RESEARCH ARTICLE

Prediction Models and Their External Validation Studies for Mortality of Patients with Acute Kidney Injury: A Systematic Review

Tetsu Ohnuma[1], Shigehiko Uchino[2]*

1 Intensive Care Unit, Department of Anesthesiology, Saitama Medical Center, Jichi Medical University, Saitama, Japan, **2** Intensive Care Unit, Department of Anesthesiology, Jikei University School of Medicine, Tokyo, Japan

* s.uchino@mac.com
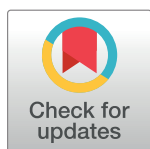
Abstract

Objectives

To systematically review AKI outcome prediction models and their external validation studies, to describe the discrepancy of reported accuracy between the results of internal and external validations, and to identify variables frequently included in the prediction models.

Methods

We searched the MEDLINE and Web of Science electronic databases (until January 2016). Studies were eligible if they derived a model to predict mortality of AKI patients or externally validated at least one of the prediction models, and presented area under the receiver-operator characteristic curves (AUROC) to assess model discrimination. Studies were excluded if they described only results of logistic regression without reporting a scoring system, or if a prediction model was generated from a specific cohort.

Results

A total of 2204 potentially relevant articles were found and screened, of which 12 articles reporting original prediction models for hospital mortality in AKI patients and nine articles assessing external validation were selected. Among the 21 studies for AKI prediction models and their external validation, 12 were single-center (57%), and only three included more than 1,000 patients (14%). The definition of AKI was not uniform and none used recently published consensus criteria for AKI. Although good performance was reported in their internal validation, most of the prediction models had poor discrimination with an AUROC below 0.7 in the external validation studies. There were 10 common non-renal variables that were reported in more than three prediction models: mechanical ventilation, age, gender, hypotension, liver failure, oliguria, sepsis/septic shock, low albumin, consciousness and low platelet count.

Conclusions

Information in this systematic review should be useful for future prediction model derivation by providing potential candidate predictors, and for future external validation by listing up the published prediction models.

Introduction

Acute kidney injury (AKI) is a common complication among critically ill patients and their mortality is high [1–4]. Reliable AKI specific scoring systems are important to predict outcome of AKI patients and to provide severity stratification for clinical studies. However, general severity scores for critically ill patients, e.g., Acute Physiology and Chronic Health Evaluation (APACHE) [5–7], Simplified Acute Physiology Score (SAPS) [8, 9], and Mortality Probability Model [10] have shown controversial results on the accuracy of predicting mortality in AKI patients [11–13], partly because those scores were generated from data that included only a few AKI patients.

Over the past three decades, multiple AKI outcome prediction models, which incorporated physiologic, laboratory, organ dysfunction and previous comorbidity, have been derived [14–20]. Even in the 21st century, five additional prediction models have been generated [12, 21–24]. Although internal validation of these prediction models has shown good accuracy, the results of external validation studies for the models have been unsatisfactory [11, 25, 26]. Currently, there is neither consensus nor guideline recommending which prediction model to apply to clinical practice.

The objectives of this study are to systematically review the AKI outcome prediction models and their external validation studies, to describe the discrepancy of reported accuracy between the results of internal and external validations, and to identify variables frequently included in the prediction models, which might be potentially useful for future prediction model derivation.

Materials and Methods

Studies eligible for review

Studies published in the medical literature were eligible if they derived a model to predict mortality of AKI patients or externally validated at least one of the prediction models, and presented area under the receiver-operator characteristic curves (AUROC) [27] or the concordance index (c-statistic) to assess model discrimination. Studies were excluded if they described only results of logistic regression without reporting a scoring system, or if a prediction model was generated from a specific cohort. Unpublished conference abstracts were also excluded. This study followed the same principal as in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (S1 PRISMA Checklist) [28].

Literature review and study selection

We searched the MEDLINE and Web of Science electronic databases (until January 2016). In the MEDLINE search, we used the terms of "acute kidney injury" (MeSH Terms), "statistical model" (MeSH Terms), "predictive value of tests" (MeSH Terms) and "validation". In the Web of Science, we used Key words of "acute kidney injury", "acute renal failure", "model", "prediction", "predictor", "validity", and "validation". References of all selected articles were searched

to identify any eligible studies. The search was restricted to human subjects. Each article selected by the primary reviewer (TO) was assessed by the second reviewer to confirm eligibility (SU).

Data extraction

A standardized data abstraction form was used to collect data on study characteristics and outcomes of interest. Data collected to describe characteristics of articles for original outcome prediction models were the type of study, study period, number of centers, sample size, mean age, gender, region, population, renal replacement therapy (RRT) requirement, hospital mortality, AKI definition, exclusion criteria, follow-up and variables included in prediction models. Following information was also collected for quality assessment of the prediction models: definition of predictors, indications for RRT defined, missing data definition, bootstrap resampling, multivariable analysis approach, event per variable ratio and internal validation cohort.

Data collected to describe characteristics of articles for external validation were type of study, study period, number of centers, sample size, mean age, hospital mortality, number of validated models and methods of discrimination and calibration. AUROCs reported in both original prediction models and external validation studies were also collected.

Results

A total of 2204 potentially relevant articles were found and screened, of which 80 were retrieved for detailed evaluation (Fig 1). We excluded five articles that had no prediction models developed by multivariate regression analysis, six articles that had no discrimination results,
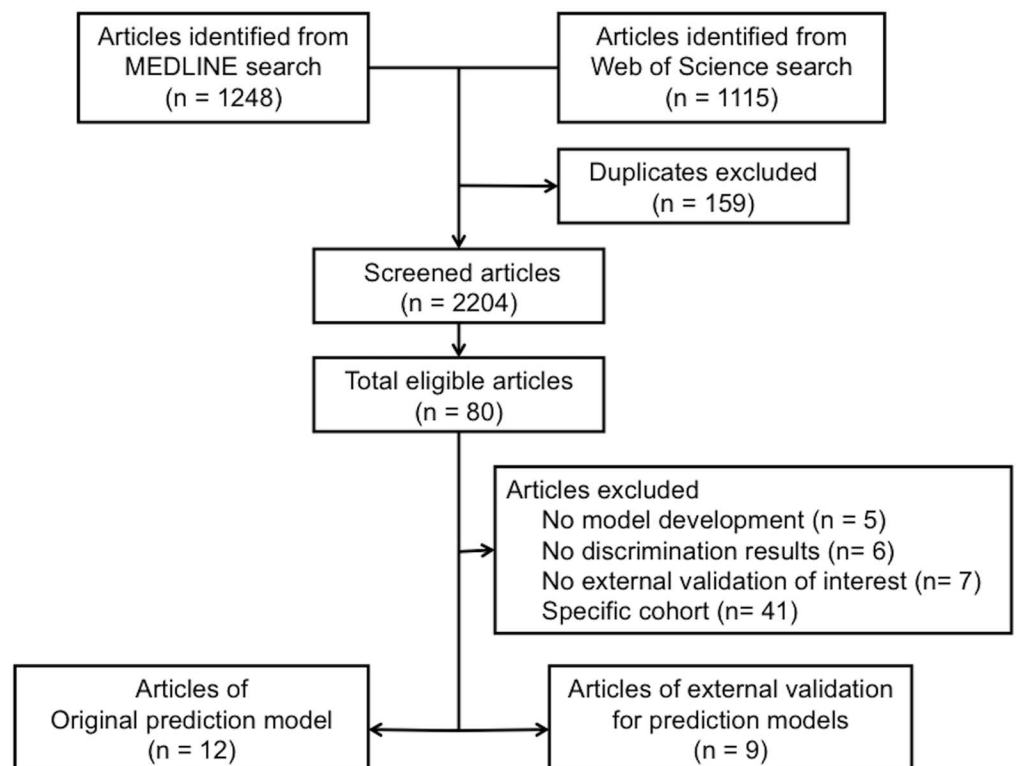

Fig 1. Selection of articles by PRISMA flow diagram

doi:10.1371/journal.pone.0169341.g001

Table 1. Characteristics of articles reporting outcome prediction models for acute kidney injury.

	Type of study	Study period	Centers, Number	Sample size	Mean age, years	Gender, Male %	Region	Population	RRT requirement	Hospital mortality
Ramussen 1985 [14]	Retrospective	1977–1981	1	148	58	NR	Australia	Hospital	50%	53%
Lohr 1988 [15]	Retrospective	1979–1985	1	126	57	73%	USA	Hospital	100%	75%
Schaefer 1991 [16]	Prospective	1985–1988	1	134	NR	NR	Germany	ICU	100%	57%
Liano 1993 [17]	Prospective	1977–1988	1	328	57	81%	Spain	Hospital	51%	53%
Paganini 1996 [18]	Retrospective	1988–1992	1	506	63	61%	USA	ICU	100%	67%
Chertow 1998 [19]	Post hoc of RCT	1993–1995	59	256	62	65%	USA, Canada	Hospital	42%	36%
Lins 2000 [20]	Prospective	1996–1997	1	197	70	60%	Belgium	ICU	26%	53%
Mehta 2002 [12]	Prospective	1989–1995	4	605	56	72%	USA	ICU	50%	52%
Lins 2004 [21]	Prospective	1997–1998	8	293	72	62%	Belgium	ICU	37%	51%
Dharan 2005 [22]	Prospective	2002	1	265	48	71%	India	Hospital	26%	38%
Chertow 2006 [23]	Prospective	1999–2001	5	618	59	59%	USA	ICU	64%	37%
Demirjian 2011 [24]	Post hoc of RCT	2003–2007	27	1,122	60	71%	USA	ICU	99%	50%

RRT: renal replacement therapy, NR: not reported, RCT: randomized control trial.

doi:10.1371/journal.pone.0169341.t001

seven articles that validated only general severity scores or had no external discrimination results and 41 articles that assessed specific cohorts (cardiac surgery: 10, contrast-induced nephropathy: eight, others: 23). Fifty-nine articles excluded from this study are listed in a supplement file (S1 File). Finally, 12 articles reporting original prediction models for hospital mortality in AKI patients [12, 14–24] and nine additional articles assessing external validation of the outcome prediction models [11, 25, 26, 29–34] were selected for analysis. Five out of 12 articles reporting original prediction models also assessed other models (14 articles in total for external validation).

Characteristics of the 12 articles reporting outcome prediction models for AKI are shown in Tables 1 and 2. The study sample size ranged from 126 to 1,122 patients and the hospital mortality ranged from 36% to 75%. Only five studies (Chertow 1998, Mehta, Lins 2004, Chertow 2006, Demirjian) included more than one center and remaining seven were conducted in single center. The definition of AKI was not uniform among the 12 articles and none used recently published consensus definitions for AKI. Quality assessment for these articles is shown in Table 3. How missing data were dealt was defined only in four articles, and all of these articles also used bootstrap resampling. Eight articles used multivariable logistic regression analysis, and the other four articles (Ramussen, Schaefer, Liano and Lins 2000) used multivariable linear regression analysis. The event per variable ratio was more than 10 in all articles except for the earliest (Ramussen).

Characteristics of the 14 external validation studies are shown in Table 4. The study sample size ranged from 197 to 17,326 patients and the hospital mortality ranged from 37% to 85%.

Table 2. AKI definitions, exclusion criteria and follow-up of articles reporting outcome prediction models for acute kidney injury.

	AKI definitions	Exclusion criteria	Follow-up
Ramussen 1985 [14]	SCR > 2.0 mg/dl, or more than 50% elevation if baseline SCR > 1.7mg/dl,	Glomerulonephritis, uric acid nephropathy, ureteric obstruction, interstitial nephritis	Hospital
Lohr 1988 [15]	Requiring RRT	Post-renal transplantation	NR
Schaefer 1991 [16]	Requiring RRT	Chronic HD and kidney transplantation	ICU
Liano 1993 [17]	SCR >2.0 mg/dl	Previous renal failure and hepato-renal syndrome, vascular, interstitial, glomerular, obstructive etiology	NR
Paganini 1996 [18]	Requiring RRT	Less than 18 years	NR
Chertow 1998 [19]	SCI of >1mg/dl	Pre-renal azotemia, urinary obstruction, glomerulonephritis, interstitial nephritis, CKD, renal transplantation	30-day
Lins 2000 [20]	SCR > 2.0 mg/dl or more than 50% elevation		Hospital
Mehta 2002 [12]	SCR ≥ 2.0 mg/dl, BUN ≥ 40 mg/dl, or SCI ≥ 1.0 mg/dl with preexisting renal insufficiency	Previous dialysis, kidney transplantation, urinary obstruction, hypovolemia	Hospital
Lins 2004 [21]	SCR >2.0 mg/dl or >50% increase in preexisting mild to moderate renal disease	Baseline SCR >3.0 mg/dl	Hospital
Dharan 2005 [22]	SCI ≥ 0.5 mg/dl with baseline SCR less than 1.9 mg/dl, or SCI ≥ 1.0 mg/dl with baseline SCR between 2.0 to 4.9 mg/dl	Baseline SCR >5.0 mg/dl, transplant recipients	NR
Chertow 2006 [23]	SCI ≥ 0.5 mg/dl with baseline SCR < 1.5 mg/dl, or SCI ≥ 1.0 mg/dl with baseline ≥ 1.5 mg/dl and < 5.0 mg/dl	Baseline SCR ≥ 5.0 mg/dl, previous HD, kidney transplantation, urinary tract obstruction	Hospital
Demirjian 2011 [24]	Ischemic or nephrotoxic ATN, oliguria, SCR ≥ 2 mg/dl in males or ≥ 1.5 mg/dl in females	Baseline SCR > 2.0 mg/dl in males, > 1.5 mg/dl in females, previous HD, kidney transplant	60-day

AKI: acute kidney injury, RRT: renal replacement therapy, SCR: serum creatinine, SCI: serum creatinine increase, ATN: acute tubular necrosis, CKD: chronic kidney disease, HD: hemodialysis, NR, not reported.

doi:10.1371/journal.pone.0169341.t002

Five studies were conducted in single center. All studies evaluated discrimination with the AUROC and nine studies evaluated calibration with the Hosmer-Lemeshow test.

AUROCs for hospital mortality reported in the original articles (internal validation) and external validation studies are shown in Fig 2. Seven recently published articles for AKI outcome prediction models reported AUROCs for internal validation and all of them had high AUROCs of above 0.7. All prediction models were externally validated by one or more studies. AUROCs in the external validation studies for these scores were generally low (less than 0.7 in most studies). In addition, seven prediction models that were validated both internally and externally had invariably lower AUROCs in external validation than those in internal validation.

Table 5 shows variables included in more than one prediction model and their odds ratios / p values. There were 10 common non-renal variables that were reported in more than three prediction models: mechanical ventilation, age, gender, hypotension, liver failure, oliguria, sepsis/septic shock, low albumin, consciousness and low platelet count. Renal variables (low creatinine and high urea) were often used in the same prediction models.

Discussion

Key findings

We have systematically reviewed AKI outcome prediction models and their external validation studies. We found 12 articles reporting original prediction models for hospital mortality in AKI patients and nine articles assessing external validation of the outcome prediction models.

Table 3. Quality assessment for articles reporting outcome prediction models for acute kidney injury.

	Definition of predictors	Indications for RRT defined	Missing data definition	Bootstrap resampling	Multivariable analysis approach	Event per variable ratio	Internal validation cohort
Ramussen 1985 [14]	Yes	Yes	NR	NR	Stepwise multiple linear regression	7.9 (10/79)	Split sample
Lohr 1988 [15]	Yes	Yes	NR	NR	Stepwise logistic regression	31.3 (3/94)	NR
Schaefer 1991 [16]	Yes	Yes	NR	NR	Stepwise linear discriminant procedure	12.7 (6/76)	NR
Liano 1993 [17]	Yes	Yes	NR	NR	Multiple linear regression	19.3 (9/174)	Cross-validation
Paganini 1996 [18]	Yes	NR	NR	NR	Stepwise logistic regression	43.0 (8/344)	Both
Chertow 1998 [19]	NR	Yes	Yes	Yes	Logistic regression	10.6 (7/74)	NR
Lins 2000 [20]	Yes	NR	NR	NR	Linear regression	20.8 (5/104)	Split sample
Mehta 2002 [12]	Yes	NR	Yes	Yes	Stepwise backward logistic regression	34.9 (9/314)	Split sample
Lins 2004 [21]	Yes	NR	NR	NR	Logistic regression	18.3 (8/146)	Split sample
Dharan 2005 [22]	Yes	NR	NR	Yes	Logistic regression	10.0 (10/100)	Both
Chertow 2006 [23]	Yes	Yes	Yes	NR	Stepwise backward logistic regression	32.7 (7/229)	NR
Demirjian 2011 [24]	Yes	Yes	Yes	Yes	Stepwise backward logistic regression	28.3 (21/595)	Split sample

NR: not recorded.

doi:10.1371/journal.pone.0169341.t003

Although good performance was reported in their internal validation, most of the prediction models had poor discrimination with an AUROC below the threshold of 0.7 in their external validation studies. We also identified 10 common variables that were frequently included in the prediction models.

Relationship to previous studies

The establishment of a clinical prediction model encompasses three consecutive research phases, namely derivation, external validation and impact analysis [35]. In this study, we conducted a systematic review for the first two phases in AKI outcome prediction. Several systematic reviews for clinical prediction models and their external validation have been conducted in other medical conditions, which consistently found methodological limitations. [36–40]. Such limitations include case mix heterogeneity, small sample sizes, insufficient description of study design, and lack of external validation. We found the same limitations in the AKI outcome prediction studies. For example, all prediction models examined in this study were relatively old (data collected more than 10 years ago) and conducted before consensus criteria for AKI were published [41–43]. Therefore, patients included in these prediction models were heterogeneous, with varied RRT requirement and mortality. We also found that more than half of the studies for AKI prediction models and their external validation were single-center (12/21, 57%), and most of them included less than 1,000 patients (19/22, 86%). Furthermore, the moment of data collection for each clinical prediction model and external validation was different. Data collection can be done at admission, at AKI diagnosis, at the start of RRT, at nephrologist consultation, and so on. Demirjian's model for instance, collected variables at

Table 4. Characteristics of external validation studies for acute kidney injury outcome prediction models.

	Type of study	Study period	Centers, Number	Sample size	Mean age, years	Hospital mortality	Validated models	Discrimination	Calibration
Douma 1997 [11]	Retrospective	1985–1993	1	238	61	76%	4	AUROC	H-L
Lins 2002 [21]	Prospective	1996–1997	1	197	70	53%	1	AUROC	NR
Martin 2002 [29]	Retrospective	1995–1996	1	349	58	59%	2	AUROC	NR
Mehta 2002 [12]	Prospective	1989–1995	4	605	56	52%	7	AUROC	H-L
d'Avila DO, 2004 [30]	Prospective	NR	1	280	54	85%	1	AUROC	H-L
Dharan 2005 [22]	Prospective	2002	1	265	48	38%	1	AUROC	H-L
Uchino 2005 [26]	Prospective	2000–2001	54	1,742	67	61%	4	AUROC	H-L
Lima 2005 [31]	Prospective	2000–2001	1	342	60	85%	1	AUROC	H-L
Chertow 2006 [23]	Prospective	1999–2001	5	618	59	37%	4	AUROC	NR
Kolhe 2008 [25]	Retrospective	1995–2004	170	17,326	63	59%	3	AUROC	H-L
Lin 2008 [32]	Retrospective	2002–2005	4	334	64	66%	4	AUROC	NR
Costa e Silva VT 2009 [33]	Prospective	2003–2005	1	366	NR	68%	3	AUROC	H-L
Demirjian 2011 [24]	Post hoc of RCT	2003–2007	27	1,122	60	50%	1	AUROC	NR
Ohnuma 2015 [34]	Retrospective	2010	14	343	69	59%	6	AUROC	H-L

RCT: randomized controlled trial, NR: not recorded, AUROC: area under the receiver operating characteristic curve, H-L: Hosmer-Lemeshow.

doi:10.1371/journal.pone.0169341.t004

RRT start [24], while other models collected variables at nephrologist consultation [12, 23], or at AKI diagnosis [21, 23]. This variable is also important for external validation, as the discrimination AUROC value can be altered if variables are collected at different moments in the new cohort. Considering the poor generalizability of currently available prediction models (AUROCs lower than 0.7 in most external validation studies), a large database collected in multicenter using consensus AKI criteria will be needed both to derive and validate AKI outcome prediction models.

Among the prediction models included in this systematic review, we found that the Liano's score [17] was the most often evaluated externally (11 studies). The range of AUROC validated externally for the Liano's score was from 0.55 to 0.90, and four of them were above 0.7. The reason why Liano showed high AUROCs in some external validation studies is unclear. It might be partially explained by that the Liano' score contained several risk factors that are frequently used in the prediction models (mechanical ventilation, age, gender, hypotension, liver failure, oliguria, consciousness disturbance), although Dharan also included nine variables, with poor discrimination by one external validation study (Table 5).

Significance and implications

To derive an accurate prediction model, choosing appropriate candidate predictors is of much importance. Previous studies have shown that clinical intuition may not be suitable for
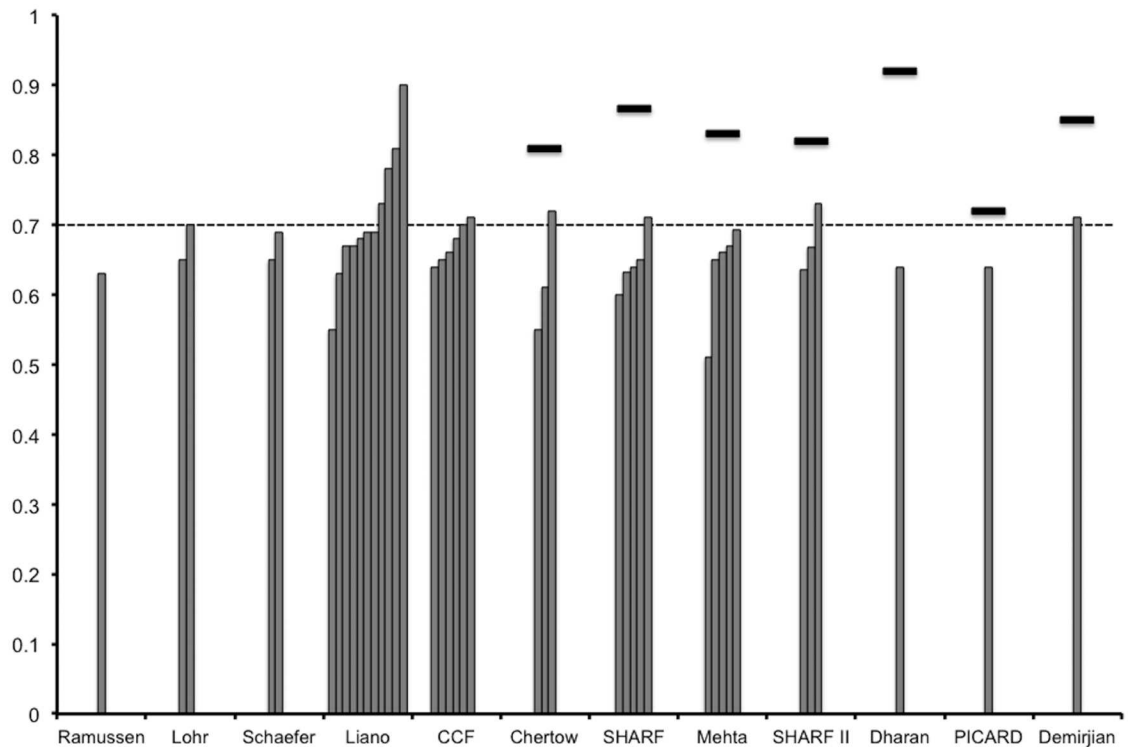
Fig 2. Area under the receiver operating characteristic curves (AUROC) for hospital mortality reported in the original articles and external validation studies. Black horizontal bars: AUROC in original studies, gray columns: AUROC in external validation studies.

identifying candidate predictors [44]. A better approach is to combine a systematic literature review of prognostic factors associated with the outcome of interest with opinions of field experts [35]. We identified 10 common variables that were frequently included in the prediction models. These variables are also often found to be related to mortality in more recent epidemiological studies using consensus AKI criteria [45–48]. We believe that our study results will be useful for future studies to derive accurate AKI outcome prediction models by including these variables for data collection.

Although often included in the prediction models, we think that including both low creatinine and high urea concentrations as independent variables can be problematic (Table 5). Low serum creatinine is included in general severity scores as one of independent variables [5]. Serum urea has been used as a marker of timing of starting RRT in several studies, which showed that patients with higher urea at start of RRT had worse outcome than patients with lower urea [49]. High urea is also included in general severity scores [5]. However, serum creatinine and urea concentrations clearly have strong co-linearity. In AKI patients, urea is almost always high when creatinine is high. Even if both variables are found to be independent variables in multivariable analysis, it seems unlikely that including both variables in a prediction model will improve prediction ability [50].

Physicians are faced with the impractical situation of having to choose among many concurrent outcome prediction models for AKI. To overcome this issue, it is recommended that investigators who have large data sets should conduct external validation studies of multiple existing models at once, in order to determine which model is most useful [51]. We believe

Table 5. Variables included in more than one prediction model and their odds ratios / p values.

	Number of studies	Ramussen	Lohr	Schaefer	Liano	Paganini	Chertow1998	Lins2000	Mehta	Lins2004	Dharan	Chertow2006	Demirjian
Mechanical ventilation	9	-	NR	P<0.01	P<0.01	NR	2.95	10.2	-	NR	2.3	-	P<0.01
Age	7	-	-	-	P<0.01	-	-	1.70	1.02	NR	1.0	1.16	P<0.01
Gender	5	-	-	-	P<0.05	NR	3.70	-	2.36	-	0.6	-	-
Hypotension	5	-	NR	P<0.01	P<0.05	-	-	-	-	NR	3.1	-	-
Liver failure	5	-	-	-	P<0.05	NR	-	-	3.06	-	3.7	1.76	-
Oliguria	5	NR	-	-	P<0.05	-	4.39	-	-	-	4.9	-	P = 0.02
Sepsis/septic shock	5	-	NR	NS	-	-	-	-	-	NR	2.2	1.69	-
Low albumin	4	-	-	-	-	-	-	1.50	-	NR	1.7	-	P = 0.02
Consciousness	4	NR	-	-	P<0.05	-	7.35	-	-	-	10.4	-	-
Low platelet count	4	-	-	-	-	NR	-	-	3.40	-	-	2.10	P<0.01
Heart failure	3	-	NR	-	-	-	-	1.88	-	NR	-	-	-
Preexisting heart disease	3	NR	-	NS	-	-	-	-	-	-	-	-	P = 0.03
Prothrombin time	3	-	-	-	-	-	-	1.29	-	NR	-	-	P = 0.01
Respiratory failure	3	NR	-	-	-	-	-	-	2.62	-	-	1.84	-
Surgical patients	3	NR	-	-	-	NR	-	-	-	-	-	-	P = 0.08
High bilirubin	2	-	-	-	-	-	-	-	-	NR	-	-	P<0.01
Acute cardiac illness	2	NR	-	-	-	-	5.90	-	-	-	-	-	-
Bicarbonate	2	-	-	-	-	-	0.93	-	-	-	-	-	P = 0.02
Malignancy	2	NR	-	-	-	-	-	-	-	-	-	-	P<0.01
Low creatinine	4	-	-	-	-	NR	-	-	0.71	-	-	0.81	P = 0.01
High urea	3	-	-	-	-	NR	-	-	1.02	-	-	1.09	-

NR: not reported, NS: not significant.

doi:10.1371/journal.pone.0169341.t005

that our study results will also be useful for future studies by providing the list of published outcome prediction models for AKI.

Strengths and limitations

The strength of our study is that, to the best of our knowledge, this is the first systematic review on AKI outcome prediction models in the medical literature. We have reviewed studies for both prediction models and their external validation, and provided potential candidate variables for future prediction models and the list of published prediction models for future external validation studies.

However, our study also contains several limitations. First, recent studies suggest that AKI biomarkers might be useful to predict outcome and could be combined with physiological and laboratory variables to improve predicting ability [52, 53]. However, prediction models should include only variables that are available at the time when the model is intended to be used, and biomarkers are not yet widely used clinically [54]. Second, we excluded six studies due to discrimination results not available [55–60]. However, these studies were generally old, small, and of poor methodological quality. We believe that including these studies would not change our main findings. Finally, the AKI definitions used in both prediction models and their external validation studies are outdated, and studies included were relatively old (the most recently published study is from 2011 and the data were collected between 2003 and 2007). There is an urgent need for a mortality prediction model based on current definitions of AKI, and this systematic review can be considered a first step to accomplish this task.

Conclusions

Multiple outcome prediction models for AKI have been derived previously. These scores had good performance in their internal validation studies, while poor performance was reported in their external validation, suggesting that there is no accurate model currently available. To generate accurate AKI prediction models, several recommendations can be provided: using a large database collected in multicenter, applying consensus AKI criteria, and collecting variables frequently used in previous models (mechanical ventilation, age, gender, hypotension, liver failure, oliguria, sepsis/septic shock, low albumin, consciousness and low platelet count). Information in this systematic review should be useful both for future prediction model derivation by providing potential candidate predictors, and for future external validation by listing up the published prediction models.

Supporting Information

S1 PRISMA Checklist.
(DOCX)

S1 File. Excluded Articles.
(DOCX)

Author Contributions

Conceptualization: SU.

Data curation: TO.

Formal analysis: TO.

Investigation: TO.

Methodology: SU.

Project administration: SU.

Resources: TO.

Software: TO.

Supervision: SU.

Validation: TO.

Visualization: TO.

Writing – original draft: TO.

Writing – review & editing: SU.

References

1. Bagshaw SM, George C, Dinu I, Bellomo R. A multi-centre evaluation of the RIFLE criteria for early acute kidney injury in critically ill patients. *Nephrol Dial Transplant* 2008; 23: 1203–1210. doi: 10.1093/ndt/gfm744 PMID: 17962378

2. Uchino S, Kellum JA, Bellomo R, Doig GS, Morimatsu H, Morgera S et al. Acute renal failure in critically ill patients: a multinational, multicenter study. *JAMA* 2005; 294: 813–818. doi: 10.1001/jama.294.7.813 PMID: 16106006

3. Liangos O, Wald R, O'Bell JW, Price L, Pereira BJ, Jaber BL. Epidemiology and outcomes of acute renal failure in hospitalized patients: a national survey. *Clin J Am Soc Nephrol* 2006; 1: 43–51. doi: 10.2215/CJN.00220605 PMID: 17699189

4. Nash K, Hafeez A, Hou S. Hospital-acquired renal insufficiency. *Am J Kidney Dis* 2002; 39: 930–936. doi: 10.1053/ajkd.2002.32766 PMID: 11979336

5. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med* 1985; 13: 818–829. PMID: 3928249

6. Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, et al. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 1991; 100: 1619–1636. PMID: 1959406

7. Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute Physiology and Chronic Health Evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. *Crit Care Med* 2006; 34: 1297–1310. doi: 10.1097/01.CCM.0000215112.84523.F0 PMID: 16540951

8. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA* 1993; 270: 2957–2963. PMID: 8254858

9. Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P, Campos RA, et al. SAPS 3—From evaluation of the patient to evaluation of the intensive care unit. Part 2: Development of a prognostic model for hospital mortality at ICU admission. *Intensive Care Med* 2005; 31: 1345–1355. doi: 10.1007/s00134-005-2763-5 PMID: 16132892

10. Higgins TL, Teres D, Copes WS, Nathanson BH, Stark M, Kramer AA. Assessing contemporary intensive care unit outcome: an updated Mortality Probability Admission Model (MPM0-III). *Crit Care Med* 2007; 35: 827–835. doi: 10.1097/01.CCM.0000257337.63529.9F PMID: 17255863

11. Douma CE, Redekop WK, van der Meulen JH, van Olden RW, Haeck J, Struijk DG et al. Predicting mortality in intensive care patients with acute renal failure treated with dialysis. *J Am Soc Nephrol* 1997; 8: 111–117. PMID: 9013455

12. Mehta RL, Pascual MT, Gruta CG, Zhuang S, Chertow GM. Refining predictive models in critically ill patients with acute renal failure. *J Am Soc Nephrol* 2002; 13: 1350–1357. PMID: 11961023

13. Costa e Silva VT, de Castro I, Liano F, Muriel A, Rodriguez-Palomares JR, Yu L. Performance of the third-generation models of severity scoring systems (APACHE IV, SAPS 3 and MPM-III) in acute kidney injury critically ill patients. *Nephrol Dial Transplant* 2011; 26: 3894–3901. doi: 10.1093/ndt/gfr201 PMID: 21505093

14. Rasmussen HH, Pitt EA, Ibels LS, McNeil DR. Prediction of outcome in acute renal failure by discriminant analysis of clinical variables. *Arch Intern Med* 1985; 145: 2015–2018. PMID: 4062452

15. Lohr JW, McFarlane MJ, Grantham JJ. A clinical index to predict survival in acute renal failure patients requiring dialysis. *Am J Kidney Dis* 1988; 11: 254–259. PMID: 3344747

16. Schaefer JH, Jochimsen F, Keller F, Wegscheider K, Distler A. Outcome prediction of acute renal failure in medical intensive care. *Intensive Care Med* 1991; 17: 19–24. PMID: 1903797

17. Liano F, Gallego A, Pascual J, García-Martín F, Teruel JL, Marcén R et al. Prognosis of acute tubular necrosis: an extended prospectively contrasted study. *Nephron* 1993; 63: 21–31. PMID: 8446248

18. Paganini EP, Halstenberg WK, Goormastic M. Risk modeling in acute renal failure requiring dialysis: the introduction of a new model. *Clin Nephrol* 1996; 46: 206–211. PMID: 8879857

19. Chertow GM, Lazarus JM, Paganini EP, Allgren RL, Lafayette RA, Sayegh MH. Predictors of mortality and the provision of dialysis in patients with acute tubular necrosis. The Auriculin Anaritide Acute Renal Failure Study Group. *J Am Soc Nephrol* 1998; 9: 692–698. PMID: 9555672

20. Lins RL, Elseviers M, Daelemans R, Zachée P, Zachée P, Gheuens E et al. Prognostic value of a new scoring system for hospital mortality in acute renal failure. *Clin Nephrol* 2000; 53: 10–17. PMID: 10661477

21. Lins RL, Elseviers MM, Daelemans R, Arnouts P, Billiouw JM, Couttenye M et al. Re-evaluation and modification of the Stuivenberg Hospital Acute Renal Failure (SHARF) scoring system for the prognosis of acute renal failure: an independent multicentre, prospective study. *Nephrol Dial Transplant* 2004; 19: 2282–2288. doi: 10.1093/ndt/gfh364 PMID: 15266030

22. Dharan KS, John GT, Antonisamy B, Kirubakaran MG, Jacob CK. Prediction of mortality in acute renal failure in the tropics. *Ren Fail* 2005; 27: 289–296. PMID: 15957545

23. Chertow GM, Soroko SH, Paganini EP, Cho KC, Himmelfarb J, Ikizler TA et al. Mortality after acute renal failure: models for prognostic stratification and risk adjustment. *Kidney Int* 2006; 70: 1120–1126. doi: 10.1038/sj.ki.5001579 PMID: 16850028

24. Demirjian S, Chertow GM, Zhang JH, O'Connor TZ, Vitale J, Paganini EP et al. Model to predict mortality in critically ill adults with acute kidney injury. *Clin J Am Soc Nephrol* 2011; 6: 2114–2120. doi: 10.2215/CJN.02900311 PMID: 21896828

25. Kolhe NV, Stevens PE, Crowe AV, Lipkin GW, Harrison DA. Case mix, outcome and activity for patients with severe acute kidney injury during the first 24 hours after admission to an adult, general critical care unit: application of predictive models from a secondary analysis of the ICNARC Case Mix Programme database. *Crit Care* 2008; 12 Suppl 1: S2.

26. Uchino S, Bellomo R, Morimatsu H, Morgera S, Schetz M, Tan I et al. External validation of severity scoring systems for acute renal failure using a multinational database. *Crit Care Med* 2005; 33: 1961–1967. PMID: 16148466

27. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143: 29–36. doi: 10.1148/radiology.143.1.7063747 PMID: 7063747

28. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009; 151: 264–269. PMID: 19622511

29. Martin C, Saran R, Leavey S, Swartz R. Predicting the outcome of renal replacement therapy in severe acute renal failure. ASAIO J 2002; 48: 640–644. PMID: 12455775

30. d'Avila DO, Cendoroglo Neto M, dos Santos OF, Schor N, Poli de Figueiredo CE. Acute renal failure needing dialysis in the intensive care unit and prognostic scores. Ren Fail 2004; 26: 59–68. PMID: 15083924

31. Lima EQ, Dirce MT, Castro I, Yu L. Mortality risk factors and validation of severity scoring systems in critically ill patients with acute renal failure. Ren Fail 2005; 27: 547–556. PMID: 16152992

32. Lin YF, Ko WJ, Wu VC, Chen YS, Chen YM, Hu FC et al. A modified sequential organ failure assessment score to predict hospital mortality of postoperative acute renal failure patients requiring renal replacement therapy. Blood Purif 2008; 26: 547–554. doi: 10.1159/000178771 PMID: 19052448

33. Costa e Silva VT, de Castro I, Liano F, Muriel A, Rodriguez-Palomares JR, Yu L. Sequential evaluation of prognostic models in the early diagnosis of acute kidney injury in the intensive care unit. Kidney Int 2009; 75: 982–986. doi: 10.1038/ki.2009.3 PMID: 19212423

34. Ohnuma T, Uchino S, Toki N, Takeda K, Namba Y, Katayama S et al.; JSEPTIC (Japanese Society for Physicians and Trainees in Intensive Care) Clinical Trial Group. External Validation for Acute Kidney Injury Severity Scores: A Multicenter Retrospective Study in 14 Japanese ICUs. Am J Nephrol. 2015; 42: 57–64. doi: 10.1159/000439118 PMID: 26337793

35. Labarère J, Renaud B, Fine MJ. How to derive and validate clinical prediction models for use in intensive care medicine. Intensive Care Med. 2014; 40: 513–27. doi: 10.1007/s00134-014-3227-6 PMID: 24570265

36. Ettema RG, Peelen LM, Schuurmans MJ, Nierich AP, Kalkman CJ, Moons KG. Prediction models for prolonged intensive care unit stay after cardiac surgery: systematic review and validation study. Circulation. 2010; 122: 682–9 doi: 10.1161/CIRCULATIONAHA.109.926808 PMID: 20679549

37. Wlodzimirow KA, Eslami S, Chamuleau RA, Nieuwoudt M, Abu-Hanna A. Prediction of poor outcome in patients with acute liver failure-systematic review of prediction models. PLoS One. 2012; 7: e50952. doi: 10.1371/journal.pone.0050952 PMID: 23272081

38. Jaja BN, Cusimano MD, Etminan N, Hanggi D, Hasan D, Ilodigwe D et al. Clinical prediction models for aneurysmal subarachnoid hemorrhage: a systematic review. Neurocrit Care. 2013; 18: 143–53. doi: 10.1007/s12028-012-9792-z PMID: 23138544

39. Warnell I, Chincholkar M, Eccles M. Predicting perioperative mortality after oesophagectomy: a systematic review of performance and methods of multivariate models. Br J Anaesth. 2015; 114: 32–43. doi: 10.1093/bja/aeu294 PMID: 25231768

40. Silver SA, Shah PM, Chertow GM, Harel S, Wald R, Harel Z. Risk prediction models for contrast induced nephropathy: systematic review. BMJ. 2015 27; 351: h4395. doi: 10.1136/bmj.h4395 PMID: 26316642

41. Bellomo R, Ronco C, Kellum JA, Mehta RL, Palevsky P; Acute Dialysis Quality Initiative workgroup. Acute renal failure—definition, outcome measures, animal models, fluid therapy and information technology needs: the Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group. Crit Care, 8: R204–12, 2004. doi: 10.1186/cc2872 PMID: 15312219

42. Mehta RL, Kellum JA, Shah SV, Molitoris BA, Ronco C, Warnock DG et al.: Acute Kidney Injury Network: report of an initiative to improve outcomes in acute kidney injury. Crit Care, 11: R31, 2007. doi: 10.1186/cc5713 PMID: 17331245

43. Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group: KDIGO Clinical Practice Guideline for Acute Kidney Injury. Kidney Int Suppl., 2: 1–138, 2012.

44. Randolph AG, Guyatt GH, Calvin JE, Doig G, Richardson WS. Understanding articles describing clinical prediction tools. Evidence Based Medicine in Critical Care Group. Crit Care Med. 1998; 26: 1603–12. PMID: 9751601

45. Nisula S, Kaukonen KM, Vaara ST, Korhonen AM, Poukkanen M, Karlsson S, et al.; FINNAKI Study Group. Incidence, risk factors and 90-day mortality of patients with acute kidney injury in Finnish intensive care units: the FINNAKI study. Intensive Care Med. 2013; 39: 420–8. doi: 10.1007/s00134-012-2796-5 PMID: 23291734

46. Bouchard J, Acharya A, Cerda J, Maccariello ER, Madarasu RC, Tolwani AJ, et al. A Prospective International Multicenter Study of AKI in the Intensive Care Unit. Clin J Am Soc Nephrol. 2015; 10: 1324–31. doi: 10.2215/CJN.04360514 PMID: 26195505

47. Xu X, Nie S, Liu Z, Chen C, Xu G, Zha Y, et al. Epidemiology and Clinical Correlates of AKI in Chinese Hospitalized Adults. Clin J Am Soc Nephrol. 2015; 10: 1510–8. doi: 10.2215/CJN.02140215 PMID: 26231194

48. Hoste EA, Bagshaw SM, Bellomo R, Cely CM, Colman R, Cruz DN, et al. Epidemiology of acute kidney injury in critically ill patients: the multinational AKI-EPI study. Intensive Care Med. 2015; 41: 1411–23. doi: 10.1007/s00134-015-3934-7 PMID: 26162677

49. Gettings LG, Reynolds HN, Scalea T. Outcome in post-traumatic acute renal failure when continuous renal replacement therapy is applied early vs. late. Intensive Care Med. 1999; 25: 805–13. PMID: 10447537

50. Uchino S. Outcome prediction for patients with acute kidney injury. Nephron Clin Pract. 2008; 109: c217–23. doi: 10.1159/000142931 PMID: 18802370

51. Collins GS, Moons KG. Comparing risk prediction models. BMJ. 2012; 344: e3186. doi: 10.1136/bmj.e3186 PMID: 22628131

52. McIlroy DR, Farkas D, Matto M, Lee HT. Neutrophil gelatinase-associated lipocalin combined with delta serum creatinine provides early risk stratification for adverse outcomes after cardiac surgery: a prospective observational study. Crit Care Med. 2015; 43: 1043–52. doi: 10.1097/CCM.0000000000000927 PMID: 25768681

53. Pike F, Murugan R, Keener C, Palevsky PM, Vijayan A, Unruh M et al.; Biological Markers for Recovery of Kidney (BioMaRK) Study Investigators. Biomarker Enhanced Risk Prediction for Adverse Outcomes in Critically Ill Patients Receiving RRT. Clin J Am Soc Nephrol. 2015; 10: 1332–9. doi: 10.2215/CJN.09911014 PMID: 26048891

54. Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? BMJ. 2009; 338: b375. doi: 10.1136/bmj.b375 PMID: 19237405

55. Cioffi WG, Ashikaga T, Gamelli RL. Probability of surviving postoperative acute renal failure. Development of a prognostic index. Ann Surg 1984; 200: 205–11. PMID: 6465976

56. Lien J, Chan V. Risk factors influencing survival in acute renal failure treated by hemodialysis. Arch Intern Med 1985; 145: 2067–9. PMID: 4062459

57. Corwin HL, Teplick RS, Schreiber MJ, Fang LS, Bonventre JV, Coggins CH. Prediction of outcome in acute renal failure. Am J Nephrol 1987; 7:8–12. PMID: 3578381

58. Barton IK, Hilton PJ, Taub NA, Warburton FG, Swan AV, Dwight J et al. Acute renal failure treated by haemofiltration: factors affecting outcome. Q J Med 1993; 86: 81–90. PMID: 8464996

59. Chertow GM, Christiansen CL, Cleary PD, Munro C, Lazarus JM. Prognostic stratification in critically ill patients with acute renal failure requiring dialysis. Arch Intern Med 1995; 155: 1505–11. PMID: 7605152

60. Radovic M, Ostric V, Djukanovic L. Validity of prediction scores in acute renal failure due to polytrauma. Ren Fail 1996; 18: 615–20. PMID: 8875687