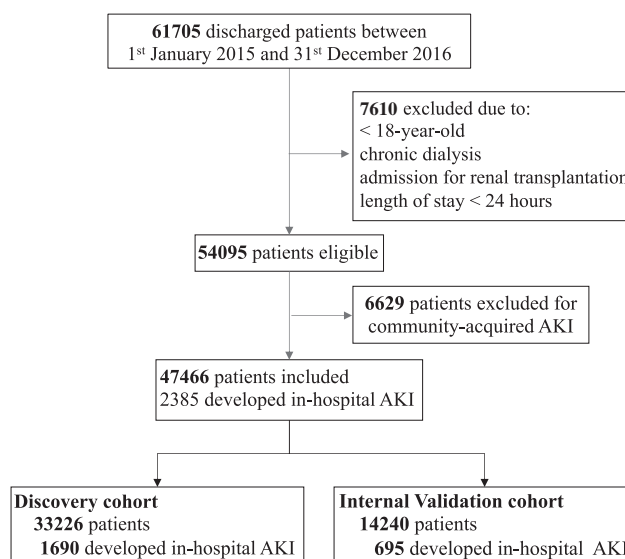

FIGURE 1: Disposition of patients.

patients (P < 0.0001). Mean All Patients Refined Diagnosis-Related Groups (APR-DRG) weight (1.97 ± 2.09 versus 0.99 ± 0.76 units; P < 0.0001) and mortality were also higher [16% (382/2385) versus 2.6% (1151/45 081, P < 0.0001)] in AKI than in non-AKI. Overall mortality was 3.2% in the study period. Supplementary data, Table S5, shows the admission department. A higher frequency of AKI was observed among patients admitted to the ICU (20.7% of admitted patients had AKI), nephrology (17.5%), cardiology and cardiac surgery (12.7% each), and vascular and endovascular surgery (8.3%) than from other departments.

Table 1 presents comorbidity and admission characteristics. Mean age of the general hospitalized population was 62.1 years. AKI patients were older (74.3 ± 15.0 versus 61.4 ± 20.1 years, P < 0.0001), more frequently male (53% versus 43%, P < 0.0001) and admitted urgently more frequently (72% versus 54%, P < 0.0001) than non-AKI patients. Past myocardial infarction, congestive heart failure, peripheral vascular disease, chronic pulmonary disease, connective tissue disease, peptic ulcer, liver disease and kidney disease were also more frequent in AKI patients.

Table 2 presents baseline analytical values. AKI patients had a lower baseline eGFR and higher serum creatinine, urinary protein-to-creatinine and urinary albumin-to-creatinine, although median values were within the normal laboratory range. Very small nominal differences for some biochemistry and haematology exams that remained within the normal range were statistically significant between AKI and non-AKI patients.

Multivariate analysis

Several initial multivariate models using comorbidities, baseline laboratory values and type of surgery and admission had AUC close to 0.80 for the prediction of hospital-acquired AKI. A random forest model using the same variables as in the initial logistic model yielded similar AUC-ROCs: AUC for predicting hospital-acquired AKI was 0.818 (95% CI 0.8098–0.8252).

The final model, which we termed the Madrid Acute Kidney Injury Prediction Score (MAKIPS), was obtained by step-wise regression followed by BMA and bootstrap validation (Table 3). The original C-index and corrected C-index were similar: 0.811 and 0.810, respectively. (Supplementary data, Figure S1). Table 3 presents the OR for hospital-acquired AKI conferred by

Table 1. Comorbidity and admission characteristics of the cohort patients

Variables	Total	Non-AKI	AKI	P-value
n	47 466	45 081	2385	
Men, % (n)	43.5 (20 647)	43.0 (19 389)	52.7 (1258)	<0.0001
Mean age (years), mean±SD	62.1±20.1	61.4±20.1	74.3±15.0	<0.0001
Diabetes, % (n)	12.2 (5786)	11.5 (5200)	24.6 (586)	<0.0001
Hypertension, % (n)	30.3 (14 392)	28.8 (13 027)	57.2 (1365)	<0.0001
Cardiovascular disease, % (n)	7.6 (3596)	7.0 (3157)	18.4 (439)	<0.0001
Cerebrovascular disease, % (n)	6 (2842)	5.7 (2607)	9.8 (235)	<0.0001
Anaemia, % (n)	11 (5205)	10.2 (4605)	25.1 (600)	<0.0001
Myocardial infarction, % (n)	2.8 (1363)	2.6 (1172)	8.0 (191)	<0.0001
Congestive heart failure, % (n)	6.7 (3222)	5.8 (2622)	25.1 (600)	<0.0001
Peripheral vascular disease, % (n)	3.9 (1867)	3.7 (1679)	7.8 (188)	<0.0001
Dementia, % (n)	0.6 (319)	0.6 (298)	0.8 (21)	0.25
Chronic pulmonary disease, % (n)	13.4 (6385)	13.0 (5869)	21.6 (516)	<0.0001
Connective tissue disease, % (n)	1.7 (809)	1.6 (732)	3.2 (77)	<0.0001
Peptic ulcer disease, % (n)	0.5 (265)	0.5 (237)	1.1 (28)	<0.0001
Liver disease, % (n)	5.3 (2535)	5.1 (2317)	9.1 (218)	<0.0001
Hemiplegia, % (n)	1.0 (506)	1.0 (454)	2.1 (52)	<0.0001
Renal disease, % (n)	6.0 (2849)	5.1 (2336)	21.5 (513)	<0.0001
Malignancy, % (n)	15.0 (7142)	14.7 (6652)	20.5 (490)	<0.0001
Metastatic solid tumour, % (n)	6.5 (3107)	6.4 (2901)	8.6 (206)	<0.0001
AIDS/HIV, % (n)	0.6 (294)	0.6 (281)	0.5 (13)	0.73
Urgent admission, % (n)	54.6 (25 916)	53.6 (24 200)	71.9 (1716)	<0.0001
Surgical patients, % (n)	45.6 (21 633)	45.7 (20 626)	42.2 (1007)	<0.0001
ASA classification, % (n)[a]				<0.0001
Total (n)	21 568	20 787	781	
I	14.5 (3132)	14.9 (3107)	3.2 (25)	
II	53.6 (11 560)	54.2 (11 263)	38.0 (297)	
III	29.7 (6401)	28.8 (5987)	53.0 (414)	
IV	2.2 (475)	2.1 (430)	5.8 (45)	

[a]This was available in many surgical patients. Slight discrepancies between surgical patients and ASA classification explained by emergency surgery without ASA or ASA assessment with later cancelled surgery. ASA, American Society of Anesthesiologists.

individual variables in the model. The highest ORs (all >2.0) were conferred by abdominal surgery (OR 3.92; 95% CI 3.13–4.89), cardiovascular surgery (OR 2.94; 95% CI 2.94–4.29) and urological surgery (OR 2.91; 95% CI 2.34–3.60) followed by congestive heart failure (OR 2.73; 95% CI 2.42–3.08), hemiplegia/paraplegia (OR 2.10; 95% CI 1.52–2.84) and urgent admission (OR 2.13; 95% CI 1.89–2.40). Among continuous variables, the highest OR was conferred by increasing age (OR 162; 95% CI 1.51–1.74) and urea and uric acid levels (OR 1.06 for both; 95% CI 1.04–1.09 and 1.01–1.11, respectively). The year of discharge was not significant when added to the model.

To correct the overestimation of the model without affecting its discrimination capacity, a penalized logistic regression model was estimated by the lasso method of coefficient shrinkage after estimation (Table 4). The AUC-ROC for the penalized logistic regression model was 0.809 (95% CI 0.801–0.816) and had 21 variables (Supplementary data, Figure S2).

A web-based calculator (http://www.bioestadistica.net/MAKIPS.aspx) is available to calculate the MAKIPS and calculate whether the risk of AKI is >20%. We envision that the model will be most useful in settings in which the information can be directly and automatically obtained from electronic clinical records.

DISCUSSION

The main finding is the description of the MAKIPS, a risk score to predict hospital-acquired AKI at admission from electronic medical records. It comprises 21 baseline variables, including

comorbidities, laboratory values and elective surgical interventions if applicable, that are easily accessible and available from electronic records and have a good predictive ability (ROC-AUC = 0.81).

Clinical risk scores for AKI in non-critical populations are often limited to very specific diseases or populations, such as cirrhosis, contrast nephropathy or patients receiving cisplatin [36–40]. The MAKIPS was developed and internally validated for a general hospitalized population, with a large variety of comorbidities, and medical and surgical conditions. In the field of general non-critical emergency admissions, there is, to the extent of our knowledge, one clinical risk score that has been externally validated: the AKI prediction score (APS), which comprises seven clinical variables: age, respiratory rate, the AVPU (alert, voice, pain or unresponsive) scale of responsiveness, CKD Categories G3–5, heart failure, diabetes and liver disease. It was externally validated in a single UK non-specialist acute hospital, yielding an AUC-ROC of 0.65 (95% CI 0.62–0.67) in patients with known baseline creatinine [41]. The incidence of hospital-acquired AKI in the aforementioned study was 8.1%. The AUC of the MAKIPS equation (0.811; 95% CI 0.795–0.825 in the validation cohort) compares favourably with the APS. While the APS is a simple score designed to be calculated manually or by filling a checklist by clinicians, the MAKIPS may be calculated automatically from electronic clinical records.

The incidence of hospital-acquired AKI in our study (5%) is below that reported in part of the literature for global incidence for AKI as defined with the KDIGO criteria, which were estimated in a meta-analysis to be 21.6% [42]. However, it is in line

Table 2. Baseline analytical values

Serum biochemistry	n	Total	Non-AKI	AKI	P-value
sCr (mg/dL)	38 350	0.80 (0.60–1.00)	0.80 (0.60–1.00)	0.93 (0.70–1.30)	<0.0001
eGFR (mL/min/1.73 m^2)	38 350	91.33 (72.23–116.42)	92.15 (73.10–117.11)	72.40 (50.36–99.89)	<0.0001
Uric acid (mg/dL)	26 541	5.10 (4.00–6.50)	5.10 (3.90–6.40)	6.00 (4.60–7.50)	<0.0001
Albumin (g/dL)	32 952	4.00 (3.70–4.30)	4.10 (3.70–4.30)	3.80 (3.40–4.10)	<0.0001
Calcium (mg/dL)	27 891	9.00 (8.50–9.30)	9.00 (8.50–9.30)	8.70 (8.20–9.20)	<0.0001
Ionic calcium (mg/dL)	6018	4.46 (4.29–4.65)	4.46 (4.29–4.65)	4.45 (4.21–4.65)	0.13
Phosphate (mg/dL)	24 272	3.40 (3.00–3.80)	3.40 (3.00–3.80)	3.40 (2.90–3.80)	0.78
Alkaline phosphatase (UI/l)	26 346	76.50 (61.00–99.00)	76.00 (60.00–99.00)	80.00 (64.00–108.00)	<0.0001
Glucose (mg/dL)	45 629	98.00 (87.00–120.00)	98.00 (86.00–119.00)	111.00 (93.00–144.00)	<0.0001
HbA1c (%)	9022	6.00 (5.50–6.80)	6.00 (5.50–6.80)	6.30 (5.70–7.30)	<0.0001
LDH (UI/l)	26 767	371.00 (371.00–452.00)	368.00 (316.00–449.00)	410.00 (342.00–517.00)	<0.0001
CRP (mg/dL)	26 040	2.40 (0.90–6.30)	2.30 (0.90–6.30)	3.00 (1.30–7.10)	<0.0001
Total proteins (g/dL)	26 225	6.60 (6.10–7.10)	6.60 (6.10–7.10)	6.40 (5.90–6.90)	<0.0001
Sodium (mEq/l)	43 690	139.00 (137.00–141.00)	139.00 (137.00–141.00)	138.00 (136.00–141.00)	<0.0001
Potassium (mEq/l)	34 882	4.10 (3.80–4.40)	4.10 (3.80–4.40)	4.20 (3.80–4.60)	<0.0001
CO_2 (mEq/l)	1795	29.00 (26.00–31.00)	29.00 (26.00–31.00)	28.00 (25.00–31.00)	0.0448
Urea (mg/dL)	38 355	36.00 (27.00–49.00)	36.00 (27.00–48.00)	48.00 (35.00–68.00)	<0.0001
Haematology					
Haemoglobin (g/dL)	46 405	12.90 (11.60–14.20)	13.00 (11.60–14.20)	12.20 (10.70–13.50)	<0.0001
Leucocytes (per μL)	46 398	7.94 (6.19–10.35)	7.90 (6.18–10.28)	8.62 (6.42–11.74)	<0.0001
Urinary biochemistry					
Density	29 803	1.010 (1.010–1.020)	1.010 (1.010–1.020)	1.010 (1.010–1.020)	<0.0001
Creatinine (mg/dL)	11 002	74.00 (50.00–112.00)	74.00 (50.00–113.00)	69.00 (48.00–96.00)	<0.0001
Albumin (mg/l)	8998	7.90 (3.30–32.40)	7.50 (3.20–29.60)	17.80 (4.60–89.90)	<0.0001
Sodium (mEq/l)	3212	62.00 (35.00–92.00)	62.00 (35.00–92.00)	61.00 (39.00–89.00)	<0.0001
UACR (mg/g)	8999	10.00 (3.97–40.35)	9.32 (3.85–35.17)	24.74 (6.81–148.00)	<0.0001
UPCR (mg/g)	4669	87.20 (52.10–251.30)	83.30 (51.10–210.10)	169.60 (70.02–751.83)	<0.0001

Values expressed as median (IQR 25–75%).
IQR, interquartile range; sCr, serum creatinine; eGFR (mL/min/1.72 m^2) assessed by the MDRD-4 equation; HbA1c, glycated haemoglobin; LDH, lactate dehydrogenase; CRP, C-reactive protein; UACR, urinary albumin:creatinine ratio; UPCR: urinary protein:creatinine ratio.

Table 3. Final multivariate model selected by step-wise logistic regression followed by BMA: MAKIPS

Variables	Estimate	SE	Z-value	OR (2.5–97.5%)	P-value
Intercept	−4.5460	0.0695	−65.37		<0.0001
Abdominal surgery	1.3662	0.1141	11.97	3.92 (3.13–4.89)	<0.0001
Cardiovascular surgery	1.2688	0.0963	13.17	3.56 (2.94–4.29)	<0.0001
Urological surgery	1.0682	0.1104	9.68	2.91 (2.34–3.60)	<0.0001
Congestive heart failure	1.0043	0.0621	16.17	2.73 (2.42–3.08)	<0.0001
Hemiplegia	0.7409	0.1596	4.64	2.10 (1.52–2.84)	<0.0001
Renal disease	0.6595	0.0679	9.71	1.93 (1.69–2.21)	<0.0001
Rheumatic disease	0.3848	0.1294	2.97	1.47 (1.13–1.88)	0.001
Liver disease	0.4832	0.0800	6.04	1.62 (1.38–1.89)	<0.0001
Malignancy	0.3281	0.0607	5.41	1.39 (1.23–1.56)	<0.0001
Cardiovascular disease	0.1637	0.0672	2.44	1.18 (1.03–1.34)	0.01
Cerebrovascular disease	0.1908	0.0770	2.48	1.21 (1.04–1.40)	0.01
Anaemia	0.477	0.054	8.73	1.61 (1.45–1.79)	<0.0001
Diabetes	0.150	0.0597	2.52	1.16 (1.03–1.31)	0.01
Surgical admission	0.5338	0.0626	8.53	1.71 (1.51–1.93)	<0.0001
Urgent admission	0.7554	0.0603	12.54	2.13 (1.89–2.40)	<0.0001
Age (years)	0.4842	0.0361	13.40	1.62 (1.51–1.74)	<0.0001
Uric acid (mg/dL)	0.0614	0.0239	2.57	1.06 (1.01–1.11)	0.01
Urea (mg/dL)	0.0624	0.0119	5.23	1.06 (1.04–1.09)	<0.0001
Calcium (mg/dL)	−0.1887	0.0210	−8.97	0.83 (0.79–0.86)	<0.0001
Leucocytes (n/μL)	0.0388	0.0080	4.85	1.04 (1.02–1.05)	<0.0001
Sodium (mEq/L)	−0.0455	0.0142	−3.20	0.96 (0.93–0.98)	0.001
Glucose (mg/dL)	0.0465	0.0097	4.77	1.05 (1.03–1.07)	<0.0001
(Potassium)^2 (mEq/L)	0.0188	0.0060	3.11	1.02 (1.01–1.03)	0.001

The comorbidities variables refer to all diseases in that category of ICD-9 code (see Supplementary data, Table S3, for more information).

Table 4. Penalized multivariate model selected by step-wise logistic regression followed by BMA: MAKIPS

Variables	Coefficient
Intercept	−3.7508
Abdominal surgery	1.2187
Cardiovascular surgery	1.2955
Urological surgery	0.6865
Congestive heart failure	0.8753
Hemiplegia	0.2393
Renal disease	0.6158
Liver disease	0.1879
Malignancy	0.1365
Cardiovascular disease	0.0940
Anaemia	0.4046
Diabetes	0.0844
Surgical admission	0.0186
Urgent admission	0.3341
Age (years)	0.3988
Uric acid (mg/dL)	0.0150
Urea (mg/dL)	0.063
Calcium (mg/dL)	−0.1415
Leucocytes (n/μL)	0.0203
Sodium (mEq/L)	−0.0107
Glucose (mg/dL)	0.0380
(Potassium)^2 (mEq/L)	0.0104

The comorbidities variables refer to all diseases in that category of ICD-9 code (see Supplementary data, Table S3, for more information). The penalized model does not include rheumatic disease and cerebrovascular disease.

with that reported in general hospitals. In this regard, it is useful to note that 54% of the patients included in the widely cited meta-analysis were ICU and cardiac surgery patients [42]. AKI incidence in these high-risk groups is reported between 24.1% and 76.6% [43] and they are often over-represented in AKI epidemiological studies. Our incidence is closer to that reported across European population cohorts. The global incidence of AKI (CA-AKI plus hospital-acquired AKI) was 8.4% for patients with baseline eGFR >60 mL/min/1.73 m^2 and 17.6% in those with <60 mL/min/1.73 m^2 in Scotland [44] and 12% in Ireland [45]. This is similar to the 16.7% (9014/54 095) combined incidence of CA-AKI and hospital-acquired AKI in our study. A Swiss study excluding critical patients and, similar to our study, CA-AKI, reported an incidence of hospital-acquired AKI of 4.11% [46]. The design of this Swiss study is the most comparable to our design and the incidence of hospital-acquired AKI was also very similar. In multicentre studies from China on general-hospitalized population and using the KDIGO criteria, hospital-acquired AKI incidence varies between 3.0% and 11.6% [7, 47]. AKI incidence, even if defined with the same KDIGO criteria, may vary with case-mix (primary, secondary or tertiary care centres), exclusion or inclusion of CA-AKI, availability of baseline creatinine determination and clinical application of the criteria (revised versus non-revised by nephrologist) [48]. The incidence of CA-AKI, defined as AKI by KDIGO criteria that are already present when the patient arrives at the emergency room, has been reported at 8.3% in the recent ICE-AKI (Impact analysis of a Clinical prediction rule and Electronic AKI) study [49]. Previous reports from the UK, Canada and Portugal had reported CA-AKI incidences of 4.6, 19.6 and 23.6%, respectively, of urgent admissions [5, 9, 20, 50]. CA-AKI incidence in our cohort (12%) was similar to that reported in the literature for similar study populations. Of note, our 17.5% incidence of in-hospital AKI may

seem low for a Nephrology Department, but CA-AKI, a frequent cause of admission in Nephrology, was excluded from the analysis.

Among the study strengths were that this was a large study with internal validation. In our study, models derived from logistic and machine-learning techniques such as random forest approaches had similar AUC-ROCs. The logistic model was chosen as clinicians are more familiar with its methodology and interpretation. The model can predict baseline risk for hospital-acquired AKI using variables that are widespread in clinical practice and in electronic clinical records, so it would ideally be automatically calculated upon admission (elective or urgent), to maximize prophylactic measures, which according to the NCEPOD (National Confidential Enquiry into Patient Outcome and Death) report, remains an unmet clinical need [51]. Thus, poor recognition of AKI risk factors in routine clinical practice led to inadequate clinical management in 29% of AKI cases [51], including failures in physiological monitoring, timely laboratory tests, intravenous fluids and recognition of acute illness, sepsis and hypovolemia. A risk score for AKI may help identify the patients in which these basic actions are absolutely paramount. Furthermore, the dataset was generated in a tertiary academic hospital caring for all types of medical and surgical patients and was not limited to ICU or high-risk surgical patients.

Our study is limited to a single centre. Although data were obtained prospectively through electronic medical records, part of the medical records was based on previously coded events in the hospital database. Identification of comorbidities was based on ICD-9 codes, which may be unreliable, and the absence of urine output data prevented a more precise AKI definition. In those patients with no available baseline creatinine, AKI may have been misdiagnosed. An external validation of the model is needed to address its generalizability to other centres or countries with different case-mix or healthcare systems. In this regard, in Spain, primary care and specialized care and hospitalization are free at the point-of-care and there are no barriers to access specialized care. Specifically, in Madrid, primary care and specialized care are integrated. Thus, the model should be validated in settings of limited access to healthcare. In this sense, an external validation with a validation dataset from another hospital, or a new prospectively collected dataset within our own institution temporally separated from the development cohort, would strengthen the study, as the penalized model was developed using the current dataset.

In conclusion, we have generated the MAKIPS score, which can be automatically calculated from electronic clinical records to predict at admission the risk of hospital-acquired AKI. Prediction of AKI risk may be more useful in decreasing the incidence of AKI than current electronic alerts that do alert, but only after AKI has already occurred. However, external validation in other health-care and hospital settings is required. For this purpose, an online tool has been set up. Furthermore, future research should focus on impact analysis and the use of machine-learning techniques to evaluate AKI risk and prediction.

SUPPLEMENTARY DATA

Supplementary data are available at ckj online.

FUNDING

ERA-PerMed-JTC2018 (KIDNEY ATTACK AC18/00064 and PERSTIGAN AC18/00071) and Sociedad Española de Nefrología, FRIAT and Comunidad de Madrid B2017/BMD-3686 CIFRA2.

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES

1. Jones J, Holmen J, De Graauw J *et al*. Association of complete recovery from acute kidney injury with incident CKD stage 3 and all-cause mortality. *Am J Kidney Dis* 2012; 60: 402–408
2. Basile DP, Donohoe D, Roethe K *et al*. Renal ischemic injury results in permanent damage to peritubular capillaries and influences long-term function. *Am J Physiol Renal Physiol* 2001; 281: F887–F899
3. Sawhney S, Marks A, Ali T *et al*. Maximising acute kidney injury alerts - a cross-sectional comparison with the clinical diagnosis. *PLoS One* 2015; 10: e0131909
4. Singh P, Rifkin DE, Blantz RC. Chronic kidney disease: an inherent risk factor for acute kidney injury? *Clin J Am Soc Nephrol* 2010; 5: 1690–1695
5. Wonnacott A, Meran S, Amphlett B *et al*. Epidemiology and outcomes in community-acquired versus hospital-acquired AKI. *Clin J Am Soc Nephrol* 2014; 9: 1007–1014
6. Schissler MM, Zaidi S, Kumar H *et al*. Characteristics and outcomes in community-acquired versus hospital-acquired acute kidney injury. *Nephrology* 2013; 18: 183–187
7. Xu X, Nie S, Liu Z *et al*. Epidemiology and clinical correlates of AKI in Chinese hospitalized adults. *Clin J Am Soc Nephrol* 2015; 10: 1510–1518
8. NCEPOD. NCEPOD - Acute Kidney Injury: Adding Insult to Injury Report. London, 2009 (cited 1 May 2017). http://www.ncepod.org.uk/2009aki.html (2 May 2018, date last accessed)
9. Aitken E, Carruthers C, Gall L *et al*. Acute kidney injury: outcomes and quality of care. *QJM* 2013; 106: 323–332
10. Levey A, Greene T, Kusek J *et al*. A simplified equation to predict glomerular filtration rate from serum creatinine. *J Am Soc Nephrol* 2000; 11: 155A
11. MDRD Study Equation. *National Kidney Foundation* (cited 2018 November 15). https://www.kidney.org/content/mdrd-study-equation (1 October 2018, date last accessed)
12. Kellum JA, Lameire N, Aspelin P *et al*. KDIGO Clinical Practice Guideline for Acute Kidney Injury. *Kidney Int Suppl* 2012; 2: 1–138
13. Hsu C, McCulloch CE, Fan D *et al*. Community-based incidence of acute renal failure. *Kidney Int* 2007; 72: 208–212
14. Patschan D, Müller GA. Acute kidney injury in diabetes mellitus. *Int J Nephrol* 2016; 2016: 6232909
15. Rosner MH, Okusa MD. Acute kidney injury associated with cardiac surgery. *Clin J Am Soc Nephrol* 2006; 1: 19–32
16. Danziger J, Chen KP, Lee J *et al*. Obesity, acute kidney injury, and mortality in critical illness. *Crit Care Med* 2016; 44: 328–334
17. Bagshaw SM, Laupland KB, Doig CJ *et al*. Prognosis for long-term survival and renal recovery in critically ill patients with severe acute renal failure: a population-based study. *Crit Care* 2005; 9: R700–R709
18. Anderson S, Eldadah B, Halter JB *et al*. Acute kidney injury in older adults. *J Am Soc Nephrol* 2011; 22: 28–38
19. Cho K, Hsu C. Quantifying severity of chronic kidney disease as a risk factor for acute kidney injury. *J Am Soc Nephrol* 2010; 21: 1602–1604

20. James MT, Hemmelgarn BR, Wiebe N *et al*. Glomerular filtration rate, proteinuria, and the incidence and consequences of acute kidney injury: a cohort study. *Lancet* 2010; 376: 2096–2103
21. Grams ME, Astor BC, Bash LD *et al*. Albuminuria and estimated glomerular filtration rate independently associate with acute kidney injury. *J Am Soc Nephrol* 2010; 21: 1757–1764
22. Levy EM, Viscoli CM, Horwitz RI. The effect of acute renal failure on mortality. *JAMA* 1996; 275: 1489
23. Fox CS, Muntner P, Chen AY *et al*. Short-term outcomes of acute myocardial infarction in patients with acute kidney injury a report from the national cardiovascular data registry. *Circulation* 2012; 125: 497
24. Bakris GL, Ritz E. The message for World Kidney Day 2009: hypertension and kidney disease - a marriage that should be prevented. *J Hypertens* 2009; 27: 666–669
25. Quan H, Sundararajan V, Halfon P *et al*. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 2005; 43: 1130–1139
26. Quan H, Parsons GA, Ghali WA. Validity of information on comorbidity derived rom ICD-9-CCM administrative data. *Med Care* 2002; 40: 675–685
27. Clyde M. Bayesian model averaging and model search strategies. *Bayesian Stat* 1999; 6: 309–322
28. Harrell FE. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York: Springer, 2001
29. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. Switzerland: Springer, 2009
30. Fox J, Monette G. Generalized collinearity diagnostics. *J Am Stat Assoc* 1992; 87: 178–183
31. Fox J. *Applied Regression Analysis and Generalized Linear Models*, 2nd edn. CA: SAGE, 2008, 791
32. Fox J, Weisberg S. *An R Companion to Applied Regression: Appendices*, 2nd edn. Thousand Oaks, CA: Sage, 2011. http://socserv.socsci.mcmaster.ca/jfox/Books/Companion (5 February 2018, date last accessed)
33. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc B Stat Methodol* 1996; 58: 267–288
34. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 2010; 33: 1–22
35. Robin X, Turck N, Hainard A *et al*. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011; 12: 77
36. Pan H-C, Jenq C-C, Tsai M-H *et al*. Risk models and scoring systems for predicting the prognosis in critically ill cirrhotic patients with acute kidney injury: a prospective validation study. *PLoS One* 2012; 7: e51094
37. Yang LK, Ping ZW, Jie BW *et al*. A novel risk score model for prediction of contrast-induced nephropathy after emergent percutaneous coronary intervention. *Int J Cardiol* 2017; 230: 402–412
38. Duan C, Cao Y, Liu Y *et al*. A new preprocedure risk score for predicting contrast-induced acute kidney injury. *Can J Cardiol* 2017; 33: 714–723
39. Liu YH, Liu Y, Zhou YL *et al*. Comparison of different risk scores for predicting contrast induced nephropathy and outcomes after primary percutaneous coronary intervention in patients with ST elevation myocardial infarction. *Am J Cardiol* 2016; 117: 1896–1903

40. Motwani SS, McMahon GM, Humphreys BD *et al.* Development and validation of a risk prediction model for acute kidney injury after the first course of cisplatin. *J Clin Oncol* 2018; 36: 682–688

41. Hodgson LE, Dimitrov BD, Roderick PJ *et al.* Predicting AKI in emergency admissions: an external validation study of the acute kidney injury prediction score (APS). *BMJ Open* 2017; 7: e013511

42. Susantitaphong P, Cruz DN, Cerda J *et al.* World incidence of AKI: a meta-analysis. *Clin J Am Soc Nephrol* 2013; 8: 1482–1493

43. Xiong J, Tang X, Hu Z *et al.* The RIFLE versus AKIN classification for incidence and mortality of acute kidney injury in critical ill patients: a meta-analysis. *Sci Rep* 2015; 5: 17917

44. Sawhney S, Robinson HA, Van Der Veer SN *et al.* Acute kidney injury in the UK: a replication cohort study of the variation across three regional populations. *BMJ Open* 2018; 8: e019435

45. Sawhney S, Marks A, Fluck N *et al.* Intermediate and long-term outcomes of survivors of acute kidney injury episodes: a large population-based cohort study. *Am J Kidney Dis* 2017; 69: 18–28

46. Meier P, Bonfils RM, Vogt B *et al.* Referral patterns and outcomes in noncritically ill patients with hospital-acquired acute kidney injury. *Clin J Am Soc Nephrol* 2011; 6: 2215–2225

47. Yang L, Xing G, Wang L *et al.* Acute kidney injury in China: a cross-sectional survey. *Lancet* 2015; 386: 1465–1471

48. Hoste EAJ, Kellum JA, Selby NM *et al.* Global epidemiology and outcomes of acute kidney injury. *Nat Rev Nephrol* 2018; 1: 607–625

49. Hodgson LE, Roderick PJ, Venn RM *et al.* The ICE-AKI study: impact analysis of a clinical prediction rule and electronic AKI alert in general medical patients. *PLoS One* 2018; 13: e0200584

50. Challiner R, Ritchie JP, Fullwood C *et al.* Incidence and consequence of acute kidney injury in unselected emergency admissions to a large acute UK hospital trust. *BMC Nephrol* 2014; 15: 84

51. Stewart J, Findlay G, Smith N *et al.* Adding insult to injury: a review of the care of patients who died in hospital with a primary diagnosis of acute kidney injury. *Natl Confid Enq into Patient Outcome Death* 2009; 22: 1–22