BMJ Open

A systematic review of prognostic prediction models for acute kidney injury (AKI) in general hospital populations.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-016591
Article Type:	Research
Date Submitted by the Author:	23-Feb-2017
Complete List of Authors:	Hodgson, Luke; University of Southampton, Faculty of Medicine; Western Sussex Hospitals NHS Trust, Intensive Care Sarnowski, Alexander; Royal Surrey County Hospital NHS Foundation Trust Roderick, Paul; University of Southampton, Faculty of Medicine Dimitrov, Borislav; University of Southampton, Primary Care and Population Sciences Venn, Richard; Western Sussex Hospitals NHS Trust, Intensive Care Forni, Lui; University of Surrey Faculty of Health and Medical Sciences; Royal Surrey County Hospital NHS Foundation Trust
Primary Subject Heading:	Renal medicine
Secondary Subject Heading:	Research methods
Keywords:	Nephrology < INTERNAL MEDICINE, Acute renal failure < NEPHROLOGY, STATISTICS & RESEARCH METHODS

Title Page

Title: A systematic review of prognostic prediction models for acute kidney injury (AKI) in general hospital populations.

Corresponding author: Dr Luke E Hodgson, Academic Unit of Primary Care and Population Sciences, Faculty of Medicine, University of Southampton, Southampton General Hospital, Tremona Road, Southampton, SO16 6YD.

Email: drlhodgson@gmail.com Telephone: +44 7734883040

Authors:

Luke E Hodgson,[1,2] Intensive Care Registrar,

Alexander Sarnowski[3], Core Medical Trainee,

Paul J Roderick[1], Professor of Public Health,

Borislav D Dimitrov[1], Associate Professor Medical Statistics†,

Richard M Venn[2], Consultant Intensivist and,

Lui G Forni[3,4], Professor of Intensive Care Medicine.

1. Academic Unit of Primary Care and Population Sciences, Faculty of Medicine, University of Southampton, Southampton General Hospital, Tremona Road, Southampton, SO16 6YD, UK.
2. Western Sussex Hospitals NHS Foundation Trust, Anaesthetics Department, Worthing Hospital, Lyndhurst Rd, Worthing, BN11 2DH, UK.
3. The Royal Surrey County Hospital NHS Foundation Trust, Egerton Road, Guildford, GU2 7XX, UK.
4. Faculty of Health and Medical Sciences, University of Surrey, Guildford, Surrey, GU2 7XH, UK.

† Deceased 15th January 2017.

Word count main text: 2715

Abstract

Objective Critically appraise prediction models for hospital-acquired AKI (HA-AKI) in general populations.

Design Systematic review.

Data sources Medline, EMBASE & Web of Science until November 2016.

Eligibility for study selection Studies describing development of a multivariable model for predicting HA-AKI in non-specialised adult hospital populations. Published guidance for systematic reviews and reporting followed for data extraction and appraisal.

Results 14046 references were screened. Of 53 HA-AKI prediction models, 11 met inclusion criteria (general medicine and/or surgery populations, 474478 patient episodes), five with external validation. The most common predictors were chronic kidney disease (n=10 models), age (n=9), diabetes (n=5), drugs (diuretics n=4 and/or Angiotensin-converting enzyme inhibitors/Angiotensin-receptor blockers n=3), serum bicarbonate and heart failure (4 models each). Substantial heterogeneity was identified between studies for outcome definition (timing and marker). Deficiencies in recommended reporting included handling of candidate predictors and missing data, blinding of outcome assessment and sample size considerations. Area under the receiver operating characteristic curves to predict HA-AKI ranged 0.71-0.80 in derivation (reported in 8/11 studies), 0.66-0.80 for internal validation studies (n=7) and 0.65-0.71 in 5 external validations. For calibration the Hosmer-Lemeshow test or a calibration plot were provided in only 4/11 derivations, 3/11 internal and 3/5 external

validations. A minority of the models allow easy bedside calculation and potential electronic automation. No impact analysis studies were found.

Conclusions AKI prediction models may help address shortcomings in risk assessment, however, in general hospital populations few have external validation. Similar candidate predictors reflect an elderly demographic with chronic co-morbidities. Reporting deficiencies mirror prediction research more broadly. Future research should focus on validation, impact analysis and potential electronic linkage between primary and secondary care. An impact analysis could combine a prediction model with AKI alerting to address prevention and early recognition of evolving AKI.

Key words: acute kidney injury, clinical prediction models, systematic review

Summary

Strengths

- This is the first systematic review of prediction models for hospital-acquired AKI (HA-AKI) in general hospital populations who account for the majority of hospital admissions and AKI cases.
- The models were selected following an extensive literature search and critical appraisal guidance.
- The large number of patient episodes provides important insights into AKI prediction and complements other recent reviews in specialised areas (cardiac surgery, CI-AKI and liver transplantation).

Weaknesses

- Lack of access to individual participant data (IPD) prevented a meta-analysis of the studies, an avenue of future research.
- The small number of externally validated models and absence of impact analysis limits recommendation of an individual model.

Introduction

Acute kidney injury (AKI) is defined as an acute increase in serum creatinine (SCr) or reduction in urine volume.[1] The incidence of AKI is increasing, affecting up to one in five hospitalised adults worldwide.[2] A continuum of injury exists long before sufficient loss of excretory kidney function can be measured with standard laboratory tests (i.e. SCr).[3 4] Associated mortality remains high, in part reflecting the severity of the underlying disease, but may also be due to the limitations of conventional markers to detect early injury.[5]

Deficits in recognition and management of patients with AKI,[6] has led to practice guidance calling for improved risk assessment, at which point interventions could be most beneficial.[7] One suggested strategy to achieve this aim is through the implementation of clinical prediction models.[8 9] Though development and validation of AKI prediction models is desirable,[7 10] clinical application in this and other fields has been hampered for a number of reasons:

- Potential predictors and models continuously increase with new studies often finding conflicting results,[11]
- Substandard reporting of methodology and results make conclusions problematic,[12 13]
- Few general hospital population studies exist - specialist fields (cardiac and transplant surgery and contrast-induced [CI-AKI]) account for the majority of AKI models and all systematic reviews, but are unlikely to be generalisable and,[14-17]
- Models rarely enable electronic automation as part of clinical workflow, known to influence uptake.[18]

High quality systematic reviews of prediction models have been called for.[19] Following recent reporting guidance (CHARMS[20] and TRIPOD),[12] this review appraises hospital-acquired AKI (HA-AKI) prediction models in general populations, who in the UK account for the majority of hospital episodes,[21] and AKI cases.[22 23]

Methods

Published guidance helped frame the review question, data extraction, reporting and appraisal.[12 20] The research question was: what are the available prognostic prediction models for the development of HA-AKI in adult general populations? Using explicit, systematic methods to minimise bias and provide reliable findings from which conclusions can be drawn and decisions made,[24 25] the review aimed to collate empirical evidence for AKI prediction models across general hospital settings, fitting pre-specified eligibility criteria (online supplementary file eTable 1). Performance was assessed by discrimination and calibration including validation studies. The presence of any impact analysis studies was also investigated. The review aimed to provide recommendations for the most robust, usable models, including the ability to incorporate future electronic data linkage, for example between the community (primary care) and hospital.

Data sources, study selection and data extraction

We searched MEDLINE, Embase & Web of Science databases (inception to November 2016) using recommended filters (online supplementary file eTables 2-4).[26 27] Titles and abstracts were screened by two reviewers (LH, AS) and full articles reviewed if eligible. Disagreements were resolved by iterative screening rounds. Reference lists from retrieved articles, systematic reviews, National[7] and International guidance[1] and our own literature files were also analysed. Data extraction, and quality assessment was performed by two investigators (LH, AS) with disagreements resolved by a third reviewer (LF). A data extraction form was used based on previous reviews and

guidance (summary online supplementary file eTable 5).[12 13 20] Items extracted included design (eg, cohort, case-control), population, location, outcome (definition duration of follow-up, blinding of assessment), modelling method (eg, logistic), method of internal validation (eg, bootstrapping), number of participants and events, number and type of predictors, model presentation, and predictive performance (calibration, discrimination). Finally, the presence of external validation was recorded. As no tool is currently available in the field, a formal risk of bias assessment was not explicitly performed.

Outcome, model performance and clinical utility

It was anticipated that study outcome, HA-AKI, would vary given the numerous definitions in use prior to KDIGO in 2012.[1] Discrimination and calibration are the most common methods to assess model performance. Discrimination is usually assessed graphically by the area under the receiver operating characteristic curve (AUROC), representing how well a model separates and ranks patients who experienced the outcome, from those who did not. For prediction models, the AUROC, which focuses solely on accuracy has a number of shortcomings, such as a lack of information on consequences and when used in populations where the outcome prevalence is rare.[28 29] Calibration describes how well predicted results agree with observed results.[12 29] The Hosmer-Lemeshow (H-L) test, despite limitations, is the most commonly used calibration statistic.[30 31] It is also recommended to graphically plot expected and actual outcomes, for example, with a calibration slope.[12] In addition to performance, ease of bedside use and whether the models could be electronically automated - factors known to

influence successful uptake - were recorded.[18] A quantitative synthesis of the models was not performed, being beyond the scope of review and formal methods for meta-analysis of prediction models are yet to be fully developed.

Reporting quality assessment

A global TRIPOD score for each study was calculated to gauge methodological reporting quality, consisting of the sum of the scores for each individual item (out of a maximum 37, with a score of 1 for criterion met, score of 0 for each item not met, or unclear).[12] As yet there has been no suggested cut-off for what represents a high quality study, though it would be reasonable to judge that those studies with the most significant gaps in reporting are those at higher risk of bias.

Patient involvement

Patients were not involved in setting the research question, outcome, design and implementation of the study. There are no plans to involve patients in dissemination.

Results

From 14046 articles identified by the search strategy, 254 full articles were reviewed (PRISMA flow chart, figure 1). Specialised fields (predominantly cardiac surgery, transplantation or CI-AKI) accounted for 61 of 74 (82%) of all studies. This review included eleven general model studies (n=474478 patient episodes), in General Surgery,[32 33] Trauma and Orthopaedics (T&O),[34] General Hospital cohorts (predominantly Medicine and Surgery),[35-39] and Heart Failure (summarised in table 1 and online supplementary file eTable 6).[40-42] Two further studies were purely external validations.[43 44] HA-AKI incidence was 7% (21641 events), though this varied from <1% in the General Surgery models,[32 33] to 28% across the Heart Failure studies and heterogeneous definitions (timeframe and marker) were employed. Mortality was significantly higher in those who developed the outcome in the six studies where data was available (ranging 6-42%). No impact analyses were retrieved.

Study reporting

A median 28 (interquartile range 25-30) of 37 recommended items were reported, suggesting significant shortcomings, increasing the risk of bias (reporting summarised in online supplementary file eTable 8). By design, eight studies were retrospective, two were prospective and one was a case control. Five studies were single-centre. The USA (n=6) and UK (n=3) accounted for the majority of the models. Two studies used imputation techniques for missing data. Definitions were heterogenous (table 1) with five using RIFLE,[36] AKIN,[41] or KDIGO criteria for changes in SCr.[34 35 38] One

study used KDIGO SCr change within a 24-hour timeframe of predictors being measured.[39]

Candidate predictors, model building and sample size

A median of 29 (interquartile range 19-35) predictors were considered, though frequently studies only reported those significant on univariate or multivariate analysis. Blinding of assessment of predictors and study outcome was not mentioned. Continuous predictors were dichotomized in two studies and ten studies used univariate analysis to select for multivariate analysis. No models mentioned shrinkage techniques or sample size calculations. Median number of outcome events was 271 (121-672). For statistical power, all of the studies had more than ten events per variable (EPV) included in the model. However the EPV was <10 in six studies, when accounting for the total number of candidate predictors assessed.[32 35 37 38 40 42] Of a total of 56 different predictors a median of 7 (7-12) were included per model, including demographics, past history, procedure information, laboratory parameters, observations and hospital admission diagnoses (most common presented in figure 2, full details online eTables 9-10).

Model performance (table 1)

Median AUROC (or C-Statistic) was 0.74 (range 0.67-0.80) for derivation and 0.75 (range 0.66-0.80) for internal validations reporting discrimination (seven studies). Only one model study presented a calibration plot for derivation and validation.[34] The H-L statistic was used in three derivations,[35 36 41] and two internal validations.[38 41] Five models have been externally validated on

separate populations within the same study,[34 38] other model studies,[42] or stand alone external validations,[32 44] with moderate AUROCs ranging 0.65-0.71. One validation provided a calibration plot,[34] one the H-L statistic,[38] and one reported both.[44] In the Bell external validation cohort calibration suggested the model over-predicted the outcome requiring recalibration.[34] In the external validation of the Forni study calibration plots showed agreement at low probability rates whilst at higher rates calibration deviated in the medical cohort.[44] Two of the three surgical models have been externally validated: the Kheterpal model,[32] in a Chinese population (AUROC 0.66),[43] and the UK T&O study used a third centre for external validation.[34] Two of five mixed general population models have external validation,[35 38] the later having been derived on medical patients and externally validated in medical and surgical cohorts.[44] The first of the three heart failure studies was externally validated in the subsequent studies with inferior discrimination (AUROC 0.65 in both validations).[40-42] No model updating was reported.

Discussion

Principal findings

In this first systematic review of HA-AKI prediction in general hospital settings, the most common predictors were age, CKD, diabetes, drugs, heart failure and serum bicarbonate. Modest discrimination performance is unsurprising when attempting at a single time point to predict a future event reflecting diverse aetiologies, affecting heterogeneous patient groups. Significant shortcomings mirror those described elsewhere:[13 45-47]

- Multiple similar models, rarely externally validated,
- No impact analysis or evidence of clinical implementation,
- Incomplete reporting and,
- Little consideration of electronic automation (allowing presentation without additional data input beyond usual clinical care), which influences uptake.[18]

Methodological and reporting shortcomings in the studies are summarised in table 2. Published after TRIPOD, Bell and colleagues' model provides researchers with a good template for adherence to guidance and demonstrates the utility of data linkage (for example between community and hospital), though lack of validation in other populations tempers recommendation for implementation.[34]

Strengths and limitations of this review

This review summarises the currently available AKI prediction models in

general populations who account for the majority of hospital admissions and AKI cases.[21-23] The models were selected following an extensive literature search and critical appraisal guidance.[12 20] The large number of patient episodes provides important insights into AKI prediction complementing other recent reviews in cardiac surgery, CI-AKI, liver transplantation and non-cardiac surgery.[14-17] In-patient mortality in those who developed the outcome ranged 6-42% (in the six studies reporting mortality) emphasising this is a crucial group to promptly identify.

The first limitation is the small number of externally validated models, which tempers recommending one model over another. Secondly, though we aimed to include general populations, caution should be employed, for example, when comparing a model derived on Heart Failure patients to one from an Orthopaedic cohort. However, in many UK hospitals, such populations share similarities (predominantly elderly demographic with co-morbidities) and if one aim of a prediction model is generalizability, a model should be tested in these different fields. Thirdly, as study outcome definitions were heterogenous, model comparisons are problematic, though recent studies were more likely to use KDIGO SCr change. Fourthly, no studies included urine output, probably reflecting the small number of patients who have this marker closely monitored. Fifthly, TRIPOD recommendations were used as a reporting benchmark, however, the relative importance of individual items and what constitutes an acceptable 'score' is arguable. A risk of bias assessment was not performed as none exists in this area, however, each study was critiqued

against reporting guidance. The absence of impact analysis limits the recommendation of one model over another. Finally, a meta-analysis was not performed without access to individual participant data (IPD). Expert guidance now exists in this area and offers opportunities to improve the scope of external validation research.[48 49]

Comparison with previous systematic reviews

Both this study and a review of CI-AKI models found pre-existing predictors - age, CKD, diabetes and heart failure to be the most commonly included.[15] A Cardiac surgery review reported specialty specific predictors in addition to these chronic co-morbidities. A non-Cardiac surgery review (5 of 6 studies in liver transplantation or resection) reported age, CKD and diabetes in at least two models.[17] Finally, a liver transplantation review highlighted the importance of CKD and (unsurprisingly) liver dysfunction.[16] The present review found drugs or acute laboratory values frequently included, though few models included acute physiological parameters. Our study and the non-Cardiac surgery review included adherence to recommended TRIPOD reporting with similar shortcomings. Across the other reviews, only in the fields of CI-AKI and Cardiac surgery were external validations reported.[14 15] Ease of use (including if necessary a calculator) and potential for electronic automation were rarely considered across the models reviewed. No impact analysis studies have been described.

Future directions

Management of HA-AKI presents a significant challenge, that could be helped by robust prediction models to risk stratify, encourage prevention and prompt recognition, key healthcare priorities.[6 10] Appraisal and synthesis of prediction studies may enable clinicians and policymakers judge model utility however, this is problematic when key study details are not reported.[12] Though much of the AKI literature is on (often assumed) hospital-acquired AKI, the majority of cases arise from the community (CA-AKI).[50 51] Indeed, a recent study demonstrated a significant proportion of such patients are never hospitalised.[52] This review suggests even in HA-AKI, the strongest predictors are pre-existing patient factors. The only laboratory measure frequently included – serum bicarbonate – may also reflect a chronic component. It is likely a proportion of cases classed as HA-AKI represent (evolving) community cases, thus, models using such pre-existing risk factors makes clinical sense. This continuum of harm between community and hospital could suggest that a risk prediction model in place at, or even *before* hospital admission, combined with early flagging of those who have met AKI criteria, may be required to improve outcomes. Electronic linkage of patient records between community and hospital data is desirable to ensure accurate inclusion of predictors (chronic morbidity, medication, laboratory and physiological parameters). This may also enable bedside automation as part of clinical workflow, where there is evidence beneficial implementation can be achieved.[18 53]

Impact analysis in prediction research is sparse making it difficult to conclude whether a model is worth implementing alongside, or replacing, usual care.[54] This is important as for example, one study suggested clinical acumen may be superior to prediction models,[55] whilst another found the combination of a model with clinical acumen was better than either alone.[56] Some impact analyses have suggested benefit, but conclusions are limited due to their rarity and design (mostly before-after without control).[57] There are a number of potential areas for impact analysis and clinical implementation (summarised in table 3). First, in specific populations a model could influence location of peri-operative care of surgical patients or drug and/or contrast dosing in patients with heart failure. Second, in a wider hospital setting the effects of highlighting those at highest risk to teams (ward, outreach critical care or Nephrology) with an adequate effector arm could be investigated. This has been demonstrated by existing AKI alerts in *established* AKI where outcome benefit has been limited to patients who had best practice delivered.[58-60] Third, as healthcare embraces complex technology, the inclusion of physiological (including urine output) or laboratory trends may be the only way to significantly improve model performance. Fourth, a model could identify a high risk group to be further risk stratified by employing one of the (increasing number of) available renal biomarkers.[61] Finally, one external validation study found those patients high risk on the prediction model who did develop AKI had a higher rate of mortality than the low risk group who developed HA-AKI, indicating the model predicts disease severity.[44] This could allow early review of such patients to help inform whether escalation of care may be required, or indeed be appropriate in the

increasing number of frail elderly patients admitted to hospitals.

To conclude, improving the management of patients to prevent AKI, or reduce associated complications, is a global health priority. This systematic review suggests there are few externally validated prediction models to help identify those at risk of AKI across general hospital populations. Future research should concentrate on validation, impact analysis and finally exploration of electronic implementation to enable clinical uptake.

Tables

Table 1 Summary of HA-AKI prediction models.

Population	General Surgery		T&O	General (Medical & Surgical)					Heart failure		
Author, year (n=derivation)	**Kheterpal 2007** (n=14,066)	**Kheterpal 2009** (n=57,080)	**Bell 2015** (n=6,220)	**Drawz 2008** (n=360)	**Matheny 2010** (n=26,107)	**Koyner 2016** (n=202,961)	**Bedford 2016** (n=7,556)	**Forni 2013** (n=1,867)	**Forman 2004** (n=1,004)	**Breidthardt 2011** (n=657)	**Wang 2013** (n=1,010)
Centres, Design	1, R	121, R	3, R	3, CC	1, R	5, R	3, R	1, P	11, R	1, P	1, R
Outcome predicted	eGFR <50 <7 days	↑SCr ≥177µmol/l, RRT	KDIGO ↑SCr	↑SCr*	RIFLE ↑SCr	KDIGO ↑SCr 24hr	KDIGO ↑SCr 72hr	KDIGO ↑SCr <7 days	↑SCr >26.5µmol/l**	↑SCr >26.5µmol/l**	AKIN SCr <48hr
Events	121	561	672	120	1,352	17,541	222	95	271	136	550
Mortality with outcome	15%	42%	-	-	-	6%	-	20%	27%	17%	17%
Mortality no outcome	3%^	8%^	-	-	-	1%	-	4%	-	6%	2%
Predictors tested	30	19	11	19	23	29	35	25	29	48	35
Predictors included	7	9	7	7	27	29	12	7	4	3	8
Derivation AUROC	0.77	0.80	0.74	0.73	0.75	-	-	0.72	-	0.71	0.76
IV AUROC	-	0.80	0.73	0.66	-	0.74	0.67	0.76	-	-	0.76
EV AUROC	0.67	-	0.71	-	-	-	0.71	0.65-0.71	0.65, 0.65	-	-
Derivation Calibration	RR	RR	Plot	-	H-L P=0.29	-		H-L P=0.96	-	-	H-L P=0.98
IV Calibration	-	RR	Plot	RR	-	-	H-L P=0.04	-	-	-	H-L P=0.13
EV Calibration	RR	-	Plot	-	-	-	H-L P=0.12	H-L P=0.06-0.09, Plot	RR	-	-
TRIPOD	25	28	34	26	28	24	29	29	26	23	30
Bedside calculation	-	-	-	-	-	-	-	Yes	Yes	Yes	Yes
Electronic Automation	-	-	Yes***	Yes	-	Yes	-	Yes	Yes	Yes	Yes

Design: R - retrospective, P – prospective, CC – case-control, Mortality - In-hospital. AUROC – area under the receiver operating characteristic curve, Plot – Calibration plot, EV – external validation, H-L – Hosmer-Lemeshow test, IV – internal validation, RR – risk range, RRT – renal replacement therapy, SCr – serum creatinine, TRIPOD – how many of the 37 recommended items were reported. *Increase sCr ≥44μmol/L if baseline SCr of ≤168μmol/L, ≥88μmol/L baseline 177-433μmol/L & ≥133μmol/L baseline >442μmol/L. **During admission, ***Used linked community and hospital data, ^Propensity matched.

Table 2 Summary of limitations in methodology and reporting

Area of concern	Description
Missing data	Multiple imputation recommended to avoid bias, rarely described.[12 62]
Definitions of outcome and predictors	No consistent strategy used to differentiate CA-AKI from HA-AKI. Two studies excluded patients with pre-existing CKD,[32 36] admission SCr frequently taken to be baseline; some studies defined co-morbidities from admission diagnoses whilst others used coded history.
Blinding of predictors or outcome	Not reported.
Sample size	Calculations not described, five studies had <10 EPV. Small sample increases risk of overfitting and underfitting.[12]
Univariate to select for multivariate analysis	Technique not recommended,[12] but used in 10 of 11 models.
Bootstrapping	Can adjust for optimism, without losing information - rarely described.[34 42]
Calibration plots	Important part of model performance,[12] present in only one model and one external validation.[34 44]
External validation and model updating	Validation adjusts for optimism, assesses generalizability. but was scarce, whilst model updating is recommended but not described.[12]
Newer performance measures	Techniques such as decision curve analysis offer insight into clinical consequences - not described.[28]
Use of data linkage	Only one study utilised data linkage.[34]

CA-AKI – community-acquired AKI, CKD – chronic kidney disease, EPV – events per variable, HA-AKI – hospital-acquired AKI, SCr – serum Creatinine.

Table 3. Potential areas for future Impact analysis of AKI prediction models

Population	Impact analysis to inform Clinical use
General Surgery	Peri-operative: haemodynamic targets, place of care, drugs, contrast delivery
Trauma & Orthopaedics	
General Populations	Risk stratification of large populations: intensity of observations, remote monitoring, application of biomarkers in subgroups at high-risk
Heart failure	Optimise haemodynamic status: diuretic dosing, use/volume of contrast

Legends to Figures

Figure 1 – PRISMA study flow chart

Figure 2 – Predictors most frequently included in the 11 AKI prediction models. ACEi – Angiotensin-converting enzyme inhibitors, ARBs – Angiotensin-receptor blockers, Bloods – laboratory parameters, CKD – chronic kidney disease, ↓HCO_3 – reduced serum bicarbonate, ↑WCC - raised white cell count.

Competing interests statement

"All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf and declare that all have no relationships with companies that might have an interest in the submitted work in the previous 3 years; their spouses, partners, or children have no financial relationships that may be relevant to the submitted work; and have no non-financial interests that may be relevant to the submitted work."

All authors had full access to all of the data (including statistical reports and tables) in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported and no important aspects of the study have been omitted.

Contributors: LH, LF, RV BDD and PR developed the idea for the study. LH, AS, LF, BDD and PR were involved in the study conception, preliminary literature review and design of the search strategy and the study protocol. LH, AS and LF were involved in screening and data extraction of papers. All

authors reviewed data extraction output. LH drafted the manuscript, which was critically reviewed and approved by all authors.

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors'.

References

1. KDIGO Clinical Practice Guideline for Acute Kidney Injury. Kidney Int 2012;**Suppl 2**(1):1-136 doi: 10.1038/kisup.2012.6published Online First: Epub Date]|.
2. Susantitaphong P, Cruz DN, Cerda J, et al. World incidence of AKI: a meta-analysis. Clin J Am Soc Nephrol 2013;**8**(9):1482-93 doi: 10.2215/cjn.00710113published Online First: Epub Date]|.
3. Bellomo R, Kellum JA, Ronco C. Acute kidney injury. Lancet 2012;**380**(9843):756-66 doi: 10.1016/s0140-6736(11)61454-2published Online First: Epub Date]|.
4. Bellomo R, Kellum J, Ronco C. Acute renal failure: time for consensus. Intensive Care Med 2001;**27**(11):1685-8 doi: 10.1007/s00134-001-1120-6published Online First: Epub Date]|.
5. Bagshaw SM, Gibney RT. Conventional markers of kidney function. Crit Care Med 2008;**36**(4 Suppl):S152-8 doi: 10.1097/CCM.0b013e318168c613published Online First: Epub Date]|.
6. National Confidential Enquiry into Patient Outcome and Death. Adding insult to injury: a review of the care of patients who died in hospital with a primary diagnosis of acute kidney injury (acute renal failure) http://www.ncepod.org.uk/2009report1/Downloads/AKI_report.pdf: NCEPOD, 2009.
7. *National Institute for Health and Care Excellence. Acute Kidney Injury: Prevention, Detection and Management Up to the Point of Renal Replacement Therapy. (Clinical guideline CG169).* London: National Clinical Guideline Centre., 2013.
8. Rabar S, Lau R, O'Flynn N, Li L, Barry P. Risk assessment of fragility fractures: summary of NICE guidance. BMJ 2012;**345**:e3698 doi: 10.1136/bmj.e3698published Online First: Epub Date]|.
9. Goff DC, Jr., Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation 2014;**129**(25 Suppl 2):S49-73 doi: 10.1161/01.cir.0000437741.48606.98published Online First: Epub Date]|.
10. Mehta RL, Cerdá J, Burdmann EA, et al. International Society of Nephrology's 0by25 initiative for acute kidney injury (zero preventable deaths by 2025): a human rights case for nephrology. The Lancet doi: 10.1016/S0140-6736(15)60126-Xpublished Online First: Epub Date]|.
11. Kyzas PA, Denaxa-Kyza D, Ioannidis JP. Almost all articles on cancer prognostic markers report statistically significant results. Eur J Cancer 2007;**43**(17):2559-79 doi: 10.1016/j.ejca.2007.08.030published Online First: Epub Date]|.
12. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015;**162**(1):W1-73 doi: 10.7326/M14-0698published Online First: Epub Date]|.
13. Bouwmeester W, Zuithoff NP, Mallett S, et al. Reporting and methods in clinical prediction research: a systematic review. PLoS Med 2012;**9**(5):1-

12 doi: 10.1371/journal.pmed.1001221published Online First: Epub Date]|.
14. Huen SC, Parikh CR. Predicting acute kidney injury after cardiac surgery: a systematic review. Ann Thorac Surg 2012;**93**(1):337-47 doi: 10.1016/j.athoracsur.2011.09.010published Online First: Epub Date]|.
15. Silver SA, Shah PM, Chertow GM, Harel S, Wald R, Harel Z. Risk prediction models for contrast induced nephropathy: systematic review. BMJ 2015;**351**:h4395 doi: 10.1136/bmj.h4395published Online First: Epub Date]|.
16. Caragata R, Wyssusek KH, Kruger P. Acute kidney injury following liver transplantation: a systematic review of published predictive models. Anaesth Intensive Care 2016;**44**(2):251-61
17. Wilson T, Quan S, Cheema K, et al. Risk prediction models for acute kidney injury following major noncardiac surgery: systematic review. Nephrol Dial Transplant 2016;**31**(2):231-40 doi: 10.1093/ndt/gfv415published Online First: Epub Date]|.
18. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 2005;**330**(7494):765 doi: 10.1136/bmj.38398.500764.8Fpublished Online First: Epub Date]|.
19. Altman DG. Systematic reviews of evaluations of prognostic variables. BMJ 2001;**323**(7306):224-8
20. Moons KG, de Groot JA, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Med 2014;**11**(10):e1001744 doi: 10.1371/journal.pmed.1001744published Online First: Epub Date]|.
21. Hospital Episode Statistics Analysis HaSCIC. Hospital Episode Statistics: Admitted patient care - 2014-15. http://www.hscic.gov.uk/pubs/hes1415: Health and Social Care Information Centre, 2015.
22. Selby NM, Crowley L, Fluck RJ, et al. Use of electronic results reporting to diagnose and monitor AKI in hospitalized patients. Clin J Am Soc Nephrol 2012;**7**(4):533-40 doi: 10.2215/cjn.08970911published Online First: Epub Date]|.
23. Porter CJ, Juurlink I, Bisset LH, Bavakunji R, Mehta RL, Devonald MA. A real-time electronic alert to improve detection of acute kidney injury in a large teaching hospital. Nephrol Dial Transplant 2014;**29**(10):1888-93 doi: 10.1093/ndt/gfu082published Online First: Epub Date]|.
24. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. JAMA 1992;**268**(2):240-8
25. Oxman AD, Guyatt GH. The science of reviewing research. Ann N Y Acad Sci 1993;**703**:125-33; discussion 33-4
26. Geersing GJ, Bouwmeester W, Zuithoff P, Spijker R, Leeflang M, Moons KG. Search filters for finding prognostic and diagnostic prediction studies in Medline to enhance systematic reviews. PLoS One 2012;**7**(2):e32844 doi: 10.1371/journal.pone.0032844published Online First: Epub Date]|.

27. Wilczynski NL, McKibbon KA, Walter SD, Garg AX, Haynes RB. MEDLINE clinical queries are robust when searching in recent publishing years. J Am Med Inform Assoc 2013;**20**(2):363-8 doi: 10.1136/amiajnl-2012-001075published Online First: Epub Date]|.
28. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;**26**(6):565-74 doi: 10.1177/0272989x06295361published Online First: Epub Date]|.
29. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007;**115**(7):928-35 doi: 10.1161/circulationaha.106.672402published Online First: Epub Date]|.
30. Hosmer DW, Hjort NL. Goodness-of-fit processes for logistic regression: simulation results. Stat Med 2002;**21**(18):2723-38 doi: 10.1002/sim.1200published Online First: Epub Date]|.
31. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited. Crit Care Med 2007;**35**(9):2052-6 doi: 10.1097/01.ccm.0000275267.64078.b0published Online First: Epub Date]|.
32. Kheterpal S, Tremper KK, Englesbe MJ, et al. Predictors of postoperative acute renal failure after noncardiac surgery in patients with previously normal renal function. Anesthesiology 2007;**107**(6):892-902 doi: 10.1097/01.anes.0000290588.29668.38published Online First: Epub Date]|.
33. Kheterpal S, Tremper KK, Heung M, et al. Development and validation of an acute kidney injury risk index for patients undergoing general surgery: results from a national data set. Anesthesiology 2009;**110**(3):505-15 doi: 10.1097/ALN.0b013e3181979440published Online First: Epub Date]|.
34. Bell S, Dekker FW, Vadiveloo T, et al. Risk of postoperative acute kidney injury in patients undergoing orthopaedic surgery-development and validation of a risk score and effect of acute kidney injury on survival: observational cohort study. BMJ 2015;**351**:h5639 doi: 10.1136/bmj.h5639published Online First: Epub Date]|.
35. Forni LG, Dawes T, Sinclair H, et al. Identifying the patient at risk of acute kidney injury: a predictive scoring system for the development of acute kidney injury in acute medical patients. Nephron Clin Pract 2013;**123**(3-4):143-50 doi: 10.1159/000351509published Online First: Epub Date]|.
36. Matheny ME, Miller RA, Ikizler TA, et al. Development of inpatient risk stratification models of acute kidney injury for use in electronic health records. Med Decis Making 2010;**30**(6):639-50 doi: 10.1177/0272989X10364246published Online First: Epub Date]|.
37. Drawz PE, Miller RT, Sehgal AR. Predicting hospital-acquired acute kidney injury--a case-controlled study. Ren Fail 2008;**30**(9):848-55 doi: 10.1080/08860220802356515published Online First: Epub Date]|.
38. Bedford M, Stevens P, Coulton S, et al. *Development of risk models for the prediction of new or worsening acute kidney injury on or during hospital admission: a cohort and nested study*. Southampton UK: Queen's Printer and Controller of HMSO., 2016.
39. Koyner JL, Adhikari R, Edelson DP, Churpek MM. Development of a Multicenter Ward-Based AKI Prediction Model. Clin J Am Soc Nephrol

2016;**11**(11):1935-43 doi: 10.2215/cjn.00280116published Online First: Epub Date]|.
40. Breidthardt T, Socrates T, Noveanu M, et al. Effect and clinical prediction of worsening renal function in acute decompensated heart failure. Am J Cardiol 2011;**107**(5):730-5 doi: 10.1016/j.amjcard.2010.10.056published Online First: Epub Date]|.
41. Wang YN, Cheng H, Yue T, Chen YP. Derivation and validation of a prediction score for acute kidney injury in patients hospitalized with acute heart failure in a Chinese cohort. Nephrology (Carlton) 2013;**18**(7):489-96 doi: 10.1111/nep.12092published Online First: Epub Date]|.
42. Forman DE, Butler J, Wang Y, et al. Incidence, predictors at admission, and impact of worsening renal function among patients hospitalized with heart failure. J Am Coll Cardiol 2004;**43**(1):61-7
43. Xing XZ, Wang HJ, Huang CL, et al. Two acute kidney injury risk scores for critically ill cancer patients undergoing non-cardiac surgery. World J Emerg Med 2012;**3**(4):278-81 doi: 10.5847/wjem.j.1920-8642.2012.04.007published Online First: Epub Date]|.
44. Hodgson L, Dimitrov, BD, Roderick, PJ, Venn, R, LG, Forni. Predicting AKI in Emergency Admissions: An external validation study of the acute kidney injury prediction score (APS). Bmj Open 2017;**In Press**
45. Noble D, Mathur R, Dent T, Meads C, Greenhalgh T. Risk models and scores for type 2 diabetes: systematic review. BMJ 2011;**343**:d7163 doi: 10.1136/bmj.d7163published Online First: Epub Date]|.
46. Collins GS, Mallett S, Omar O, Yu LM. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. BMC Med 2011;**9**:103 doi: 10.1186/1741-7015-9-103published Online First: Epub Date]|.
47. Mallett S, Royston P, Dutton S, Waters R, Altman DG. Reporting methods in studies developing prognostic models in cancer: a review. BMC Med 2010;**8**:20 doi: 10.1186/1741-7015-8-20published Online First: Epub Date]|.
48. Riley RD, Ensor J, Snell KIE, et al. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. BMJ 2016;**353** doi: 10.1136/bmj.i3140published Online First: Epub Date]|.
49. Debray TPA, Damen JAAG, Snell KIE, et al. A guide to systematic review and meta-analysis of prediction model performance. BMJ 2017;**356** doi: 10.1136/bmj.i6460published Online First: Epub Date]|.
50. Wonnacott A, Meran S, Amphlett B, Talabani B, Phillips A. Epidemiology and outcomes in community-acquired versus hospital-acquired AKI. Clin J Am Soc Nephrol 2014;**9**(6):1007-14 doi: 10.2215/cjn.07920713published Online First: Epub Date]|.
51. 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**(9):1510-8 doi: 10.2215/cjn.02140215published Online First: Epub Date]|.
52. Sawhney S, Fluck N, Fraser SD, et al. KDIGO-based acute kidney injury criteria operate differently in hospitals and the community-findings from a large population cohort. Nephrol Dial Transplant 2016;**31**(6):922-9 doi: 10.1093/ndt/gfw052published Online First: Epub Date]|.

53. Kannry J, McCullagh L, Kushniruk A, Mann D, Edonyabo D, McGinn T. A Framework for Usable and Effective Clinical Decision Support: Experience from the iCPR Randomized Clinical Trial. EGEMS (Wash DC) 2015;**3**(2):1150 doi: 10.13063/2327-9214.1150published Online First: Epub Date]|.
54. Toll DB, Janssen KJ, Vergouwe Y, Moons KG. Validation, updating and impact of clinical prediction rules: a review. J Clin Epidemiol 2008;**61**(11):1085-94 doi: 10.1016/j.jclinepi.2008.04.008published Online First: Epub Date]|.
55. Sinuff T, Adhikari NK, Cook DJ, et al. Mortality predictions in the intensive care unit: comparing physicians with scoring systems. Crit Care Med 2006;**34**(3):878-85 doi: 10.1097/01.ccm.0000201881.58644.41published Online First: Epub Date]|.
56. Brabrand M, Hallas J, Knudsen T. Nurses and physicians in a medical admission unit can accurately predict mortality of acutely admitted patients: a prospective cohort study. PLoS One 2014;**9**(7):e101739 doi: 10.1371/journal.pone.0101739published Online First: Epub Date]|.
57. Fillmore CL, Bray BE, Kawamoto K. Systematic review of clinical decision support interventions with potential for inpatient cost reduction. BMC Med Inform Decis Mak 2013;**13**:135 doi: 10.1186/1472-6947-13-135published Online First: Epub Date]|.
58. Wilson FP, Shashaty M, Testani J, et al. Automated, electronic alerts for acute kidney injury: a single-blind, parallel-group, randomised controlled trial. Lancet 2015 doi: 10.1016/s0140-6736(15)60266-5published Online First: Epub Date]|.
59. Kolhe NV, Reilly T, Leung J, et al. A simple care bundle for use in acute kidney injury: a propensity score-matched cohort study. Nephrol Dial Transplant 2016;**31**(11):1846-54 doi: 10.1093/ndt/gfw087published Online First: Epub Date]|.
60. Kolhe NV, Staples D, Reilly T, et al. Impact of Compliance with a Care Bundle on Acute Kidney Injury Outcomes: A Prospective Observational Study. PLoS One 2015;**10**(7):e0132279 doi: 10.1371/journal.pone.0132279published Online First: Epub Date]|.
61. Murray PT, Mehta RL, Shaw A, et al. Current Use of Biomarkers in Acute Kidney Injury: Report and Summary of Recommendations from the 10(th) Acute Dialysis Quality Initiative Consensus Conference. Kidney international 2014;**85**(3):513-21 doi: 10.1038/ki.2013.374published Online First: Epub Date]|.
62. Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. J Clin Epidemiol 2006;**59**(10):1087-91 doi: 10.1016/j.jclinepi.2006.01.014published Online First: Epub Date]|.

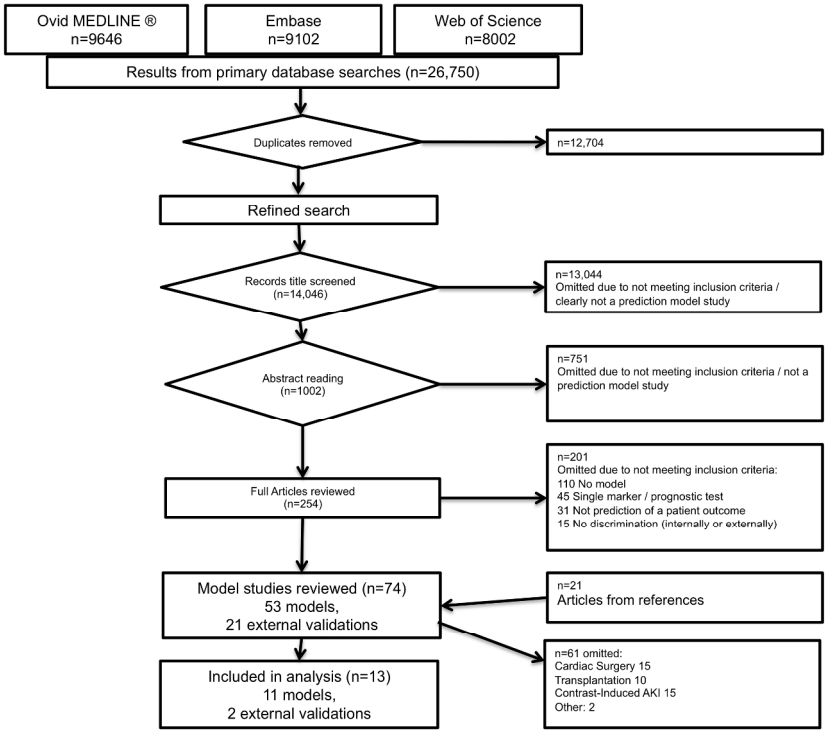

Figure 1 – PRISMA study flow chart

1057x793mm (72 x 72 DPI)

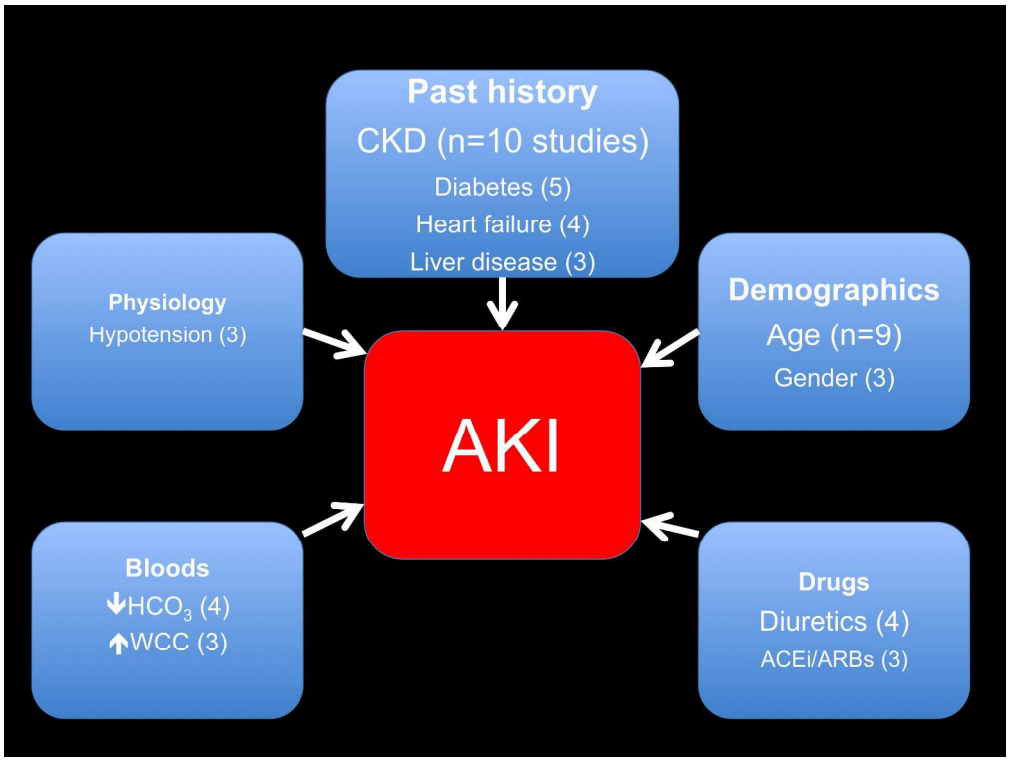

Figure 2 – Predictors most frequently included in the 11 AKI prediction models. ACEi – Angiotensin-converting enzyme inhibitors, ARBs – Angiotensin-receptor blockers, Bloods – laboratory parameters, CKD – chronic kidney disease, ⌊HCO3 – reduced serum bicarbonate, ⌈WCC - raised white cell count.

1057x793mm (72 x 72 DPI)

Online Supplementary file

eTable 1 Study inclusion criteria
eTable 2 Embase Search
eTable 3 Ovid MEDLINE® search
eTable 4 - Web of Science search
eTable 5 - CHARMS checklist and data extracted for systematic review
eTable 6(i-iv) – Full details of models reviewed
eTable 7 – Abbreviations used
eTable 8 - TRIPOD items reported in the 11 studies
eTable 9 – Most common predictors used in the 11 models
eTable 10 – All predictors included in the 11 models

eTable 1 - Study inclusion criteria.

Inclusion Criteria
• Articles in peer-reviewed journals reporting a prognostic multivariable prediction model (scoring system or algorithm) identifying patients who developed HA-AKI (or other measures of renal dysfunction in older studies)
• Validation studies (and updating) of an existing model
• Retrospective, prospective and case-control designs
• Adults (≥18 years) in general hospital settings
• Statistical measures of discrimination (AUROC or c-statistic)
Exclusion Criteria
• Patients <18 years old
• Cardiac surgery, other specialised surgery (e.g. transplantation), CI-AKI
• Non-human studies
• Case reports or conference abstracts
• Only logistic regression without a prediction model
• Lack of discrimination statistics (unless model validated elsewhere)
• Studies that investigated a single predictor, test, or marker
• Studies that investigated only causality between one or more predictors & an outcome
• Use of patients already with the outcome (e.g. AKI present at hospital admission)
• Patients in primary care
• Novel, not widely available tests, such as biomarkers

AUROC - area under the receiver-operating characteristic curve, CI-AKI – Contrast-Induced AKI, HA-AKI - hospital-acquired-AKI.

eTable 2 - Embase Search

RESULTS	LINE	SEARCH TERM
16946	1	(acute AND kidney AND injury).ti,ab
7601	2	AKI.ti,ab
42560	3	(acute AND renal AND failure).ti,ab
10263	4	ARF.ti,ab
2911	5	(contrast AND induced AND nephropathy).ti,ab
31958	6	ACUTE KIDNEY INJURY/
78043	7	1 OR 2 OR 3 OR 4 OR 5 OR 6
1277060	8	predict*.ti,ab
71645	9	PREDICTIVE VALUE OF TESTS/
835041	10	scor*.ti,ab
2999344	11	observ*.ti,ab
17216	12	OBSERVER VARIATION/
4625679	13	8 OR 9 OR 10 OR 11 OR 12
20315	14	7 AND 13
16946	15	(acute AND kidney AND injury).ti,ab
7601	16	AKI.ti,ab
42560	17	(acute AND renal AND failure).ti,ab
10263	18	ARF.ti,ab
2911	19	(contrast AND induced AND nephropathy).ti,ab
48915	20	ACUTE KIDNEY FAILURE/
84125	21	15 OR 16 OR 17 OR 18 OR 19 OR 20
1277060	22	predict*.ti,ab
3711942	23	exp METHODOLOGY/
400667	24	validat*.ti,ab
4889839	25	22 OR 23 OR 24
18916	26	21 AND 25
9102	27	14 [Limit to: Human and (Publication Types Article)]

eTable 3 Ovid MEDLINE® search

Results	Line	Search term
10612	1	(acute AND kidney AND injury).ti,ab
4343	2	AKI.ti,ab
30834	3	(acute AND renal AND failure).ti,ab
8629	4	ARF.ti,ab
1811	5	(contrast AND induced AND nephropathy).ti,ab
33267	6	ACUTE KIDNEY INJURY/
60802	7	28 OR 29 OR 30 OR 31 OR 32 OR 33
1005899	8	predict*.ti,ab
149557	9	PREDICTIVE VALUE OF TESTS/
581297	10	scor*.ti,ab
2581563	11	observ*.ti,ab
33069	12	OBSERVER VARIATION/
3877048	13	35 OR 36 OR 37 OR 38 OR 39
12929	14	34 AND 40
9646	15	41 [Limit to: (Document type Classical Article or Clinical Trial or Clinical Trial, Phase I or Clinical Trial, Phase Ii or Clinical Trial, Phase Iii or Clinical Trial, Phase Iv or Comment or Comparative Study or Controlled Clinical Trial or Corrected And Republished Article or English Abstract or Evaluation Studies or Historical Article or Journal Article or Multicenter Study or Observational Study or Pragmatic Clinical Trial or Randomized Controlled Trial or Report or Validation Studies) and Humans]

eTable 4 - Web of Science search

RESULTS	WEB OF SCIENCE SEARCH
8002	(TS=((acute kidney injury) OR (aki) OR (acute renal failure) OR (arf) OR (contrast induced nephropathy)) AND TS=(predict* OR scor* OR observ* OR validat*)) AND DOCUMENT TYPES: (Article) Refined by: WEB OF SCIENCE CATEGORIES: (UROLOGY NEPHROLOGY OR SURGERY OR CARDIAC CARDIOVASCULAR SYSTEMS OR TRANSPLANTATION OR CRITICAL CARE MEDICINE OR MEDICINE GENERAL INTERNAL OR MEDICAL INFORMATICS OR GASTROENTEROLOGY HEPATOLOGY OR ANESTHESIOLOGY) AND DOCUMENT TYPES: (ARTICLE) AND WEB OF SCIENCE CATEGORIES: (UROLOGY NEPHROLOGY OR SURGERY OR CARDIAC CARDIOVASCULAR SYSTEMS OR TRANSPLANTATION OR MEDICAL INFORMATICS OR CRITICAL CARE MEDICINE OR HEALTH CARE SCIENCES SERVICES OR MEDICINE GENERAL INTERNAL OR MEDICAL LABORATORY TECHNOLOGY OR GASTROENTEROLOGY HEPATOLOGY OR ANESTHESIOLOGY OR EMERGENCY MEDICINE) Indexes=SCI-EXPANDED Timespan=All years

eTable 5 - CHARMS checklist and data extracted for systematic review.

Item	Explanation in the Review
1. Type of studies	Prognostic prediction models
2. Scope	Published prognostic prediction models for development of AKI in general hospital settings; to inform risk stratification & potential uses in decision-making in different patient groups
3. Type of studies	Model development +/- external validation in independent data; external model validation & model updating, if present
4. Target population	Adult (≥18) Patients in acute hospital environment
5. Outcome predicted	Development of AKI (or equivalent definition, including RRT) after an admission to hospital or Surgery
6. Time span of prediction	In-hospital development of the outcome
7. Intended moment of using the model	Pre-operatively to predict the risk of post-op AKI or need for RRT; at admission to risk stratify or guide therapy
Summary of Data extracted	
• Data source (years, retrospective, prospective; cohort, case-control, trial data) • Participants & setting (eg cardiac surgery, single or multi-centre, country • Primary outcome (and any blinding) • Candidate predictors (definitions; continuous data dichotomised? & how selected for modelling) • Sample size, EPV (including all predictors considered) • Type of model(s) evaluated - derivation, validation (internal, external) • Missing data, number included & excluded (criteria) • Type of model (eg full model approach), shrinkage • Incidence of outcome & mortality data • Predictors in the model(s) • Performance: discrimination (AUC or C-Statistic) & calibration (eg H-L P value, slope/curve), risk groups • Internal & external validation (in same study) • External validation studies with relevant performance measures • Additional resources, funding	

AKI – acute kidney injury, EPV – events per variable, H-L – Hosmer-Lemeshow goodness-of-fit test, RRT – renal replacement therapy.

eTable 6(i) Surgery

Author, Type, TRIPOD detail	Population, Outcome, AKI Definitions, Methods	Outcomes, Predictors & Model Performance	External Validation
Kheterpal 2007	USA single centre, retrospective cohort study (n=65,043). Data collected 2003-6.	Outcome (AKI) in 0.8% (n=121), 0.1% (n=14) required RRT. Propensity matched 30-day mortality with outcome 15% (n=17/118) vs. 2.7% (n=9/352) without. AKI associated with significant increase in 30-day, 60-day, 1-yr mortality.	
General surgery	Inclusion: pre-op eGFR (Cockcroft-Gault) ≥80 ml/min; major surgery (≥2 days in-patient).	7 pre-op predictors: age, emergent surgery, liver disease, BMI, high-risk surgery, PVD & COPD.	
TRIPOD 1A - Derivation	Exclusions (n=49,941): cardiac, transplant, urology & ECT, suprarenal aortic cross-clamping; pre-op AKI & IV contrast <7 days post-op, no pre-op SCr (n=6,534), pre-op eGFR <80 (n=5659). Included: n=15,102.	Weighted c-Statistic 0.77 (95% CIs 0.75-0.79). Un-weighted risk factor scale (cut-off Age >59, BMI ≥32) c-Statistic 0.73 (0.7-0.76).	Xing 2012 - AUC 0.66
(25/37 pts)	Outcome: reduction of eGFR to ≤50ml/min <7 days post-op.	With intra-op: vasopressor dose, infusion & diuretic: AUC 0.79 (0.77-0.81)	
	Predictors: 24 pre, 6 intra-op.	No calibration statistics.	
	Collinearity predictors evaluated; bivariate correlation matrix; remaining predictors entered into logistic regression full model fit. Missing data: excluded from full model. After exclusions n=14,066 included. Un-weighted model continuous predictors dichotomised.		
Kheterpal 2009	USA multi-centre (121) retrospective database study (n=152,244). 2005-6.	Outcome in 1% (n=762/75,952) – n=561 derivation, n=201 in validation sets.	
General surgery	Included n=75,952. Random split derivation 75% (n=57,080) & validation (25% n= 18,872).	Mortality 42% (n=320) in those with outcome vs 8% in a propensity matched group without outcome.	
TRIPOD 2A - Derivation, Validation	Exclusions (n=76,292): vascular, cardiac, urology, ophthalmology, obstetric, or urologic procedures; day case; pre-op AKI (rapidly increasing azotaemia & SCr ≥265 µmol/L <24h of surgery) or previous RRT (n=1637).	9 predictors (simplified risk index): age ≥56 yr, male, emergency, intraperitoneal surgery, diabetes, CCF, ascites, HTN, mild or moderate pre-op renal insufficiency.	-
(28/37 pts)	Mild pre-op renal insufficiency defined SCr 106-168 µmol/L; moderate >177 µmol/L.	c-Statistic 0.80 (0.79-0.81) in derivation & internal validation cohorts.	
	Outcome: AKI defined as increase SCr to ≥177 µmol/L or RRT <30 days.	Calibration: Risk classes reported for derivation vs validation sets.	
	Missing data: SPSS assessed impact of imputation. Continuous predictors dichotomised. Collinearity & Pearson correlations evaluated for all 19 preoperative predictors (comorbidities, drugs, type of surgery). Remaining predictors entered into full model fit logistic regression.		

eTable 6(ii) Trauma & Orthopaedics			
Author, Type, TRIPOD detail	**Population, Outcome, AKI Definitions, Methods**	**Outcomes, Predictors & Model Performance**	**External Validation**
Bell 2015	UK multi-centre (3) retrospective cohort study linking multiple prospectively collected databases (n=15,218). 2005-11.	Outcome (AKI) in 10.8% (n=672) derivation & 6.7% (n=295) validation sets. With AKI adjusted hazard ratio 1.53 (95% CI 1.38-1.70).	
T&O	Included: derivation n=6,220 (2 sites) & validation n=4,395 (1 site).	7 predictors: age, male, diabetes, number drugs, eGFR, ACEi/ARBs & ASA. Risk calculator supplied.	
TRIPOD 3,4 - Derivation, Internal & EV	Exclusions: missing SCr (n=2,688), RRT, 2nd operation (n=1,915).	Derivation AUC 0.74 (0.72-0.76), Internal validation 0.73.	Same Study different site AUC 0.70
(34/37 pts)	Outcome: KDIOGO SCr changes <7 days. eGFR: CKD-EPI. Admission SCr taken as baseline.	EV 0.70. Risk groups shown.	
	Entered 11 candidate predictors (age, sex, renal function, diabetes, number drugs, ACEi/ARB, NSAID/COX-2, statin, urgency, ASA grade & deprivation category into Backward/forward multivariable selection. Applied a conservative selection criterion of P<0.15 to limit over-fitting risk.	Calibration plot. Calibration suboptimal in validation cohort (over-predicted risk).	
	Bootstrapping for IV. To assess robustness sensitivity analyses performed: multiple imputation relaxing & restricting the backward selection removal criterion & adding non-linear & interaction terms. Categorised eGFR.	Re-calibration: correction factor, added to intercept; intercept and regression coefficient index as the only predictor used to transform prognostic index & compute recalibrated probabilities.	

eTable 6(iii)
General admissions

Author, Type, TRIPOD detail	Population, Outcome, AKI Definitions, Methods	Outcomes, Predictors & Model Performance	External Validation
Drawz 2008	USA multi-centre (3), retrospective, case-controlled study (n=180 cases, n=360 controls). 2003.	No information on mortality.	
Medicine, surgery, obstetrics	Hospital-acquired AKI (HA-AKI) defined: increase SCr ≥44µmol/L if baseline SCr ≤168µmol/L), ≥88µmol/L baseline 177-433µmol/L & ≥133µmol/L baseline >442µmol/L. Admission SCr presumed to be baseline.	7 predictors: age, SBP, HR, HCO_3, urea, albumin & drugs (NSAIDs, ACE-I, ARBs or diuretic).	
TRIPOD 2A - Derivation, Internal Validation	'Control' cases - mix of same discharge diagnosis or next patient admitted to clinical team. Inclusions: age ≥18 & normal admission SCr or admission SCr not qualifying as AKI vs known baseline. Exclusions: RRT, no repeat SCr performed.	Derivation c-statistic 0.73. Simplified: HR ≥70/min, HCO_3 (<24 or >30mmol/L), SCr ≥88µmol/L & drugs. c-statistic derivation 0.69. Internal validation both models - 0.66.	-
(26/37 pts)	19 predictors assessed: demographics (age, sex, race), medical history, medications & admission observations (BP, HR, HCO_3, urea, SCr, & albumin). Predictors with p value <0.20 univariate analysis entered into multiple logistic regression model. Cases with missing data excluded.	No H-L p-value. Risk range in validation set plotted: 0/1 risk factor = 16% risk developing HA-AKI, vs 4 risk factors = 62% risk HA-AKI.	
Matheny 2010	USA single centre, retrospective cohort study (n=61,179). 1999-2003.	AKI Risk 5.2% (n=1,352), AKI Injury 2.8% (n=726).	
General admissions	Inclusions: adult admissions ≥2 days (n=26,107).	No mortality data.	
TRIPOD 1B - Derivation	Exclusions: missing data, baseline eGFR <60 (n=11,342), AKI on admission, no SCr available <48 hrs of admission (n=10,378) or no repeat SCr (n=13,352).	27 predictors: Female, Age, Race, 11 classes of drugs, Contrast, bacterial infection (use of antibiotics), SCr, MI, rhabdomyolysis, hepatitis, pancreatitis, ammonia, AST/ALT ratio, thrombocytopenia, leucocytosis, hypercalcaemia, glucose.	-
(28/37 pts)	Outcome (<30 days post admission): AKI Risk = ≥2 SCr results ≥150% of baseline. AKI Injury = ≥200% baseline. eGFR using MDRD equation.	AKI Risk: AUC 0.75 (0.73–0.76). H-L P = 0.29. AKI Injury: AUC 0.78 (0.76–0.79), H-L P=0.12.	
	27 predictors assessed: coded diagnoses (including admission diagnosis) & drugs following univariate analysis placed in multivariable model. Missing values captured as a separate category.	Calibration plotted by deciles.	
	10-fold cross-validation employed to estimate overfitting.		

eTable 6(iii) General admissions			
Author, Type, TRIPOD detail	**Population, Outcome, AKI Definitions, Methods**	**Outcomes, Predictors & Model Performance**	**External Validation**
Forni 2013 **General medical TRIPOD 2B, 4 – Derivation, Internal & EV (29/37 pts)**	UK single centre. Prospective cohort study (n=3,707). 2012. Inclusion: medical patients staying >1 night in hospital (n=1,867). Exclusions: RRT, non-medical patients, age <18, AKI on admission (n=184), missing data (n=553). Included n=3,523. Derivation n=1,867. Outcome: AKI (KDIGO SCr change <7 days). Pre-admission SCr measured >1 month & <6 months. Internal validation: patients with no previous SCr result, but with a SCr on admission within normal range (defined 80-120µmol/L) (n=1,656). CKD defined - eGFR <60. 25 predictors on univariate, If P <0.05 variable entered into multivariable analysis. No missing data information.	Derivation group developed AKI 7% (n=95) - mortality 20% vs 3.5% (n=62) without outcome. In validation cohort n=60 developed AKI. 7 predictors: Age 60-79 (1 point) ≥80 (3 pts), CCF, CKD, Diabetes (2 pts), Liver disease (3 pts), respiratory rate ≥20/min, <alert on AVPU score (3 pts). Derivation AUC 0.72 (0.66–0.77). H-L P=0.96. Risks plotted. Validation AUC 0.76 (0.71–0.82). No H-L reported.	Hodgson 2017, AUC 0.65-0.71
Bedford 2016 **General admissions TRIPOD 2A, 3 – Derivation, EV** **(29/37 pts)**	UK multi-centre (3), 2011. Retrospective cohort study (n=11,655). Included: derivation n=7,556 admissions & internal validation n=2,514. Exclusions: non-emergency, pre-admission AKI, AKI at admission, obstetrics, patients with no info on AKI at 72 hours. Outcomes: AKI & AKI Stage 2/3. AKI <72 hours, using KDIGO change in SCr. Ordinal logistic regression with univariable analysis for development of multivariable analysis. Backwards selection used for retention of statistically significant predictors. Missing data excluded or given own category. 3:1 random split for internal validation. External validation n=1,585, single centre.	Derivation AKI 9.6% (n=241), AKI 2/3: n=40. No mortality data. EV AKI 7.6% (n=120), AKI 2/3 n=12. 12 predictors: age, primary diagnosis, previous hospital admissions, Charlson co-morbidity index score, HbA1C, troponin, proteinuria, baseline eGFR, K^+, WCC, Mg^{2+}, CRP. IV AUC 0.67 (0.64-0.71) any AKI, 0.68 for AKI 2/3. No derivation AUC H-L P=0.04 any AKI model, P=0.005 for AKI 2/3.	Same study AUC 0.71 (0.63 AKI 2/3). H-L P=0.12 AKI, P=0.14 for AKI 2/3.
Koyner 2016 **General admissions** **TRIPOD 2A - Derivation, Internal Validation** **(24/37 pts)**	USA multi-centre (5) Retrospective cohort study (n=269,999). 2008-2013. Included: n=202,961. Exclusions: SCr >354 µmol/L on admission (n=11,305), those without SCr measurement (n=52,508) & AKI prior to arrival on ward (n=3,225). Model included 29 predictors. Continuous predictors modelled using restricted cubic splines with knot placement. Variable importance plot created. Lab values & vital signs updated periodically during admission therefore separated into time intervals & logistic regression used for model estimation. Values closest to beginning of that time variable used to predict outcome for that interval, if no values available during an interval, most recent value used, if no previous value available, median value across entire cohort for that variable imputed. Split derivation (60%) and internal validation (40%) by time. Admission SCr defined as baseline.	AKI 8.6% (n=17,541). Mortality with outcome 6% (n=1031) vs 1% (n=1,419) without 29 predictors: SCr, Urea, HR, anion gap, Urea/SCr, RR, glucose, WCC, K^+, O_2 Sats, age, HCO_3, Na^+, temperature, prior ICU, albumin, bilirubin, Ca^{2+}, platelets, time, SBP/ DBP, pulse pressure, sex, AVPU, alk phosphatase, Hb, total protein, AST For AKI AUC 0.74 (0.74-0.74), AKI Stage 3 AUC 0.83 (0.83-0.84) Model including only SCr, BUN & their ratio AUC 0.69 (0.68-0.69). Internal validation	-

eTable 6(iv)
Heart failure

Author, Type, TRIPOD detail	Population, Outcome, AKI Definitions, Methods	Outcomes, Predictors & Model Performance	External Validation
Forman 2004 **TRIPOD 1B - Derivation, Internal Validation** **(26/37 pts)**	USA multi-centre (11) retrospective cohort study (n=1,009). 1997-8. Exclusion (number not given): elective, <2 days, severe aortic stenosis, anticipated transplant, RRT, LVAD, high output failure, age <20, chemotherapy. Excluded n=5 with missing charts. Included n=1004. Outcome: worsening renal function (WRF) - rise SCr >26.5µmol/l during admission. 29 predictors assessed: demographics, history, drugs, symptoms, signs. Multivariable Cox regression models, stepwise selection. Bootstrapping for IV. Missing data: predictors missing >15% excluded; categorical data assumed "not present" & separate dummy indicator used if >5% of values missing.	'WRF' 27% (271/1,004). Mortality: risk ratio 7.5 with outcome (number not reported). 4 predictors: CCF, diabetes & BP >160 mmHg (1 point), SCr 132.6-212µmol/l (2 points) & SCr >221µmol/l (3 points). Risk 'WRF': 0 pts = 10%, 1 = 19%, 2 = 20%, 3 = 30%, 4+ = 53%. 22% of total sample with risk score ≥4 had 53% likelihood WRF vs 10% risk among 12% with risk score 0 points (p<0.001). **No AUC or Calibration statistics.**	Breidthardt 2011 - 0.65 Wang 2013 - 0.65
Breidthardt 2011 **TRIPOD 1A – Derivation, (EV Forman)** **(23/37 pts)**	Swiss multi-centre (3) prospective analysis, with derivation (Basel score) & external validation of Forman score (n=767). 2001-2, 2006-2010. Included n=657. Exclusions (n=110): stay <2 days, incomplete SCr. Outcome: WRF = in-hospital increase SCr ≥26.5µmol/L. eGFR – MDRD equation. CKD = eGFR <60 for >3/12 pre-admission. 48 predictors assessed on univariate & those with P value <0.05 entered into multivariable analysis. No missing data information. n=223 had blood gas analysis.	Outcome 21% (136/657). In-hospital mortality with outcome 17% (n=23) vs 6% (n=33) without (P <0.01). 3 predictors (n=223): HCO_3 <21 mmol/L, Diuretics, CKD - AUC 0.71 (0.63-0.79). No H-L calibration data. Scores & percentage developing outcome: 0 - 1%, 1 –35%, 2 –27%, 3 – 35%.	-

eTable 6(iv)
Heart failure

Author, Type, TRIPOD detail	Population, Outcome, AKI Definitions, Methods	Outcomes, Predictors & Model Performance	External Validation
Wang 2013 **TRIPOD 2A – Derivation, External validation (Forman)** **(30/37 pts)**	China, single centre, retrospective cohort study (n=1,709). 2004-11.	AKI 32% (n=550). Mortality 16.5% (n=91) vs.1.9% (n=22) without AKI (P <0.01). Stay with AKI 14 vs. 11 days without (P <0.01).	
	Inclusion: CCF admission diagnosed by 2 cardiologists using European Society of Cardiology guidelines.	8 predictors: Age ≥70; ≥3 CCF admissions, systolic BP <90mmHg, sodium <130mmol/L, NYHA IV, proteinuria, SCr ≥104 µmol/L & furosemide dose ≥80 mg/day.	
	Exclusions: age <18, stay <2 days, missing data, hospital transfer, use LVAD, ESRD or RRT & septic or haemorrhagic shock; cardiac op, pacemaker or cardioversion & contrast.	Derivation AUC 0.76 (0.73–0.79) H-L P=0.98. Calibration plots by deciles. Validation 0.76 (0.72-0.8), H-L P=0.13.	-
	Split derivation (60%, n=1010) & validation (40%, n=699).	≥8 points high risk - 55.1% incidence vs. 18% if <8 points No calibration slope.	
	Outcome: AKI (AKIN): increase SCr ≥26.4 µmol/L or ≥50% in <48 hrs. eGFR - MDRD.	Forman - 0.65 (0.62–0.69). vs Forman score, improvement of 0.11 AUC, (P <0.001(DeLong.(9)	
	35 predictors – those with P value <0.1 on unviariate analysis placed in multivariate analysis (n=932).		

eTable 7 – Abbreviations used in eTable 6(i-iv)

ACEi – Angiotensin-converting enzyme inhibitors
AKI – Acute kidney injury
AKIN – Acute kidney injury network
ALT – Alanine transferase
ARB – Angiotensin receptor blockers
AUC/AUROC – Area under the receiver operating characteristic curve
BMI – Body mass index
CA-AKI – Community-acquired AKI
CI-AKI – Iodinated contrast AKI
CKD – Chronic kidney disease
COPD - Chronic obstructive pulmonary disease
CKD-EPI – CKD Epidemiology collaborative equation
COX – Cyclo-oxygenase
D – Derivation study
eGFR – estimated glomerular filtration rate
ESRD – end-stage renal disease
EV – External validation study
HA-AKI – Hospital-acquired AKI
HCO_3 – sodium bicarbonate
H-L – Hosmer-Lemeshow goodness-of-fit test (Calibration statistic)
HTN – Hypertension
ICU – Intensive Care Unit
IHD – Ischaemic heart disease
KDIGO – Kidney disease improving global outcomes
LOS – length of stay
LVAD – Left ventricular assist device
LVEF – Left ventricular ejection fraction
MAP – Mean arterial pressure
MDRD – Modification of diet in renal disease equation
MI – Myocardial infarction
Na^+ – sodium
NSAID – Non-steroidal anti-inflammatory agent
NYHA – New York Heart Association Classification for heart failure (I-IV)
PVD – Peripheral vascular disease
RIFLE – Risk, Injury, failure, loss of kidney function
RRT – Renal replacement therapy
SCr – serum creatinine
TRIPOD – Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis. Checklist in derivation (37 points – 1 point for each recommended item reported).
TRIPOD Study types
Type 1a: Development only
Type 1b: Development and validation using resampling
Type 2a: Random split-sample development and validation,
Type 2b: Non-random split-sample development and validation
Type 3: Development and validation using separate data
Type 4: Validation only.

eTable 8 - TRIPOD items reported in the 11 studies.

Title & Abstract		TRIPOD Item description	Reported ?
Title	1	Identify study as developing &/or validating a multivariable prediction model, target population & outcome.	10
Abstract	2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results & conclusions.	9
Introduction			
Background & objectives	3a	Explain medical context (including whether diagnostic or prognostic) & rationale for developing or validating the multivariable prediction model, including references to existing models.	11
	3b	Specify objectives, including whether the study describes development or validation of the model or both.	11
Methods			
Source of data	4a	Describe study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development & validation data sets, if applicable.	11
	4b	Specify key study dates, including start of accrual; end of accrual; & if applicable, end of follow-up.	11
Participants	5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number & location of centres.	11
	5b	Describe eligibility criteria for participants.	11
	5c	Give details of treatments received, if relevant.	9
Outcome	6a	Clearly define outcome predicted by the prediction model, including how & when assessed.	11
	6b	Report any actions to blind assessment of the outcome to be predicted.	0
Predictors	7a	Clearly define all predictors used in developing or validating the model, including how & when they were measured.	11
	7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	0
Sample size	8	Explain how the study size was arrived at.	2
Missing data	9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	7
Statistical analysis methods	10a	Describe how predictors were handled in the analyses.	11
	10b	Specify type of model, model-building procedures (including predictor selection) & method for internal validation.	11
	10c	For validation, describe how the predictions were calculated.	6
	10d	Specify all measures used to assess model performance & if relevant, to compare multiple models.	8
	10e	Describe any model updating (e.g., recalibration) arising from the validation, if done.	1
Risk groups	11	Provide details on how risk groups were created, if done.	10
Development vs. validation	12	For validation, identify differences from development data in setting, eligibility criteria, outcome & predictors.	4
Results			
Participants	13a	Describe flow of participants through the study, including number of participants with & without the outcome & if applicable, a summary of the follow-up time. A diagram may be helpful.	11
	13b	Describe characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors & outcome.	9
	13c	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	3
Model development	14a	Specify the number of participants & outcome events in each analysis.	11
	14b	If done, report the unadjusted association between each candidate predictor & outcome.	11
Model specification	15a	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients & model intercept or baseline survival at a given time point).	8
	15b	Explain how to the use the prediction model.	11
Model performance	16	Report performance measures (with CIs) for the prediction model.	8
Model-updating	17	If done, report the results from any model updating (i.e., model specification, model performance).	1
Discussion			
Limitations	18	Discuss any limitations of the study (non-representative sample, few events per predictor, missing data).	11
Interpretation	19a	For validation, discuss results with reference to performance in development data & any other validation data.	8
	19b	Give overall interpretation of results, considering objectives, limitations, results from similar studies & other relevant evidence.	11
Implications	20	Discuss potential clinical use of the model & implications for future research.	11
Other information			
Suppl info	21	Provide information about availability of supplementary resources, (study protocol, Web calculator, & data sets).	3
Funding	22	Give the source of funding & role of the funders for the present study.	9

Red colouring highlights items reported in less than 50% of the 11 AKI prediction model studies.

eTable 9 – Most common predictors included in the 11 models

Field	Total	General Surgery		T&O	General					Heart Failure		
Study		Kheterpal 2007	Kheterpal 2009	Bell 2015	Drawz 2008	Matheny 2010	Koyner 2016	Bedford 2016	Forni 2013	Forman 2004	Breidthardt 2011	Wang 2013
Demographics												
Age	**9**	x	x	x	x	x	x	x	x			x
Male/gender	**3**		x	x			x					
Past history												
CKD or SCr	**10**		x	x	x	x	x	x	x	x	x	x
Diabetes	**5**		x	x				x	x	x		
Heart failure	**4**		x						x	x		x
Liver disease	**3**	x				x			x			
Drugs												
Diuretics	**4**				x	x					x	x
ACEi/ARBs	**3**			x	x	x						
Observations												
Hypotension/ Shock	**3**				x		x					x
Bloods												
Bicarbonate	**4**				x	x	x				x	
↑WCC	**3**					x	x	x				

ACEi – angiotensin-converting enzyme inhibitor drugs, ARB – Angiotensin 2 receptor blocker drugs, CKD = chronic kidney disease, Bloods – laboratory parameters, SCr – serum creatinine, T&O – Trauma and Orthopaedics, WCC – white cell count.

eTable 10 – All predictors included in the 11 models

Field	General Surgery		T&O	General					Heart Failure			Total
Study	Kheterpal 2007	Kheterpal 2009	Bell 2015	Drawz 2008	Matheny 2010	Koyner 2016	Bedford 2016	Forni 2013	Forman 2004	Breidthardt 2011	Wang 2013	
Demographics												
Age	x	x	x	x	x	x	x	x			x	9
Male/gender		x	x			x						3
BMI	x											1
Race					x							1
Past history												
CKD or SCr		x	x	x	x	x	x	x	x	x	x	10
Diabetes		x	x				x	x	x			5
Heart failure		x						x	x		x	4
Liver disease	x				x			x				3
Hypertension		x										1
PVD	x											1
Ascites		x										1
COPD	x											1
Previous admissions						x (ICU)	x					2
Charlson co-morbidity index							x					1
ASA Grade			x									1

Field	General Surgery		T&O	General					Heart Failure			Total
Study	Kheterpal 2007	Kheterpal 2009	Bell 2015	Drawz 2008	Matheny 2010	Koyner 2016	Bedford 2016	Forni 2013	Forman 2004	Breidthardt 2011	Wang 2013	
Drugs												
Diuretics				x	x					x	x	4
ACEi/ARBs			x	x	x							3
NSAIDs				x	x							2
Contrast					x							1
Number of drugs			x									1
Other drugs[1]					x							1
Observations												
Hypotension / Shock				x		x					x	3
Pulse pressure						x						1
Hypertension				x					x			2
Heart rate[2]				x		x						2
Temperature						x						1
Respiratory rate[3]						x		x				2
O2 saturations						x						1
Consciousness[4]						x		x				2
Surgery, Other												
Type of surgery	x	x										2
Emergency	x	x										2
Time						x						1

Field	General Surgery		T&O	General					Heart Failure			Total
Study	Kheterpal 2007	Kheterpal 2009	Bell 2015	Drawz 2008	Matheny 2010	Koyner 2016	Bedford 2016	Forni 2013	Forman 2004	Breidthardt 2011	Wang 2013	
Labs, Diagnosis												
Primary diagnosis							x					1
Haemoglobin						x						1
↓Platelets[5]					x	x						2
CRP							x					1
↑WCC					x	x	x					3
Bacterial infection[6]					x							1
MI[7]					x		x					2
Rhabdomyolysis[8]					x							2
Hepatitis/AST[9]					x	x						2
Alk Phosphatase						x						1
Bilirubin						x						1
Pancreatitis[10]					x							1
Bicarbonate				x	x	x				x		4
Anion gap						x						1
BUN/Cr						x						1
Urea				x		x						2
↑Ca^{2+}/Ca^{2+}[11]					x	x						2
↑glucose					x	x						2
Magnesium							x					1
Potassium						x	x					2
↓Na^{+}/Na^{+}[12]						x					x	2
Albumin				x		x						2
Total protein						x						1
Proteinuria							x				x	2

ACEi – angiotensin-converting enzyme inhibitor drugs, ARB – Angiotensin 2 receptor blocker drugs, ASA grade – American Anesthesiology Association grade, CKD = chronic kidney disease, COPD – Chronic obstructive pulmonary disease, NSAIDs – non-steroidal anti-inflammatory drugs, PVD – peripheral vascular disease. [1]Other drugs: Aminoglycoside, Amphotericin B, Cyclosporine, Acyclovir, Cisplatin, [2]= ≥70/min, [3]= Respiratory rate ≥20/min, [4]= not alert on AVPU scale (best response: Alert, to Voice, Pain, Unresponsive), [5] =platelet count <75% lower limit of normal, [6] =acute use of antibiotics, [7] =elevated CK-MB or Troponin-I or T, [8] = increase CK x5 in absence of myocardial infarction, [9] =peak AST or ALT >400IU/L, [10] =>x3 lipase normal range, [11] Serum Ca^{2+} =>upper limit normal, Na – serum Sodium, [12] =serum Sodium <130mmol/L, *SCr. T&O – Trauma and Orthopaedics. NB Koyner et al did not specify ranges for the predictors included.

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	Title page
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	Page 2-3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	5-6
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	7-8
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	NA
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	8, online supplementary eTable 1
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	7
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	online supplementary eTable 2-4
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	7-8
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	8-9, online supplementary eTable 5
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	8-9 online supplementary eTable
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	9
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	NA

PRISMA 2009 Checklist

Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency ($e.g., I^2$) for each meta-analysis.	Table 1

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	online supplementary eTables 6, 8
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	NA
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	10, Figure 1,
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	supplementary eTables 6
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	supplementary eTables 6, 8
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	supplementary eTable 6
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	NA
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	NA
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	NA
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	13-15
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	supplementary eTable 6, 8, pg.14
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	15-18
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	NA

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097.

PRISMA 2009 Checklist

doi:10.1371/journal.pmed1000097 For more information, visit: www.prisma-statement.org.

BMJ Open

A systematic review of prognostic prediction models for acute kidney injury (AKI) in general hospital populations.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-016591.R1
Article Type:	Research
Date Submitted by the Author:	07-Jun-2017
Complete List of Authors:	Hodgson, Luke; University of Southampton, Faculty of Medicine; Western Sussex Hospitals NHS Trust, Intensive Care Sarnowski, Alexander; Royal Surrey County Hospital NHS Foundation Trust Roderick, Paul; University of Southampton, Faculty of Medicine Dimitrov, Borislav; University of Southampton, Primary Care and Population Sciences Venn, Richard; Western Sussex Hospitals NHS Trust, Intensive Care Forni, Lui; University of Surrey Faculty of Health and Medical Sciences; Royal Surrey County Hospital NHS Foundation Trust
Primary Subject Heading:	Renal medicine
Secondary Subject Heading:	Research methods
Keywords:	Nephrology < INTERNAL MEDICINE, Acute renal failure < NEPHROLOGY, STATISTICS & RESEARCH METHODS

Title Page

Title: A systematic review of prognostic prediction models for acute kidney injury (AKI) in general hospital populations.

Corresponding author: Dr Luke E Hodgson, Academic Unit of Primary Care and Population Sciences, Faculty of Medicine, University of Southampton, Southampton General Hospital, Tremona Road, Southampton, SO16 6YD.

Email: drlhodgson@gmail.com Telephone: +44 7734883040

Authors:

Luke E Hodgson,[1,2] Intensive Care Registrar,

Alexander Sarnowski[3], Core Medical Trainee,

Paul J Roderick[1], Professor of Public Health,

Borislav D Dimitrov[1], Associate Professor Medical Statistics†,

Richard M Venn[2], Consultant Intensivist and,

Lui G Forni[3,4], Professor of Intensive Care Medicine.

1. Academic Unit of Primary Care and Population Sciences, Faculty of Medicine, University of Southampton, Southampton General Hospital, Tremona Road, Southampton, SO16 6YD, UK.
2. Western Sussex Hospitals NHS Foundation Trust, Anaesthetics Department, Worthing Hospital, Lyndhurst Rd, Worthing, BN11 2DH, UK.
3. The Royal Surrey County Hospital NHS Foundation Trust, Egerton Road, Guildford, GU2 7XX, UK.
4. Faculty of Health and Medical Sciences, University of Surrey, Guildford, Surrey, GU2 7XH, UK.

† Deceased 15th January 2017.

Word count main text: 3717

Abstract

Objective Critically appraise prediction models for hospital-acquired AKI (HA-AKI) in general populations.

Design Systematic review.

Data sources Medline, EMBASE & Web of Science until November 2016.

Eligibility Studies describing development of a multivariable model for predicting HA-AKI in non-specialised adult hospital populations. Published guidance followed for data extraction reporting and appraisal.

Results 14046 references were screened. Of 53 HA-AKI prediction models, 11 met inclusion criteria (general medicine and/or surgery populations, 474478 patient episodes), five externally validated. The most common predictors were age (n=9 models), diabetes (5), admission serum creatinine (SCr) (5), chronic kidney disease (CKD) (4), drugs (diuretics, 4 and/or Angiotensin-converting enzyme inhibitors/Angiotensin-receptor blockers, 3), bicarbonate and heart failure (4 models each). Heterogeneity was identified for outcome definition. Deficiencies in reporting included handling of predictors, missing data and sample size. Admission SCr was frequently taken to represent baseline renal function. Most models were considered at high risk of bias. Area under the receiver operating characteristic curves to predict HA-AKI ranged 0.71-0.80 in derivation (reported in 8/11 studies), 0.66-0.80 for internal validation studies (n=7) and 0.65-0.71 in 5 external validations. For calibration the Hosmer-Lemeshow test or a calibration plot were provided in 4/11 derivations, 3/11 internal and 3/5 external validations. A minority of the models allow easy bedside calculation and potential electronic automation. No impact analysis studies were found.

Conclusions AKI prediction models may help address shortcomings in risk assessment, however, in general hospital populations few have external validation. Similar predictors reflect an elderly demographic with chronic co-morbidities. Reporting deficiencies mirror prediction research more broadly, with handling of SCr (baseline function and use as a predictor) a concern. Future research should focus on validation, exploration of electronic linkage and impact analysis. The later could combine a prediction model with AKI alerting to address prevention and early recognition of evolving AKI.

Key words: acute kidney injury, clinical prediction models, systematic review

Summary

Strengths

- This is the first systematic review of prediction models for hospital-acquired AKI (HA-AKI) in general hospital populations who account for the majority of hospital admissions and AKI cases.
- The models were selected following an extensive literature search; the review followed the latest critical appraisal guidance and assessed validity of the models in terms of risk of bias and applicability, highlighting important shortcomings such as handling of serum creatinine.
- The large number of patient episodes provides important insights into AKI prediction and complements other recent reviews in specialised areas (cardiac surgery, CI-AKI and liver transplantation).

Weaknesses

- Lack of access to individual participant data (IPD) prevented a meta-analysis of the studies, an avenue of future research.
- The small number of externally validated models and absence of impact analysis limits recommendation and implementation of an individual model.

Introduction

Acute kidney injury (AKI) is defined as an acute increase in serum creatinine (SCr) or reduction in urine volume.[1] The incidence of AKI is increasing, affecting up to one in five hospitalised adults worldwide.[2] A continuum of injury exists long before sufficient loss of excretory kidney function can be measured with standard laboratory tests (i.e. SCr).[3 4] Associated mortality remains high, in part reflecting the severity of the underlying disease, but may also be due to the limitations of conventional markers to detect early injury.[5]

Deficits in recognition and management of patients with AKI,[6] has led to practice guidance calling for improved risk assessment, at which point interventions could be most beneficial.[7] One suggested strategy to achieve this aim is through the implementation of clinical prediction models.[8 9] Though development and validation of AKI prediction models is desirable,[7 10] clinical application in this and other fields has been hampered for a number of reasons:

- Potential predictors and models continuously increase with new studies often finding conflicting results,[11]
- Substandard reporting of methodology and results make conclusions problematic,[12 13]
- Few general hospital population studies exist - specialist fields (cardiac and transplant surgery and contrast-induced [CI-AKI]) account for the majority of AKI models and all systematic reviews, but are unlikely to be generalisable and,[14-17]
- Models rarely enable electronic automation as part of clinical workflow, known to influence uptake.[18]

High quality systematic reviews of prediction models have been called for.[19] Following recent reporting guidance (CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies, CHARMS[20] and Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis, TRIPOD),[12] this review appraises hospital-acquired AKI (HA-AKI) prediction models in general populations, who in the UK account for the majority of hospital episodes,[21] and AKI cases.[22 23]

Methods

Published guidance (CHARMS, TRIPOD and Preferred reporting items for systematic reviews and meta-analyses, PRISMA) helped frame the review question, data extraction, reporting and appraisal.[12 20 24] The research question was: what are the available prognostic prediction models for the development of HA-AKI in adult general populations? Using explicit, systematic methods to minimise bias and provide reliable findings from which conclusions can be drawn and decisions made,[25 26] the review aimed to collate empirical evidence for AKI prediction models across general hospital settings, fitting pre-specified eligibility criteria (online supplementary file eTable 1). Performance was assessed by discrimination and calibration including validation studies. The presence of any impact analysis studies was also investigated. The review aimed to provide recommendations for the most robust, usable models, including the ability to incorporate future electronic data linkage, for example between the community (primary care) and hospital.

Data sources, study selection and data extraction

We searched MEDLINE, Embase & Web of Science databases (inception to November 2016) using recommended filters (online supplementary file eTables 2-4).[27 28] Titles and abstracts were screened by two reviewers (LH, AS) and full articles reviewed if eligible. Disagreements were resolved by iterative screening rounds. Reference lists from retrieved articles, systematic reviews, National[7] and International guidance[1] and our own literature files were also analysed. Data extraction, and quality assessment was performed by two investigators (LH, AS) with disagreements resolved by a third reviewer

(LF). A data extraction form was used based on previous reviews and guidance (summary online supplementary file eTable 5).[12 13 20] Items extracted included design (eg, cohort, case-control), population, location, outcome (definition duration of follow-up, blinding of assessment), modelling method (eg, logistic), method of internal validation (eg, bootstrapping), number of participants and events, number and type of predictors, model presentation, and predictive performance (calibration, discrimination). The presence of external validation was recorded.

Outcome, model performance and clinical utility

It was anticipated that study outcome, HA-AKI, would vary given the numerous definitions in use prior to KDIGO in 2012.[1] Thus, during the search strategy we included studies with a SCr around admission and repeated during a hospital admission to diagnose HA-AKI. Information was gathered on how a study defined a patients baseline renal function, how community AKI cases were handled, whether SCr was used as a predictor in analysis and finally the magnitude and timeframe used to define the outcome. Discrimination and calibration are the most common methods to assess model performance. Discrimination is usually assessed graphically by the area under the receiver operating characteristic curve (AUROC), representing how well a model separates and ranks patients who experienced the outcome, from those who did not. For prediction models, the AUROC, which focuses solely on accuracy has a number of shortcomings, such as a lack of information on consequences and when used in populations where the outcome prevalence is rare.[29 30] Calibration describes how well predicted

results agree with observed results.[12 30] The Hosmer-Lemeshow (H-L) test, despite limitations, is the most commonly used calibration statistic.[31 32] It is also recommended to graphically plot expected and actual outcomes, for example, with a calibration slope.[12] In addition to performance, ease of bedside use and whether the models could be electronically automated - factors known to influence successful uptake - were recorded.[18] A quantitative synthesis of the models was not performed, being beyond the scope of review and formal methods for meta-analysis of prediction models are yet to be fully developed.

Study quality assessment

A global TRIPOD score for each study was calculated to quantify reporting, consisting of the sum of the scores for each individual item (out of a maximum 37, with a score of 1 for criterion met, score of 0 for each item not met, or unclear).[12] As yet there has been no suggested cut-off for what represents a high quality study, though it would be reasonable to judge that those studies with the most significant gaps in reporting are likely to be at higher risk of bias. Furthermore the quality (risk of bias) of each study was assessed by piloting a version of PROBAST (Prediction study Risk Of Bias Assessment Tool), a tool for assessing risk of bias and applicability of prognostic model studies, nearing completion and ready for piloting when this review was undertaken (Wolff R, Whiting P, Mallett S, et al, personal communication, website: http://s371539711.initial-website.co.uk/probast/). Elements were considered in the following domains: study participants, predictors, outcome, sample size

and missing data, statistical analysis and overall judgement of bias and applicability.

Patient involvement

Patients were not involved in setting the research question, outcome, design and implementation of the study. There are no plans to involve patients in dissemination.

Results

From 14046 articles identified by the search strategy, 254 full articles were reviewed (PRISMA flow chart, figure 1). Specialised fields (predominantly cardiac surgery, transplantation or CI-AKI) accounted for 61 of 74 (82%) of all studies. This review included eleven general model studies (n=474478 patient episodes), in General Surgery,[33 34] Trauma and Orthopaedics (T&O),[35] General Hospital cohorts (predominantly Medicine and Surgery),[36-40] and Heart Failure (summarised in table 1 and online supplementary file eTable 6, with abbreviations in eTable 7).[41-43] Two further studies were purely external validations.[44 45] HA-AKI incidence was 7% (21641 events), though this varied from <1% in the General Surgery models,[33 34] to 28% across the Heart Failure studies and heterogeneous definitions (timeframe and marker) were employed (see table 1 for definitions used with further information in eTable 6). For example, five studies took admission SCr to represent a patients baseline, potentially confusing CKD, established and emerging AKI.[34 38 40 42 43] Of note one study produced a model to predict admission AKI as well as HA-AKI at 72 hours with the former not considered suitable for analysis in this review.[39]

In seven of the nine studies reporting age, this was significantly higher in the group with the outcome, with eight studies reporting a mean or median age over 65 years in the outcome group (table 1). Mortality was significantly higher in those who developed the outcome in the six studies where data were available (ranging 6-42%). No impact analyses were retrieved.

Study reporting

A median 28 (interquartile range 25-30) of 37 recommended items were reported, suggesting significant shortcomings (key shortcomings are summarised in table 2 with TRIPOD reporting summarised in online supplementary file eTable 8). By design, eight studies were retrospective, two prospective and one was a case control. Five studies were single-centre. The USA (n=6) and UK (n=3) accounted for the majority of the models. Only three studies used imputation techniques for missing data.[34 35 38] Definitions were heterogenous (table 1) with five using RIFLE,[37] AKIN,[42] or KDIGO criteria for changes in SCr.[35 36 39] One study used KDIGO SCr change within a 24-hour timeframe of predictors being measured.[40]

Candidate predictors, model building and sample size

A median of 29 (interquartile range 19-35) predictors were considered, though frequently studies only reported those significant on univariate or multivariate analysis. Blinding of assessment of predictors and study outcome was not mentioned. Continuous predictors were dichotomized in four studies and ten studies used univariate analysis to select for multivariate analysis. No models mentioned shrinkage techniques or sample size calculations. Median number of outcome events was 271 (121-672). For statistical power, all of the studies had more than ten events per predictor (EPP) *included in the model*. However the EPP was <10 in six studies, when accounting for the total number of candidate predictors assessed.[33 36 38 39 41 43] Of a total of 56 different predictors a median of 7 (7-12) were included per model, including demographics, past history, procedure information, laboratory parameters, physiological observations and hospital admission diagnoses (most common

presented in figure 2, full details online eTables 9-10). Only 4 studies included physiological parameters in their final model.[36 38 40 43] Seven studies included admission SCr as potential predictor with five including this in the final model, thus potentially confusing prediction with a diagnosis of AKI.[34 37 40 42 43] Each study's handling of SCr in terms of when a baseline was calculated (prior or at admission) and whether SCr was used as a predictor are summarised in eTable 11.

Model performance (table 1)

Median AUROC (or C-Statistic) was 0.745 (range 0.71-0.80) for derivation (eight studies) and 0.74 (range 0.66-0.80) for internal validations reporting discrimination (seven studies). Excluding the studies using non-consensus based definitions and those including admission SCr as predictor and/or baseline, left only four studies.[35 36 39 41] In these studies AUROCs ranged 0.71-0.74 in derivation (three studies), 0.67-0.76 for internal validation (three studies) and 0.65-0.71 for external validation (three studies). Only one model study presented a calibration plot for derivation and validation.[35] The H-L statistic was used in three derivations,[36 37 42] and two internal validations.[39 42]

Five models have been externally validated: on separate populations within the same study;[35 39] other model studies;[43] or stand alone external validations,[33 45] where the AUROCs were moderate, ranging 0.65-0.71. One validation provided a calibration plot,[35] one the H-L statistic,[39] and one reported both.[45] In the Bell external validation cohort calibration

suggested the model over-predicted the outcome requiring recalibration.[35] In the external validation of the Forni study calibration plots showed agreement at low probability rates whilst at higher rates calibration deviated in the medical cohort.[45] Two of the three surgical models have been externally validated: the Kheterpal model,[33] in a Chinese population (AUROC 0.66),[44] and the UK T&O study used a third centre for external validation.[35] Two of five mixed general population models have external validation,[36 39] the later having been derived on medical patients and externally validated in medical and surgical cohorts.[45] The first of the three heart failure studies was externally validated in the subsequent studies with inferior discrimination (AUROC 0.65 in both validations).[41-43] No model updating was reported.

Quality assessment and risk of bias summary

Quality assessment based on a draft version of the PROBAST tool. This suggested evidence in 9 of the 11 included studies of a high risk of bias (summarized in table 3) with shortcomings across the major domains of the assessment. For example, one study used a case-control design which is inappropriate for developing a prediction model as it does not enable calculation of absolute risks and thus yields incorrect estimates of model intercept or baseline hazard.[20] A wide variety of predictors were considered with use of univariate analysis to select for multivariate in 10/11 of the studies. Six studies were potentially underpowered having less than ten EPP assessed. Seven of the studies introduced potential bias in handling of renal function and SCr either in failing to establish a reliable baseline renal function,

excluding patients with reduced renal function, or employing it as a predictor. Finally, outcome definition frequently varied in part owing to a number of the studies preceding consensus definitions.

Discussion

Principal findings

In this first systematic review of HA-AKI prediction in general hospital settings, the most common predictors were age, diabetes, CKD, drugs, heart failure and serum creatinine and bicarbonate. Modest discrimination performance of all the models is unsurprising when attempting at a single time point to predict a future event reflecting diverse aetiologies, affecting heterogeneous patient groups. Significant shortcomings mirror those described elsewhere:[13 46-48]

- Multiple similar models, rarely externally validated,
- No impact analysis or evidence of clinical implementation,
- Incomplete reporting and,
- Little consideration of electronic automation (allowing presentation without additional data input beyond usual clinical care), which influences uptake.[18]

Methodological and reporting shortcomings in the studies (summarised in table 2) included 6 studies having less than 10 EPP potentially leading to overfitting, with only 3 employing multiple imputation to handle missing data which can increase sample size and power.[12 49 50]

Handling of SCr and CKD was of particular concern in a number of areas. First, in part due to a previous lack of a consensus definition, the outcome in question, HA-AKI, had heterogeneous definitions, both in magnitude of SCr rise and time-frame. For example, the Kheterpal study (2009)[34] used a rise in SCr ≥177 μmol/L which has been shown to significantly underestimate

rates of AKI when compared with more recent definitions.[51] Koyner et al used a rolling timeframe of 24 hours whilst others used SCr elevation at any point during an admission.[40] Indeed one study produced a separate model to predict AKI at admission to hospital.[39] This was further confused by seven studies inclusion of admission SCr as a potential predictor, (with inclusion in five models), five studies taking admission SCr to represent a patient's baseline and two studies excluding all patients with a reduced admission eGFR from their analysis. This risks confusing prediction and detection of AKI events. Issues with differing definitions have been described before in systematic reviews of prediction models and should be considered when researchers embark on future studies.[52 53]

A formal risk of bias assessment (PROBAST) suggested the majority of studies had domains placing the studies at high risk of bias. Published after TRIPOD, Bell and colleagues' model provides researchers with a good template for adherence to reporting guidance, with a low risk of bias and demonstrates the utility of data linkage (for example between community and hospital), though lack of validation in other populations tempers recommendation for implementation.[35]

Strengths and limitations of this review

This review summarises the currently available AKI prediction models in general populations who account for the majority of hospital admissions and AKI cases.[21-23] The models were selected following an extensive literature

search and the review employed the most recent critical appraisal guidance and risk of bias assessment.[12 20] The large number of patient episodes provides important insights into AKI prediction complementing other recent reviews in cardiac surgery, CI-AKI, liver transplantation and non-cardiac surgery.[14-17] In-patient mortality in those who developed the outcome ranged 6-42% (in the six studies reporting mortality) emphasising this is a crucial group to promptly identify.

The first limitation is the small number of externally validated models, which tempers recommending one model over another. Second, though we aimed to include general populations, caution should be employed, for example, when comparing a model derived on Heart Failure patients to one from an Orthopaedic cohort. However, in many UK hospitals, such populations share similarities (predominantly elderly demographic with co-morbidities) and if one aim of a prediction model is generalizability, a model should be tested in these different fields. Third, as study outcome definitions and handling of SCr (baseline and as predictor of outcome) were heterogenous, model comparisons are problematic, though recent studies were more likely to use KDIGO SCr change. Fourth, no studies included urine output, probably reflecting the small number of patients who have this marker closely monitored. Fifth, TRIPOD recommendations were used as a reporting benchmark, however, the relative importance of individual items and what constitutes an acceptable ‘score’ is arguable, though a formal risk of bias assessment was also carried (PROBAST) providing further insight into

respective study strengths and weaknesses. The absence of impact analysis limits the recommendation of one model over another. Finally, a meta-analysis was not performed without access to individual participant data (IPD). Expert guidance now exists in this area and offers opportunities to improve the scope of external validation research.[53 54]

Comparison with previous systematic reviews

Both this study and a review of CI-AKI models found pre-existing predictors - age, CKD, diabetes and heart failure to be the most commonly included.[15] A Cardiac surgery review reported specialty specific predictors in addition to these chronic co-morbidities. A non-Cardiac surgery review (5 of 6 studies in liver transplantation or resection) reported age, CKD and diabetes in at least two models.[17] Finally, a liver transplantation review highlighted the importance of CKD and (unsurprisingly) liver dysfunction.[16] The present review found drugs or acute laboratory values frequently included, though only 5 models included acute physiological parameters. Our study and the non-Cardiac surgery review included adherence to recommended TRIPOD reporting with similar shortcomings. Across the other reviews, only in the fields of CI-AKI and Cardiac surgery were external validations reported.[14 15] Ease of use (including if necessary a calculator) and potential for electronic automation were rarely considered across the models reviewed. No impact analysis studies have been described.

Future directions

Management of HA-AKI presents a significant challenge, that could be helped by robust prediction models to risk stratify, encourage prevention and prompt recognition, key healthcare priorities.[6 10] Appraisal and synthesis of prediction studies may enable clinicians and policymakers judge model utility however, this is problematic when key study details are not reported.[12] Though much of the AKI literature is on (often assumed) hospital-acquired AKI, the majority of cases arise from the community (CA-AKI).[55 56] Indeed, a recent study demonstrated a significant proportion of such patients are never hospitalised.[57] This review suggests even in HA-AKI, the strongest predictors are pre-existing patient factors. The two laboratory measures frequently included – serum creatinine and bicarbonate – may also reflect a chronic component. It is likely a proportion of cases classed as HA-AKI represent (evolving) community cases, thus, models using such pre-existing risk factors makes clinical sense. This continuum of harm between community and hospital could suggest that a risk prediction model in place at, or even *before* hospital admission, combined with early flagging of those who have met AKI criteria, may be required to improve outcomes.

Electronic linkage of patient records between community and hospital data is desirable to ensure accurate inclusion of predictors (chronic morbidity, medication, laboratory and physiological parameters). This may also enable bedside automation as part of clinical workflow, where there is evidence beneficial implementation can be achieved.[18 58] Acute physiological

parameters assessed as predictors in seven studies and subsequently included in only five studies, could be an avenue of future research to improve the modest performance of all models at a single time point (admission to hospital) described to date. As hospitals increasingly employ electronic track and trigger observation systems this may then enable the application of complex statistics (e.g. machine learning) to account for the effects of trends and repeated measures. Risk stratification using chronic comorbidity and medication(s) with trends in physiology, could be further enhanced by measurement of urine output and/or newer biomarkers. Unfortunately, to date such research has not been published, with reliance on using retrospective databases often only providing information at a single time point. A future study in this area would thus require prospective collection of rich data, with the aim to achieve accurate prediction modeling demanded by clinicians and patients prior to implementation.

Impact analysis in prediction research is sparse making it difficult to conclude whether a model is worth implementing alongside, or replacing, usual care.[59] This is important as for example, one study suggested clinical acumen may be superior to prediction models,[60] whilst another found the combination of a model with clinical acumen was better than either alone.[61] Some impact analyses have suggested benefit, but conclusions are limited due to their rarity and design (mostly before-after without control).[62] There are a number of potential areas for impact analysis and clinical implementation (summarised in table 4). First, in specific populations a model could influence location of peri-operative care of surgical patients or drug

and/or contrast dosing in patients with heart failure. Second, in a wider hospital setting the effects of highlighting those at highest risk to teams (ward, outreach critical care or Nephrology) with an adequate effector arm could be investigated. This has been demonstrated by existing AKI alerts in *established* AKI where outcome benefit has been limited to patients who had best practice delivered.[63-65] Third, as healthcare embraces complex technology, the inclusion of physiological (including urine output) or laboratory trends may be the only way to significantly improve model performance. Fourth, a model could identify a high risk group to be further risk stratified by employing one of the (increasing number of) available renal biomarkers,[66] or response to an intervention such as a frusemide stress test.[67] Finally, one external validation study found those patients high risk on the prediction model who did develop AKI had a higher rate of mortality than the low risk group who developed HA-AKI, indicating the model predicts disease severity.[45] This could allow early review of such patients to help inform whether escalation of care may be required, or indeed be appropriate in the increasing number of frail elderly patients admitted to hospitals.

To conclude, improving the management of patients to prevent AKI, or reduce associated complications, is a global health priority. This systematic review suggests there are few externally validated prediction models to help identify those at risk of AKI across general hospital populations. Future research should concentrate on validation, utility of additional markers, exploration of electronic implementation to enable clinical uptake and impact analysis.

Table 1 Summary of HA-AKI prediction models.

Population	General Surgery		T&O	General (Medical & Surgical)					Heart failure		
Author, year (n=derivation)	**Kheterpal** 2007 (n=14,066)	**Kheterpal** 2009 (n=57,080)	**Bell** 2015 (n=6,220)	**Drawz** 2008 (n=360)	**Matheny** 2010 (n=26,107)	**Koyner** 2016 (n=202,961)	**Bedford** 2016 (n=7,556)	**Forni** 2013 (n=1,867)	**Forman** 2004 (n=1,004)	**Breidthardt 2011** (n=657)	**Wang** 2013 (n=1,010)
Centres, Design	1, R	121, R	3, R	3, CC	1, R	5, R	3, R	1, P	11, R	1, P	1, R
Age (with outcome)	59	65 (±15)	77 (±11)	67	-	70 (±16)	-	80 (70-86)	68.7	79 (72-85)	73 (67-78)
Age (no outcome)	47	54 (±17)	70 (±16)	63	-	63 (±19)	-	73 (61-81)	66.8	79 (70-85)	71 (63-75)
Outcome predicted	eGFR <50 (<7 days)	↑SCr ≥177µmol/l, RRT (30 day)	KDIGO ↑SCr	↑SCr*	RIFLE ↑SCr	KDIGO ↑SCr (24hr)	KDIGO ↑SCr (72hr)	KDIGO ↑SCr (<7 days)	↑SCr >26.5µmol/l**	↑SCr >26.5µmol/l**	AKIN SCr (<48hr)
Events	121	561	672	120	1,352	17,541	222	95	271	136	341
Mortality with outcome	15%	42%	-	-	-	6%	-	20%	27%	17%	17%
Mortality no outcome	3%^	8%^	-	-	-	1%	-	4%	-	6%	2%
Predictors tested	30	19	11	19	23	29	45	25	29	48	35
Predictors included	7	9	7	7	27	29	12	7	4	3	8
EPP	4	30	61	6	59	605	5	4	9	3	10
Inappropriate handling of SCr	X	X		X	X	X			X	X	X
Derivation AUROC	0.77	0.80	0.74	0.73	0.75	-	-	0.72	-	0.71	0.76
IV AUROC	-	0.80	0.73	0.66	-	0.74	0.67	0.76	-	-	0.76
EV AUROC	0.67	X	0.71	X	X	X	0.71	0.65-0.71#	0.65, 0.65	X	X
Derivation Calibration	RR	RR	Plot	-	H-L P=0.29	-		H-L P=0.96	-	-	H-L P=0.98
IV Calibration	-	RR	Plot	RR	-	-	H-L P=0.04	-	-	-	H-L P=0.13
EV Calibration	RR	-	Plot	-	-	-	H-L P=0.12	H-L P=0.06-0.09, Plot	RR	-	-

TRIPOD items reported	25	28	34	26	28	24	29	29	26	23	30
Bedside calculation	-	-	-	-	-	-	-	Yes	Yes	Yes	Yes
Electronic Automation	-	-	Yes***	Yes	-	Yes	-	Yes	Yes	Yes	Yes

Design: R - retrospective, P – prospective, CC – case-control, Mortality - In-hospital. AUROC – area under the receiver operating characteristic curve, Plot – Calibration plot, EPP – Events per predictor, EV – external validation, H-L – Hosmer-Lemeshow test, IV – internal validation, RR – risk range, RRT – renal replacement therapy, SCr – serum creatinine, T&O – Trauma & Orthopaedics, TRIPOD – how many of the 37 recommended items were reported. *Increase sCr ≥44μmol/L if baseline SCr of ≤168μmol/L, ≥88μmol/L baseline 177-433μmol/L & ≥133μmol/L baseline >442μmol/L. **During admission, ***Used linked community and hospital data, ^Propensity matched, [#]validations in medicine/surgery with/without baseline SCr.

Table 2 Summary of limitations in methodology and reporting

Area of concern	Description
Missing data	Multiple imputation recommended to avoid bias, rarely described.[12 50]
Definitions of outcome & predictors	No consistent strategy used to differentiate CA-AKI from HA-AKI. Two studies excluded patients with pre-existing CKD;[33 37] five studies took admission SCr as baseline; five included SCr as predictor despite it forming the outcome; co-morbidities inconsistently defined: including from admission diagnoses or coded history.
Blinding of predictors or outcome	Not reported.
Sample size	Calculations not described, six studies had <10 EPP. Small sample increases risk of overfitting and underfitting.[12 49]
Univariate to select for multivariate analysis	Technique not recommended, used in 10 of 11 models.[12]
Bootstrapping	Adjust for optimism, without losing information - rarely described.[35 43]
Calibration plots	Important part of model performance,[12] present in only one model & one external validation.[35 45]
External validation & model updating	Validation adjusts for optimism, assesses generalizability. but was scarce, whilst model updating is recommended but not described.[12]
Newer performance measures	Techniques such as decision curve analysis offer insight into clinical consequences - not described.[29]
Use of data linkage	Only one study utilised data linkage.[35]

CA-AKI – community-acquired AKI, CKD – chronic kidney disease, EPP – events per predictor, HA-AKI – hospital-acquired AKI, SCr – serum Creatinine.

Table 3. Risk of bias summary based on PROBAST (Prediction study Risk Of Bias Assessment Tool, PROBAST, permission from Wolff R, personal communication).

Population	General Surgery		T&O	General (Medical & Surgical)					Heart failure		
Model, year	**Kheterpal (2007)**	**Kheterpal (2009)**	**Bell (2015)**	**Drawz (2008)**	**Matheny (2010)**	**Koyner (2016)**	**Bedford (2016)**	**Forni (2013)**	**Forman (2004)**	**Breidthardt (2011)**	**Wang (2013)**
Study participants	?	?	+	+	+	?	+	+	?	?	?
Predictors	?	?	?	?	-	?	-	+	-	-	?
Outcome	-	-	+	-	-	-	+	+	-	-	-
Sample size & missing data	-	+	+	-	-	?	?	?	-	-	?
Statistical analysis	-	-	+	-	+	-	-	?	-	-	-
Overall judgement of bias	-	-	+	-	-	-	-	+	-	-	-
Overall judgement of applicability	-	-	?	-	-	-	+	+	-	-	-
Usability of the model	+	+	+	+	+	+	+	+	+	+	+

Study participants domain - design of the included study, and inclusion and exclusion of its participants; Predictors domain - definition, timing, and measurement of predictors (also assesses whether predictors have not been measured and were therefore omitted from the model); Outcome domain - definition, timing, and measurement of predicted outcomes; Sample size and missing data domain - number of participants in the study and exclusions owing to missing data; Statistical analysis domain - methods (eg appropriate presentation of discrimination and calibration). Red = “high”, Green = “low” or Amber = “unclear” risk of bias.

Table 4. Potential areas for future Impact analysis of AKI prediction models

Population	Impact analysis to inform Clinical use
General Surgery	Peri-operative: haemodynamic targets, place of care, drugs, contrast delivery
Trauma & Orthopaedics	
General Populations	Risk stratification of large populations: eg influencing intensity of observations, remote monitoring, application of biomarkers in subgroups at high-risk
Heart failure	Optimise haemodynamic status: diuretic dosing, use/volume of contrast

Legends to Figures

Figure 1 – PRISMA study flow chart

Figure 2 – Predictors most frequently included in the 11 HA-AKI prediction models. ACEi – Angiotensin-converting enzyme inhibitors, ARBs – Angiotensin-receptor blockers, Bloods – laboratory parameters, CKD – chronic kidney disease, ↓HCO_3 – reduced serum bicarbonate, SCr – serum creatinine, ↑WCC - raised white cell count.

Competing interests statement

"All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf and declare that all have no relationships with companies that might have an interest in the submitted work in the previous 3 years; their spouses, partners, or children have no financial relationships that may be relevant to the submitted work; and have no non-financial interests that may be relevant to the submitted work."

All authors had full access to all of the data (including statistical reports and tables) in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported and no important aspects of the study have been omitted.

Contributors: LH, LF, RV BDD and PR developed the idea for the study. LH, AS, LF, BDD and PR were involved in the study conception, preliminary literature review and design of the search strategy and the study protocol. LH, AS and LF were involved in screening and data extraction of papers. All authors reviewed data extraction output. LH drafted the manuscript, which was critically reviewed and approved by all authors.

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

References

1. KDIGO Clinical Practice Guideline for Acute Kidney Injury. Kidney Int 2012;**Suppl 2**(1):1-136 doi: 10.1038/kisup.2012.6published Online First: Epub Date]|.
2. Susantitaphong P, Cruz DN, Cerda J, et al. World incidence of AKI: a meta-analysis. Clin J Am Soc Nephrol 2013;**8**(9):1482-93 doi: 10.2215/cjn.00710113published Online First: Epub Date]|.
3. Bellomo R, Kellum JA, Ronco C. Acute kidney injury. Lancet 2012;**380**(9843):756-66 doi: 10.1016/s0140-6736(11)61454-2published Online First: Epub Date]|.
4. Bellomo R, Kellum J, Ronco C. Acute renal failure: time for consensus. Intensive Care Med 2001;**27**(11):1685-8 doi: 10.1007/s00134-001-1120-6published Online First: Epub Date]|.
5. Bagshaw SM, Gibney RT. Conventional markers of kidney function. Crit Care Med 2008;**36**(4 Suppl):S152-8 doi: 10.1097/CCM.0b013e318168c613published Online First: Epub Date]|.
6. National Confidential Enquiry into Patient Outcome and Death. Adding insult to injury: a review of the care of patients who died in hospital with a primary diagnosis of acute kidney injury (acute renal failure) http://www.ncepod.org.uk/2009report1/Downloads/AKI_report.pdf: NCEPOD, 2009.
7. *National Institute for Health and Care Excellence. Acute Kidney Injury: Prevention, Detection and Management Up to the Point of Renal Replacement Therapy. (Clinical guideline CG169).* London: National Clinical Guideline Centre., 2013.
8. Rabar S, Lau R, O'Flynn N, Li L, Barry P. Risk assessment of fragility fractures: summary of NICE guidance. BMJ 2012;**345**:e3698 doi: 10.1136/bmj.e3698published Online First: Epub Date]|.
9. Goff DC, Jr., Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation 2014;**129**(25 Suppl 2):S49-73 doi: 10.1161/01.cir.0000437741.48606.98published Online First: Epub Date]|.
10. Mehta RL, Cerdá J, Burdmann EA, et al. International Society of Nephrology's 0by25 initiative for acute kidney injury (zero preventable deaths by 2025): a human rights case for nephrology. The Lancet doi: 10.1016/S0140-6736(15)60126-Xpublished Online First: Epub Date]|.
11. Kyzas PA, Denaxa-Kyza D, Ioannidis JP. Almost all articles on cancer prognostic markers report statistically significant results. Eur J Cancer 2007;**43**(17):2559-79 doi: 10.1016/j.ejca.2007.08.030published Online First: Epub Date]|.
12. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015;**162**(1):W1-73 doi: 10.7326/M14-0698published Online First: Epub Date]|.
13. Bouwmeester W, Zuithoff NP, Mallett S, et al. Reporting and methods in clinical prediction research: a systematic review. PLoS Med 2012;**9**(5):1-

12 doi: 10.1371/journal.pmed.1001221published Online First: Epub Date]|.
14. Huen SC, Parikh CR. Predicting acute kidney injury after cardiac surgery: a systematic review. Ann Thorac Surg 2012;**93**(1):337-47 doi: 10.1016/j.athoracsur.2011.09.010published Online First: Epub Date]|.
15. Silver SA, Shah PM, Chertow GM, Harel S, Wald R, Harel Z. Risk prediction models for contrast induced nephropathy: systematic review. BMJ 2015;**351**:h4395 doi: 10.1136/bmj.h4395published Online First: Epub Date]|.
16. Caragata R, Wyssusek KH, Kruger P. Acute kidney injury following liver transplantation: a systematic review of published predictive models. Anaesth Intensive Care 2016;**44**(2):251-61
17. Wilson T, Quan S, Cheema K, et al. Risk prediction models for acute kidney injury following major noncardiac surgery: systematic review. Nephrol Dial Transplant 2016;**31**(2):231-40 doi: 10.1093/ndt/gfv415published Online First: Epub Date]|.
18. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 2005;**330**(7494):765 doi: 10.1136/bmj.38398.500764.8Fpublished Online First: Epub Date]|.
19. Altman DG. Systematic reviews of evaluations of prognostic variables. BMJ 2001;**323**(7306):224-8
20. Moons KG, de Groot JA, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Med 2014;**11**(10):e1001744 doi: 10.1371/journal.pmed.1001744published Online First: Epub Date]|.
21. Hospital Episode Statistics Analysis HaSCIC. Hospital Episode Statistics: Admitted patient care - 2014-15. http://www.hscic.gov.uk/pubs/hes1415: Health and Social Care Information Centre, 2015.
22. Selby NM, Crowley L, Fluck RJ, et al. Use of electronic results reporting to diagnose and monitor AKI in hospitalized patients. Clin J Am Soc Nephrol 2012;**7**(4):533-40 doi: 10.2215/cjn.08970911published Online First: Epub Date]|.
23. Porter CJ, Juurlink I, Bisset LH, Bavakunji R, Mehta RL, Devonald MA. A real-time electronic alert to improve detection of acute kidney injury in a large teaching hospital. Nephrol Dial Transplant 2014;**29**(10):1888-93 doi: 10.1093/ndt/gfu082published Online First: Epub Date]|.
24. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;**6**(7):e1000097 doi: 10.1371/journal.pmed.1000097published Online First: Epub Date]|.
25. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. JAMA 1992;**268**(2):240-8
26. Oxman AD, Guyatt GH. The science of reviewing research. Ann N Y Acad Sci 1993;**703**:125-33; discussion 33-4

27. Geersing GJ, Bouwmeester W, Zuithoff P, Spijker R, Leeflang M, Moons KG. Search filters for finding prognostic and diagnostic prediction studies in Medline to enhance systematic reviews. PLoS One 2012;**7**(2):e32844 doi: 10.1371/journal.pone.0032844published Online First: Epub Date]|.
28. Wilczynski NL, McKibbon KA, Walter SD, Garg AX, Haynes RB. MEDLINE clinical queries are robust when searching in recent publishing years. J Am Med Inform Assoc 2013;**20**(2):363-8 doi: 10.1136/amiajnl-2012-001075published Online First: Epub Date]|.
29. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;**26**(6):565-74 doi: 10.1177/0272989x06295361published Online First: Epub Date]|.
30. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007;**115**(7):928-35 doi: 10.1161/circulationaha.106.672402published Online First: Epub Date]|.
31. Hosmer DW, Hjort NL. Goodness-of-fit processes for logistic regression: simulation results. Stat Med 2002;**21**(18):2723-38 doi: 10.1002/sim.1200published Online First: Epub Date]|.
32. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited. Crit Care Med 2007;**35**(9):2052-6 doi: 10.1097/01.ccm.0000275267.64078.b0published Online First: Epub Date]|.
33. Kheterpal S, Tremper KK, Englesbe MJ, et al. Predictors of postoperative acute renal failure after noncardiac surgery in patients with previously normal renal function. Anesthesiology 2007;**107**(6):892-902 doi: 10.1097/01.anes.0000290588.29668.38published Online First: Epub Date]|.
34. Kheterpal S, Tremper KK, Heung M, et al. Development and validation of an acute kidney injury risk index for patients undergoing general surgery: results from a national data set. Anesthesiology 2009;**110**(3):505-15 doi: 10.1097/ALN.0b013e3181979440published Online First: Epub Date]|.
35. Bell S, Dekker FW, Vadiveloo T, et al. Risk of postoperative acute kidney injury in patients undergoing orthopaedic surgery-development and validation of a risk score and effect of acute kidney injury on survival: observational cohort study. BMJ 2015;**351**:h5639 doi: 10.1136/bmj.h5639published Online First: Epub Date]|.
36. Forni LG, Dawes T, Sinclair H, et al. Identifying the patient at risk of acute kidney injury: a predictive scoring system for the development of acute kidney injury in acute medical patients. Nephron Clin Pract 2013;**123**(3-4):143-50 doi: 10.1159/000351509published Online First: Epub Date]|.
37. Matheny ME, Miller RA, Ikizler TA, et al. Development of inpatient risk stratification models of acute kidney injury for use in electronic health records. Med Decis Making 2010;**30**(6):639-50 doi: 10.1177/0272989X10364246published Online First: Epub Date]|.
38. Drawz PE, Miller RT, Sehgal AR. Predicting hospital-acquired acute kidney injury--a case-controlled study. Ren Fail 2008;**30**(9):848-55 doi: 10.1080/08860220802356515published Online First: Epub Date]|.
39. Bedford M, Stevens P, Coulton S, et al. *Development of risk models for the prediction of new or worsening acute kidney injury on or during hospital*

admission: a cohort and nested study. Southampton UK: Queen's Printer and Controller of HMSO., 2016.
40. Koyner JL, Adhikari R, Edelson DP, Churpek MM. Development of a Multicenter Ward-Based AKI Prediction Model. Clin J Am Soc Nephrol 2016;**11**(11):1935-43 doi: 10.2215/cjn.00280116published Online First: Epub Date]|.
41. Breidthardt T, Socrates T, Noveanu M, et al. Effect and clinical prediction of worsening renal function in acute decompensated heart failure. Am J Cardiol 2011;**107**(5):730-5 doi: 10.1016/j.amjcard.2010.10.056published Online First: Epub Date]|.
42. Wang YN, Cheng H, Yue T, Chen YP. Derivation and validation of a prediction score for acute kidney injury in patients hospitalized with acute heart failure in a Chinese cohort. Nephrology (Carlton) 2013;**18**(7):489-96 doi: 10.1111/nep.12092published Online First: Epub Date]|.
43. Forman DE, Butler J, Wang Y, et al. Incidence, predictors at admission, and impact of worsening renal function among patients hospitalized with heart failure. J Am Coll Cardiol 2004;**43**(1):61-7
44. Xing XZ, Wang HJ, Huang CL, et al. Two acute kidney injury risk scores for critically ill cancer patients undergoing non-cardiac surgery. World J Emerg Med 2012;**3**(4):278-81 doi: 10.5847/wjem.j.1920-8642.2012.04.007published Online First: Epub Date]|.
45. Hodgson L, Dimitrov, BD, Roderick, PJ, Venn, R, LG, Forni. Predicting AKI in Emergency Admissions: An external validation study of the acute kidney injury prediction score (APS). Bmj Open 2017;**In Press**
46. Noble D, Mathur R, Dent T, Meads C, Greenhalgh T. Risk models and scores for type 2 diabetes: systematic review. BMJ 2011;**343**:d7163 doi: 10.1136/bmj.d7163published Online First: Epub Date]|.
47. Collins GS, Mallett S, Omar O, Yu LM. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. BMC Med 2011;**9**:103 doi: 10.1186/1741-7015-9-103published Online First: Epub Date]|.
48. Mallett S, Royston P, Dutton S, Waters R, Altman DG. Reporting methods in studies developing prognostic models in cancer: a review. BMC Med 2010;**8**:20 doi: 10.1186/1741-7015-8-20published Online First: Epub Date]|.
49. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. J Clin Epidemiol 1995;**48**(12):1503-10
50. Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. J Clin Epidemiol 2006;**59**(10):1087-91 doi: 10.1016/j.jclinepi.2006.01.014published Online First: Epub Date]|.
51. Bihorac A, Brennan M, Baslanti TO, et al. NATIONAL SURGICAL QUALITY IMPROVEMENT PROGRAM UNDERESTIMATES THE RISK ASSOCIATED WITH MILD AND MODERATE POSTOPERATIVE ACUTE KIDNEY INJURY. Critical care medicine 2013;**41**(11):10.1097/CCM.0b013e31829860fc doi: 10.1097/CCM.0b013e31829860fcpublished Online First: Epub Date]|.

52. Ensor J, Riley RD, Moore D, Snell KI, Bayliss S, Fitzmaurice D. Systematic review of prognostic models for recurrent venous thromboembolism (VTE) post-treatment of first unprovoked VTE. BMJ Open 2016;**6**(5):e011190 doi: 10.1136/bmjopen-2016-011190published Online First: Epub Date]|.
53. Debray TPA, Damen JAAG, Snell KIE, et al. A guide to systematic review and meta-analysis of prediction model performance. BMJ 2017;**356** doi: 10.1136/bmj.i6460published Online First: Epub Date]|.
54. Riley RD, Ensor J, Snell KIE, et al. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. BMJ 2016;**353** doi: 10.1136/bmj.i3140published Online First: Epub Date]|.
55. Wonnacott A, Meran S, Amphlett B, Talabani B, Phillips A. Epidemiology and outcomes in community-acquired versus hospital-acquired AKI. Clin J Am Soc Nephrol 2014;**9**(6):1007-14 doi: 10.2215/cjn.07920713published Online First: Epub Date]|.
56. 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**(9):1510-8 doi: 10.2215/cjn.02140215published Online First: Epub Date]|.
57. Sawhney S, Fluck N, Fraser SD, et al. KDIGO-based acute kidney injury criteria operate differently in hospitals and the community-findings from a large population cohort. Nephrol Dial Transplant 2016;**31**(6):922-9 doi: 10.1093/ndt/gfw052published Online First: Epub Date]|.
58. Kannry J, McCullagh L, Kushniruk A, Mann D, Edonyabo D, McGinn T. A Framework for Usable and Effective Clinical Decision Support: Experience from the iCPR Randomized Clinical Trial. EGEMS (Wash DC) 2015;**3**(2):1150 doi: 10.13063/2327-9214.1150published Online First: Epub Date]|.
59. Toll DB, Janssen KJ, Vergouwe Y, Moons KG. Validation, updating and impact of clinical prediction rules: a review. J Clin Epidemiol 2008;**61**(11):1085-94 doi: 10.1016/j.jclinepi.2008.04.008published Online First: Epub Date]|.
60. Sinuff T, Adhikari NK, Cook DJ, et al. Mortality predictions in the intensive care unit: comparing physicians with scoring systems. Crit Care Med 2006;**34**(3):878-85 doi: 10.1097/01.ccm.0000201881.58644.41published Online First: Epub Date]|.
61. Brabrand M, Hallas J, Knudsen T. Nurses and physicians in a medical admission unit can accurately predict mortality of acutely admitted patients: a prospective cohort study. PLoS One 2014;**9**(7):e101739 doi: 10.1371/journal.pone.0101739published Online First: Epub Date]|.
62. Fillmore CL, Bray BE, Kawamoto K. Systematic review of clinical decision support interventions with potential for inpatient cost reduction. BMC Med Inform Decis Mak 2013;**13**:135 doi: 10.1186/1472-6947-13-135published Online First: Epub Date]|.
63. Wilson FP, Shashaty M, Testani J, et al. Automated, electronic alerts for acute kidney injury: a single-blind, parallel-group, randomised controlled trial. Lancet 2015 doi: 10.1016/s0140-6736(15)60266-5published Online First: Epub Date]|.

64. Kolhe NV, Reilly T, Leung J, et al. A simple care bundle for use in acute kidney injury: a propensity score-matched cohort study. Nephrol Dial Transplant 2016;**31**(11):1846-54 doi: 10.1093/ndt/gfw087published Online First: Epub Date]|.
65. Kolhe NV, Staples D, Reilly T, et al. Impact of Compliance with a Care Bundle on Acute Kidney Injury Outcomes: A Prospective Observational Study. PLoS One 2015;**10**(7):e0132279 doi: 10.1371/journal.pone.0132279published Online First: Epub Date]|.
66. Murray PT, Mehta RL, Shaw A, et al. Current Use of Biomarkers in Acute Kidney Injury: Report and Summary of Recommendations from the 10(th) Acute Dialysis Quality Initiative Consensus Conference. Kidney international 2014;**85**(3):513-21 doi: 10.1038/ki.2013.374published Online First: Epub Date]|.
67. Koyner JL, Davison DL, Brasha-Mitchell E, et al. Furosemide Stress Test and Biomarkers for the Prediction of AKI Severity. J Am Soc Nephrol 2015;**26**(8):2023-31 doi: 10.1681/asn.2014060535published Online First: Epub Date]|.

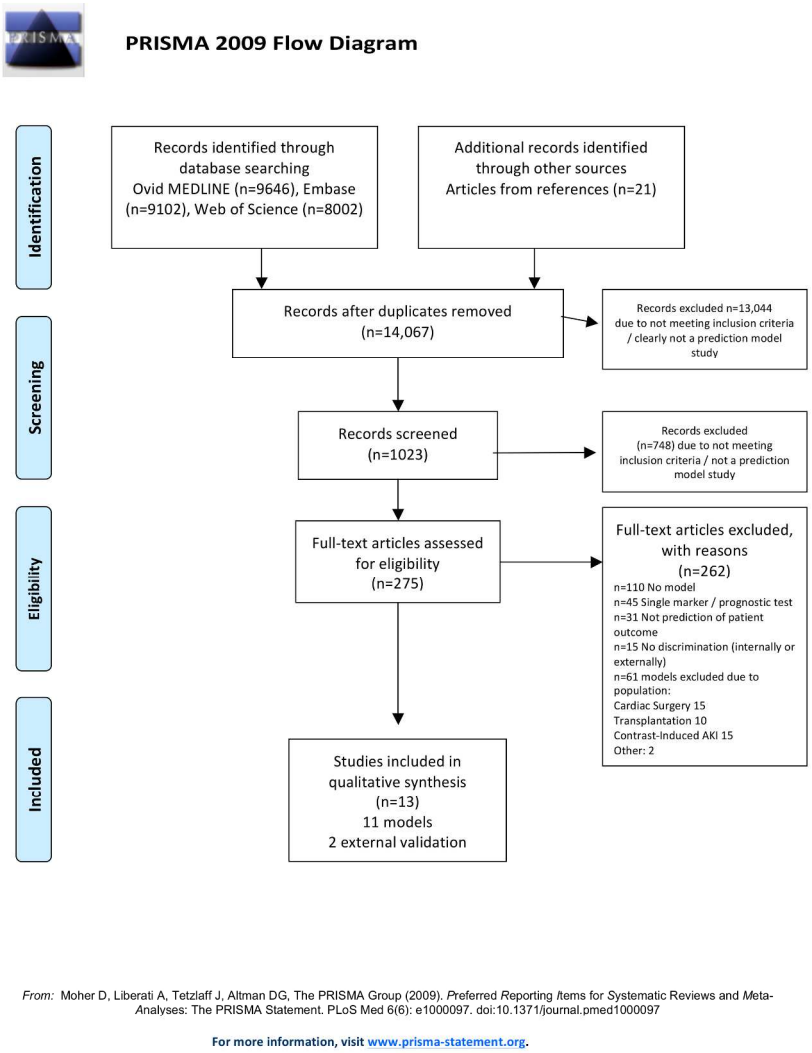

Figure 1 - PRISMA study flow chart

215x279mm (300 x 300 DPI)

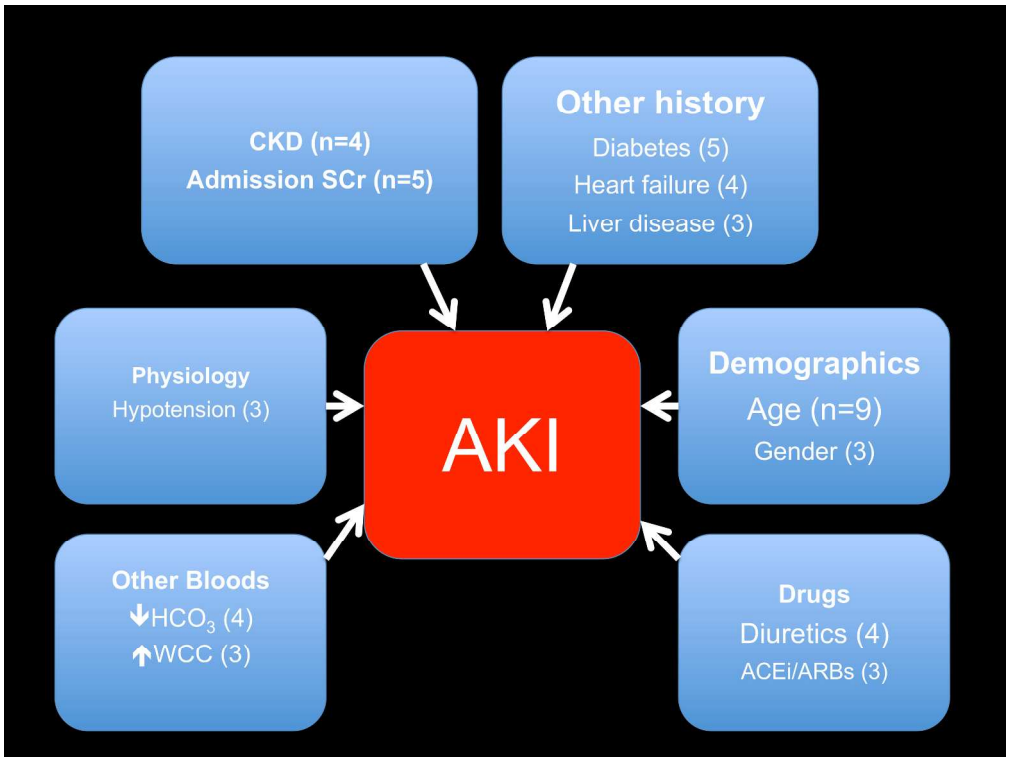

Figure 2 - Predictors most frequently included in the 11 HA-AKI prediction models. ACEi – Angiotensin-converting enzyme inhibitors, ARBs – Angiotensin-receptor blockers, Bloods – laboratory parameters, CKD – chronic kidney disease, ↓HCO3 – reduced serum bicarbonate, SCr – serum creatinine, ↑WCC - raised white cell count.

254x190mm (300 x 300 DPI)

Online Supplementary file

eTable 1 - Study inclusion criteria.

Inclusion Criteria
• Articles in peer-reviewed journals reporting a prognostic multivariable prediction model (scoring system or algorithm) identifying patients who developed HA-AKI (or other measures of renal dysfunction in older studies)
• Validation studies (and updating) of an existing model
• Retrospective, prospective and case-control designs
• Adults (≥18 years) in general hospital settings
• Statistical measures of discrimination (AUROC or c-statistic)
Exclusion Criteria
• Patients <18 years old
• Cardiac surgery, other specialised surgery (e.g. transplantation), CI-AKI
• Non-human studies
• Case reports or conference abstracts
• Only logistic regression without a prediction model
• Lack of discrimination statistics (unless model validated elsewhere)
• Studies that investigated a single predictor, test, or marker
• Studies that investigated only causality between one or more predictors & an outcome
• Use of patients already with the outcome (e.g. AKI present at hospital admission)
• Patients in primary care
• Novel, not widely available tests, such as biomarkers

AUROC - area under the receiver-operating characteristic curve, CI-AKI – Contrast-Induced AKI, HA-AKI - hospital-acquired-AKI.

eTable 2 - Embase Search

LINE	SEARCH TERM
1	(acute AND kidney AND injury).ti,ab
2	AKI.ti,ab
3	(acute AND renal AND failure).ti,ab
4	ARF.ti,ab
5	(contrast AND induced AND nephropathy).ti,ab
6	ACUTE KIDNEY INJURY/
7	OR/1-6
8	predict*.ti,ab
9	PREDICTIVE VALUE OF TESTS/
10	scor*.ti,ab
11	observ*.ti,ab
12	OBSERVER VARIATION/
13	8 OR 9 OR 10 OR 11 OR 12
14	7 AND 13
15	(acute AND kidney AND injury).ti,ab
16	AKI.ti,ab
17	(acute AND renal AND failure).ti,ab
18	ARF.ti,ab
19	(contrast AND induced AND nephropathy).ti,ab
20	ACUTE KIDNEY FAILURE/
21	OR/15-20
22	predict*.ti,ab
23	exp METHODOLOGY/
24	validat*.ti,ab
25	OR/22-24
26	21 AND 25
27	14 AND 26
	Articles: 9102

eTable 3 Ovid MEDLINE® search

Line	Search term
1	(acute AND kidney AND injury).ti,ab
2	AKI.ti,ab
3	(acute AND renal AND failure).ti,ab
4	ARF.ti,ab
5	(contrast AND induced AND nephropathy).ti,ab
6	ACUTE KIDNEY INJURY/
7	OR/1-6
8	predict*.ti,ab
9	PREDICTIVE VALUE OF TESTS/
10	scor*.ti,ab
11	observ*.ti,ab
12	OBSERVER VARIATION/
13	OR/8-12
14	7 AND 13
	Articles: 9646

eTable 4 - Web of Science search

RESULTS	WEB OF SCIENCE SEARCH
8002	(TS=((acute kidney injury) OR (aki) OR (acute renal failure) OR (arf) OR (contrast induced nephropathy)) AND TS=(predict* OR scor* OR observ* OR validat*)) AND DOCUMENT TYPES: (Article) Refined by: WEB OF SCIENCE CATEGORIES: (UROLOGY NEPHROLOGY OR SURGERY OR CARDIAC CARDIOVASCULAR SYSTEMS OR TRANSPLANTATION OR CRITICAL CARE MEDICINE OR MEDICINE GENERAL INTERNAL OR MEDICAL INFORMATICS OR GASTROENTEROLOGY HEPATOLOGY OR ANESTHESIOLOGY) AND DOCUMENT TYPES: (ARTICLE) AND WEB OF SCIENCE CATEGORIES: (UROLOGY NEPHROLOGY OR SURGERY OR CARDIAC CARDIOVASCULAR SYSTEMS OR TRANSPLANTATION OR MEDICAL INFORMATICS OR CRITICAL CARE MEDICINE OR HEALTH CARE SCIENCES SERVICES OR MEDICINE GENERAL INTERNAL OR MEDICAL LABORATORY TECHNOLOGY OR GASTROENTEROLOGY HEPATOLOGY OR ANESTHESIOLOGY OR EMERGENCY MEDICINE) Indexes=SCI-EXPANDED Timespan=All years

eTable 5 - CHARMS checklist and data extracted for systematic review.

Item	Explanation in the Review
1. Type of studies	Prognostic prediction models
2. Scope	Published prognostic prediction models for development of AKI in general hospital settings; to inform risk stratification & potential uses in decision-making in different patient groups
3. Type of studies	Model development +/- external validation in independent data; external model validation & model updating, if present
4. Target population	Adult (≥18) Patients in acute hospital environment
5. Outcome predicted	Development of AKI (or equivalent definition, including RRT) after an admission to hospital or Surgery
6. Time span of prediction	In-hospital development of the outcome
7. Intended moment of using the model	Pre-operatively to predict the risk of post-op AKI or need for RRT; at admission to risk stratify or guide therapy
Summary of Data extracted	
• Data source (years, retrospective, prospective; cohort, case-control, trial data) • Participants & setting (eg cardiac surgery, single or multi-centre, country • Primary outcome (and any blinding) • Candidate predictors (definitions; continuous data dichotomised? & how selected for modelling) • Sample size, EPV (including all predictors considered) • Type of model(s) evaluated - derivation, validation (internal, external) • Missing data, number included & excluded (criteria) • Type of model (eg full model approach), shrinkage • Incidence of outcome & mortality data • Predictors in the model(s) • Performance: discrimination (AUC or C-Statistic) & calibration (eg H-L P value, slope/curve), risk groups • Internal & external validation (in same study) • External validation studies with relevant performance measures • Additional resources, funding	

AKI – acute kidney injury, EPV – events per variable, H-L – Hosmer-Lemeshow goodness-of-fit test, RRT – renal replacement therapy.

eTable 6(i) Surgery

Author, Type, TRIPOD detail	Population, Outcome, AKI Definitions, Methods	Outcomes, Predictors & Model Performance	External Validation
Kheterpal 2007 **General surgery** **TRIPOD 1A - Derivation** **(25/37 pts)**	USA single centre, retrospective cohort study (n=65,043). Data collected 2003-6. Mean Age with outcome 59, without outcome 47 (P<0.001). Male with outcome 56%, without outcome 52% (P=0.32). Inclusion: pre-op eGFR (Cockcroft-Gault) ≥80 ml/min; major surgery (≥2 days in-patient). Exclusions (n=49,941): pre-op eGFR <80 (n=5659). cardiac, transplant, urology & ECT, suprarenal aortic cross-clamping; pre-op AKI & IV contrast <7 days post-op, no pre-op SCr (n=6,534). Included: n=15,102. Outcome: reduction of eGFR to ≤50ml/min <7 days post-op. Predictors: 24 pre, 6 intra-op. Collinearity predictors evaluated; bivariate correlation matrix; remaining predictors entered into logistic regression full model fit. Missing data: excluded from full model. After exclusions n=14,066 included. Un-weighted model continuous predictors dichotomised.	Outcome (AKI) in 0.8% (n=121), 0.1% (n=14) required RRT. Propensity matched 30-day mortality with outcome 15% (n=17/118) vs. 2.7% (n=9/352) without. AKI associated with significant increase in 30-day, 60-day, 1-yr mortality. **7 pre-op predictors: age, emergent surgery, liver disease, BMI, high-risk surgery, PVD & COPD.** Weighted c-Statistic 0.77 (95% CIs 0.75-0.79). Un-weighted risk factor scale (cut-off Age >59, BMI ≥32) c-Statistic 0.73 (0.7-0.76). With intra-op: vasopressor dose, infusion & diuretic: AUC 0.79 (0.77-0.81) No calibration statistics.	Xing 2012 - AUC 0.66
Kheterpal 2009 **General surgery** **TRIPOD 2A - Derivation, Validation** **(28/37 pts)**	USA multi-centre (121) retrospective database study (n=152,244). 2005-6. Mean age with outcome 64.8 (±14.8), without 53.5 (±17.3) (P<0.001). Male with outcome 57%, without outcome 39% (P<0.001). Included n=75,952. Random split derivation 75% (n=57,080) & validation (25% n= 18,872). Exclusions (n=76,292): vascular, cardiac, urology, ophthalmology, obstetric, or urologic procedures; day case; pre-op AKI (rapidly increasing azotaemia & SCr ≥265 µmol/L <24h of surgery) or previous RRT (n=1637). Admission SCr taken as baseline, assessed as predictor & included in the model. 'Mild' pre-op renal insufficiency defined SCr 106-168 µmol/L; 'moderate' >177 µmol/L. Outcome: AKI defined as increase SCr ≥177 µmol/L (from pre-op value) or RRT <30 days. Missing data: SPSS assessed impact of imputation. Continuous predictors dichotomised. Collinearity & Pearson correlations evaluated for all 19 preoperative predictors (comorbidities, drugs, type of surgery). Remaining predictors entered into multivariable logistic regression.	Outcome in 1% (n=762/75,952) – n=561 derivation, n=201 in validation sets. Mortality 42% (n=320) in those with outcome vs 8% in a propensity matched group without outcome. **9 predictors (simplified risk index): age ≥56 yr, male, emergency, intraperitoneal surgery, diabetes, CCF, ascites, HTN, mild or moderate pre-op renal insufficiency.** c-Statistic 0.80 (0.79-0.81) in derivation & internal validation cohorts. Calibration: Risk classes reported for derivation vs validation sets.	-

eTable 6(ii)
Trauma & Orthopaedics

Author, Type, TRIPOD detail	Population, Outcome, AKI Definitions, Methods	Outcomes, Predictors & Model Performance	External Validation
Bell 2015 **T&O** **TRIPOD 3,4 – Derivation, Internal & EV** **(34/37 pts)**	UK multi-centre (3) retrospective cohort study linking multiple prospectively collected databases (n=15,218). 2005-11. Overall mean age 70.7 (±15.3), with outcome 76.5 (±11.1) without outcome 70.0 (±15.6). Overall Male 37%, with outcome 47%, without 36%. Included: derivation n=6,220 (2 sites) & validation n=4,395 (1 site). Exclusions: missing SCr (n=2,688), RRT, 2nd operation (n=1,915). Outcome: KDIOGO SCr changes <7 days. CKD defined using eGFR from CKD-EPI. Admission SCr taken as baseline if elective admission. Entered 11 candidate predictors (age, sex, CKD (baseline eGFR), diabetes, number drugs, ACEi/ARB, NSAID/COX-2, statin, urgency, ASA grade & deprivation category into Backward/forward multivariable selection. Applied a conservative selection criterion of P<0.15 to limit over-fitting risk. Bootstrapping for IV. To assess robustness sensitivity analyses performed: multiple imputation relaxing & restricting the backward selection removal criterion & adding non-linear & interaction terms. Categorised eGFR.	Outcome (AKI) in 10.8% (n=672) derivation & 6.7% (n=295) validation sets. With AKI adjusted hazard ratio 1.53 (95% CI 1.38-1.70). **7 predictors: age, male, diabetes, number drugs, CKD (eGFR), ACEi/ARBs & ASA.** Risk calculator supplied. Derivation AUC 0.74 (0.72-0.76), Internal validation 0.73. EV 0.70. Risk groups shown. Calibration plot. Calibration suboptimal in validation cohort (over-predicted risk). Re-calibration: correction factor, added to intercept; intercept and regression coefficient index as the only predictor used to transform prognostic index & compute recalibrated probabilities.	Same Study different site AUC 0.70

eTable 6(iii)
General admissions

Author, Type, TRIPOD detail	Population, Outcome, AKI Definitions, Methods	Outcomes, Predictors & Model Performance	External Validation
Drawz 2008 **Medicine, surgery, obstetrics** **TRIPOD 2A – Derivation, Internal Validation (26/37 pts)**	USA multi-centre (3), retrospective, case-controlled study (n=180 cases, n=360 controls). 2003. Mean age with outcome 67, controls 63 (P=0.01). Males with outcome 74%, controls 80% (P=0.18). Hospital-acquired AKI (HA-AKI) defined: increase SCr ≥44µmol/L if baseline SCr ≤168µmol/L), ≥88µmol/L baseline 177-433µmol/L & ≥133µmol/L baseline >442µmol/L. Admission SCr presumed to be baseline. Random split derivation (2/3) & internal validation). 'Control' cases – mix of same discharge diagnosis or next patient admitted to clinical team. Inclusions: age ≥18 & normal admission SCr or admission SCr not qualifying as AKI vs known baseline. Exclusions: RRT, no repeat SCr performed. 19 predictors assessed: demographics (age, sex, race), medical history, medications & admission observations & blood parameters (BP, HR, HCO3, urea, SCr, & albumin). Predictors with p value <0.20 univariate analysis entered into multiple logistic regression model. Final model chosen by maximizing likelihood ratio, c-statistic & R2 while minimizing AIC. Also produced a simplified model, created by categorizing continuous variables into quartiles. Cases with missing data excluded. Multiple imputation also performed.	No information on mortality. **7 predictors: age, SBP, HR, HCO_3, urea, albumin & drugs (NSAIDs, ACE-I, ARBs or diuretic).** Derivation c-statistic 0.73. Simplified: HR ≥70/min, HCO_3 (<24 or >30mmol/L), SCr ≥88µmol/L & drugs. Internal validation 0.66. Simplified model HR ≥70, HCO_3 (<24 or >30mmol/L), SCr ≥88 µmol/L & NSAIDs/ACEi/ARBs/Diuretics - C-statistic derivation 0.69, internal validation 0.66. No H-L p-value. Risk range in validation set plotted: 0/1 risk factor = 16% risk HA-AKI, vs 4 risk factors = 62%.	-
Matheny 2010 **General admissions** **TRIPOD 1B – Derivation** **(28/37 pts)**	USA single centre, retrospective cohort study (n=61,179). 1999-2003. No data on mean age (25.6% age >65). Overall males 44.4%. Inclusions: adult admissions ≥2 days (n=26,107). Exclusions: missing data, those with a baseline eGFR <60 (n=11,342), AKI on admission, no SCr available within 48 hrs of admission (n=10,378) or no repeat SCr (n=13,352). Outcome (<30 days post admission): AKI Risk = ≥2 SCr results ≥150% of baseline. AKI Injury = ≥200% baseline. eGFR using MDRD equation. 27 predictors assessed: coded diagnoses (including admission diagnosis), blood parameters (including admission SCr) & drugs following univariate analysis placed in multivariable model. Missing values captured as a separate category. 10-fold cross-validation employed to estimate overfitting.	AKI Risk 5.2% (n=1,352), AKI Injury 2.8% (n=726). No mortality data. **27 predictors: Female, Age, Race, 11 classes of drugs, Contrast, bacterial infection (use of antibiotics), admission SCr, MI, rhabdomyolysis, hepatitis, pancreatitis, ammonia, AST/ALT ratio, thrombocytopenia, leucocytosis, hypercalcaemia, glucose.** AKI Risk: AUC 0.75 (0.73–0.76). H-L P = 0.29. AKI Injury: AUC 0.78 (0.76–0.79), H-L P=0.12. Calibration plotted by deciles.	-

eTable 6(iii) General admissions

Author, Type, TRIPOD detail	Population, Outcome, AKI Definitions, Methods	Outcomes, Predictors & Model Performance	External Validation
Forni 2013 **General medical TRIPOD 2B, 4 – Derivation, Internal & EV** **(29/37 pts)**	UK single centre. Prospective cohort study (n=3,707). 2012. Median (IQR) age with outcome 80 (70-86), without outcome 73 (61-81) (P<0.001). Males with outcome 51% without outcome 49% (P=0.834). Inclusion: medical patients staying >1 night in hospital (n=1,867). Exclusions: RRT, non-medical patients, age <18, AKI on admission (n=184), missing data (n=553). Included n=3,523. Derivation n=1,867. Outcome: AKI (KDIGO SCr change <7 days). CKD defined – eGFR <60 on Pre-admission SCr measured >1 month & <6 months. Internal validation: patients with no previous SCr result, but with a SCr on admission within normal range (defined 80-120µmol/L) (n=1,656). 25 predictors on univariate, If P <0.05 variable entered into multivariable analysis. No missing data information.	Derivation group developed AKI 7% (n=95) – mortality 20% vs 3.5% (n=62) without outcome. In validation cohort n=60 developed AKI. **7 predictors: Age 60-79 (1 point) ≥80 (3 pts), CCF, CKD, Diabetes (2 pts), Liver disease (3 pts), respiratory rate ≥20/min, <alert on AVPU score (3 pts).** Derivation AUC 0.72 (0.66–0.77). H-L P=0.96. Risks plotted. Validation AUC 0.76 (0.71–0.82). No H-L reported.	Hodgson 2017, AUC 0.65-0.71
Bedford 2016 **General admissions TRIPOD 2A, 3 – Derivation, EV** **(29/37 pts)**	UK multi-centre (3), 2011. Retrospective cohort study (n=11,655). Average age and sex not given. Included: derivation n=7,556 admissions & internal validation n=2,514. Exclusions: non-emergency, pre-admission AKI, AKI at admission, obstetrics, patients with no info on AKI at 72 hours. Outcomes: AKI & AKI Stage 2/3. AKI <72 hours, using KDIGO change in SCr. Ordinal logistic regression with univariable analysis for development of multivariable analysis. 45 Predictors included demographics, bloods, prior admissions, co-morbidity. Backwards selection used for retention of statistically significant predictors. Missing data excluded or given own category. 3:1 random split for internal validation. External validation n=1,585, single centre.	Derivation AKI 9.6% (n=241), AKI 2/3: n=40. No mortality data. EV AKI 7.6% (n=120), AKI 2/3 n=12. **12 predictors: age, primary diagnosis, previous hospital admissions, Charlson co-morbidity index score, HbA1C, troponin, proteinuria, baseline eGFR, K^+, WCC, Mg^{2+}, CRP.** IV AUC 0.67 (0.64-0.71) any AKI, 0.68 for AKI 2/3. No derivation AUC H-L P=0.04 any AKI model, P=0.005 for AKI 2/3.	Same study AUC 0.71 (0.63 AKI 2/3). H-L P=0.12 AKI, P=0.14 for AKI 2/3.
Koyner 2016 General admissions TRIPOD 2A Derivation, Internal Validation (24/37 pts)	USA multi-centre (5) Retrospective cohort study (n=269,999). 2008-2013. Mean age with outcome 70 (±16), without outcome 63 (±19) (P<0.001). Males with outcome 49%, without outcome 43% (P<0.001). Included: n=202,961. Exclusions: SCr >354 µmol/L on admission (n=11,305), those without SCr measurement (n=52,508) & AKI prior to arrival on ward (n=3,225). Admission SCr defined as baseline, assessed as predictor & included in model. Outcome: rise SCr as per KDIGO but within 24hrs period. Model included 29 predictors. Continuous predictors modelled using restricted cubic splines with knot placement. Variable importance plot created. Laboratory values & vital signs updated periodically therefore separated into time intervals & logistic regression used for model estimation. Values closest to beginning of that time variable used to predict outcome for that interval, if no values available during an interval, most recent value used, if no previous value available, median value across entire cohort imputed. Split derivation (60%) & internal validation (40%) by time.	AKI 8.6% (n=17,541). Mortality with outcome 6% (n=1031) vs 1% (n=1,419) without **29 predictors: SCr, Urea, HR, anion gap, Urea/SCr, RR, glucose, WCC, K^+, Oxygen Saturations, age, HCO_3, Na^+, temperature, prior ICU, albumin, bilirubin, Ca^{2+}, platelets, time, SBP/ DBP, pulse pressure, sex, AVPU, Alkaline phosphatase, Hb, total protein, AST.** Dsicrimination reported for validation cohort only: AKI AUC 0.74 (0.74-0.74), AKI Stage 3 AUC 0.83 (0.83-0.84) Model including only SCr, BUN & their ratio AUC 0.69 (0.68-0.69).	-

eTable 6(iv)
Heart failure

Author, Type, TRIPOD detail	Population, Outcome, AKI Definitions, Methods	Outcomes, Predictors & Model Performance	External Validation
Forman 2004 **TRIPOD 1B - Derivation, Internal Validation** **(26/37 pts)**	USA multi-centre (11) retrospective cohort study (n=1,009). 1997-8. Overall mean age 67 (±15), with outcome 68.7, without outcome 66.8 (P=0.07). Overall males 51.2%, 52% with outcome 50.9% without outcome. Exclusion (number not given): elective, <2 days, severe aortic stenosis, anticipated transplant, RRT, LVAD, high output failure, age <20, chemotherapy. Excluded n=5 with missing charts. Included n=1004. Outcome: worsening renal function (WRF) - rise SCr >26.5µmol/l during admission. 29 predictors assessed: demographics, history, drugs, symptoms, signs. Unclear method for excluding patients who had AKI at admission. Used admission SCr as baseline & as a predictor. Multivariable Cox regression models, stepwise selection. Bootstrapping for IV. Missing data: predictors missing >15% excluded; categorical data assumed "not present" & separate dummy indicator used if >5% of values missing.	'WRF' 27% (271/1,004). Mortality: risk ratio 7.5 with outcome (number not reported). **4 predictors: CCF, diabetes & BP >160 mmHg (1 point), SCr 132.6-212µmol/l (2 points) & SCr >221µmol/l (3 points).** Risk 'WRF': 0 pts = 10%, 1 = 19%, 2 = 20%, 3 = 30%, 4+ = 53%. 22% of total sample with risk score ≥4 had 53% likelihood WRF vs 10% risk among 12% with risk score 0 points (p<0.001). **No AUC or Calibration statistics.**	Breidthardt 2011 - 0.65 Wang 2013 - 0.65
Breidthardt 2011 **TRIPOD 1A – Derivation, (EV Forman)** **(23/37 pts)**	Swiss multi-centre (3) prospective analysis, with derivation (Basel score) & external validation of Forman score (n=767). 2001-2, 2006-2010. Overall median age 79 (71-85), with outcome 79 (72-85), without outcome 79 (70-85) (P=0.36). Overall males 55%, with outcome 61%, without outcome 54% (P=0.08). Included n=657. Exclusions (n=110): stay <2 days, incomplete SCr. Outcome: WRF = in-hospital increase SCr ≥26.5µmol/L. CKD from eGFR (using MDRD equation) <60 for >3/12 pre-admission. eGFR at admission to hospital included as a predictor. Unclear method for excluding patients who had AKI at admission. 48 predictors assessed on univariate & those with P value <0.05 entered into multivariable analysis. No missing data information. n=223 had blood gas analysis.	Outcome 21% (136/657). In-hospital mortality with outcome 17% (n=23) vs 6% (n=33) without (P <0.01). **3 predictors (n=223): HCO_3 <21 mmol/L, Diuretics, CKD** - AUC 0.71 (0.63-0.79). A computer-based, complex, exponential risk model AUC 0.75 (0.67-0.82). No H-L calibration data. Scores & percentage developing outcome: 0 - 1%, 1 –35%, 2 –27%, 3 – 35%.	-

eTable 6(iv)
Heart failure

Author, Type, TRIPOD detail	Population, Outcome, AKI Definitions, Methods	Outcomes, Predictors & Model Performance	External Validation
Wang 2013	China, single centre, retrospective cohort study (n=1,709). 2004-11. Median age with outcome 73 (67-78), without outcome 71 (63-75) (P<0.001). Males 56.6% with outcome, 55% without outcome (P=0.13).	Overall AKI 32% (n=550). Mortality 16.5% (n=91) vs.1.9% (n=22) without AKI (P <0.01). Stay with AKI 14 vs. 11 days without (P <0.01).	
TRIPOD 2A – Derivation, External validation (Forman) **(30/37 pts)**	Inclusion: CCF admission diagnosed by 2 cardiologists using European Society of Cardiology guidelines. Exclusions: age <18, stay <2 days, missing data, hospital transfer, use LVAD, ESRD or RRT & septic or haemorrhagic shock; cardiac op, pacemaker or cardioversion & contrast. Split derivation (60%, n=1010) & validation (40%, n=699). Outcome: AKI (AKIN): increase SCr ≥26.4 µmol/L or ≥50% in <48 hrs. eGFR – MDRD – unclear whether admission SCr was used to estimate baseline eGFR or how patients with AKI at admission were excluded. Admission SCr used as a predictor. 35 predictors – those with P value <0.1 on unviariate analysis placed in multivariate analysis (n=932).	**8 predictors: Age ≥70; ≥3 CCF admissions, systolic BP <90mmHg, Na^+ <130mmol/L, NYHA IV, proteinuria, SCr ≥104 µmol/L & furosemide dose ≥80 mg/day.** Derivation AUC 0.76 (0.73–0.79) H-L P=0.98. Calibration plots by deciles. Validation 0.76 (0.72-0.8), H-L P=0.13. ≥8 points high risk - 55.1% incidence vs. 18% if <8 points No calibration slope. Forman - 0.65 (0.62–0.69). vs Forman score, improvement of 0.11 AUC, (P <0.001(DeLong)(9)	-

eTable 7 – Abbreviations used in eTable 6(i-iv)

ACEi – Angiotensin-converting enzyme inhibitors
AKI – Acute kidney injury
AKIN – Acute kidney injury network
ALT – Alanine aminotransferase
ARB – Angiotensin receptor blockers
ASA – American Society of Anesthesiologists Physical status grading used in pre-operative assessment
AST – Aspartate transaminase
AVPU – scale of consciousness best response: **A**lert, responds to **V**oice, **P**ain, **U**nresponsive.
AUC/AUROC – Area under the receiver operating characteristic curve
BMI – Body mass index
BP – Blood pressure
CA-AKI – Community-acquired AKI
Ca^{2+} - Serum Calcium
CI-AKI – Iodinated contrast AKI
CKD – Chronic kidney disease
COPD - Chronic obstructive pulmonary disease
CCF – Congestive cardiac failure
CKD-EPI – CKD Epidemiology collaborative equation
COX – Cyclo-oxygenase
CRP – C-reactive protein
D – Derivation study
DBP – Diastolic Blood Pressure
eGFR – estimated glomerular filtration rate
ESRD – end-stage renal disease
EV – External validation study
HA-AKI – Hospital-acquired AKI
Hb - Haemoglobin
HbA1C – glycated haemoglobin (A1c) Marker of long-term glucose control
HCO_3 – serum Sodium Bicarbonate
H-L – Hosmer-Lemeshow goodness-of-fit test (Calibration statistic)
HR – Heart rate (beats per minute)
HTN – Hypertension
ICU – Intensive Care Unit
IHD – Ischaemic heart disease
IV – Internal Validation study
K^+ - Serum Potassium
KDIGO – Kidney disease improving global outcomes (Stage 1-3 AKI defined by magnitude of SCr rise or fall in ur
LOS – length of stay
LVAD – Left ventricular assist device
LVEF – Left ventricular ejection fraction
MAP – Mean arterial pressure
MDRD – Modification of diet in renal disease equation
Mg^{2+} - serum Magnesium
MI – Myocardial infarction
Na^+ – Serum sodium
NSAID – Non-steroidal anti-inflammatory agent

NYHA – New York Heart Association Classification for heart failure (I-IV)
PVD – Peripheral vascular disease
RIFLE – Risk, Injury, failure, loss of kidney function
RR – respiratory rate (breaths per minute)
RRT – Renal replacement therapy
SCr – serum creatinine
Systolic Blood Pressure
TRIPOD – Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis. Checklist in derivation (37 points – 1 point for each recommended item reported). **TRIPOD Study types** Type 1a: Development only Type 1b: Development and validation using resampling Type 2a: Random split-sample development and validation, Type 2b: Non-random split-sample development and validation Type 3: Development and validation using separate data Type 4: Validation only. WCC – White cell count WRF – 'worsening renal failure' (defined by individual study)

eTable 8 - TRIPOD items reported in the 11 studies.

Title & Abstract		TRIPOD Item description	Reported ?
Title	1	Identify study as developing &/or validating a multivariable prediction model, target population & outcome.	10
Abstract	2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results & conclusions.	9
Introduction			
Background & objectives	3a	Explain medical context (including whether diagnostic or prognostic) & rationale for developing or validating the multivariable prediction model, including references to existing models.	11
	3b	Specify objectives, including whether the study describes development or validation of the model or both.	11
Methods			
Source of data	4a	Describe study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development & validation data sets, if applicable.	11
	4b	Specify key study dates, including start of accrual; end of accrual; & if applicable, end of follow-up.	11
Participants	5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number & location of centres.	11
	5b	Describe eligibility criteria for participants.	11
	5c	Give details of treatments received, if relevant.	9
Outcome	6a	Clearly define outcome predicted by the prediction model, including how & when assessed.	11
	6b	Report any actions to blind assessment of the outcome to be predicted.	0
Predictors	7a	Clearly define all predictors used in developing or validating the model, including how & when they were measured.	11
	7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	0
Sample size	8	Explain how the study size was arrived at.	2
Missing data	9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	7
Statistical analysis methods	10a	Describe how predictors were handled in the analyses.	11
	10b	Specify type of model, model-building procedures (including predictor selection) & method for internal validation.	11
	10c	For validation, describe how the predictions were calculated.	6
	10d	Specify all measures used to assess model performance & if relevant, to compare multiple models.	8
	10e	Describe any model updating (e.g., recalibration) arising from the validation, if done.	1
Risk groups	11	Provide details on how risk groups were created, if done.	10
Development vs. validation	12	For validation, identify differences from development data in setting, eligibility criteria, outcome & predictors.	4
Results			
Participants	13a	Describe flow of participants through the study, including number of participants with & without the outcome & if applicable, a summary of the follow-up time. A diagram may be helpful.	11
	13b	Describe characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors & outcome.	9
	13c	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	3
Model development	14a	Specify the number of participants & outcome events in each analysis.	11
	14b	If done, report the unadjusted association between each candidate predictor & outcome.	11
Model specification	15a	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients & model intercept or baseline survival at a given time point).	8
	15b	Explain how to the use the prediction model.	11
Model performance	16	Report performance measures (with CIs) for the prediction model.	8
Model-updating	17	If done, report the results from any model updating (i.e., model specification, model performance).	1
Discussion			
Limitations	18	Discuss any limitations of the study (non-representative sample, few events per predictor, missing data).	11
Interpretation	19a	For validation, discuss results with reference to performance in development data & any other validation data.	8
	19b	Give overall interpretation of results, considering objectives, limitations, results from similar studies & other relevant evidence.	11
Implications	20	Discuss potential clinical use of the model & implications for future research.	11
Other information			
Suppl info	21	Provide information about availability of supplementary resources, (study protocol, Web calculator, & data sets).	3
Funding	22	Give the source of funding & role of the funders for the present study.	9

Red colouring highlights items reported in less than 50% of the 11 AKI prediction model studies.

eTable 9 – Most common predictors included in the 11 models

Field	Total	General Surgery		T&O	General					Heart Failure		
Study		Kheterpal 2007	Kheterpal 2009	Bell 2015	Drawz 2008	Matheny 2010	Koyner 2016	Bedford 2016	Forni 2013	Forman 2004	Breidthardt 2011	Wang 2013
Demographics												
Age	9	X	x	x	x	x	x	x	x			x
Male/gender	3		x	x			x					
Past history												
Diabetes	5		x	x				x	x	x		
CKD	4			x				x	x		x	
Heart failure	4		x						x	x		x
Liver disease	3	X				x			x			
Drugs												
Diuretics	4				x	x					x	x
ACEi/ARBs	3			x	x	x						
Observations												
Hypotension/ Shock	3				x		x					x
Bloods												
SCr	5		x			x	x			x		x
Bicarbonate	4				x	x	x				x	
↑WCC	3					x	x	x				

ACEi – angiotensin-converting enzyme inhibitor drugs, ARB – Angiotensin 2 receptor blocker drugs, CKD = chronic kidney disease, Bloods – laboratory parameters, SCr – serum creatinine, T&O – Trauma and Orthopaedics, WCC – white cell count.

eTable 10 – All predictors included in the 11 models

Field	General Surgery		T&O	General					Heart Failure			Total
Study	**Kheterpal 2007**	**Kheterpal 2009**	**Bell 2015**	**Drawz 2008**	**Matheny 2010**	**Koyner 2016**	**Bedford 2016**	**Forni 2013**	**Forman 2004**	**Breidthardt 2011**	**Wang 2013**	
Demographics												
Age	x	x	x	x	x	x	x	x			x	9
Male/gender		x	x			x						3
BMI	x											1
Race					x							1
Past history												
Diabetes		x	x				x	x	x			5
CKD			x				x	x		x		4
Heart failure		x						x	x		x	4
Liver disease	x				x			x				3
Hypertension		x										1
PVD	x											1
Ascites		x										1
COPD	x											1
Previous admissions						x (ICU)	x					2
Charlson co-morbidity index							x					1
ASA Grade			x									1

Field	General Surgery		T&O	General					Heart Failure			Total
Study	Kheterpal 2007	Kheterpal 2009	Bell 2015	Drawz 2008	Matheny 2010	Koyner 2016	Bedford 2016	Forni 2013	Forman 2004	Breidthardt 2011	Wang 2013	
Drugs												
Diuretics				x	x					x	x	**4**
ACEi/ARBs			x	x	x							**3**
NSAIDs				x	x							**2**
Contrast					x							**1**
Number of drugs			x									**1**
Other drugs[1]					x							**1**
Observations												
Hypotension / Shock				x		x					x	**3**
Pulse pressure						x						**1**
Hypertension				x					x			**2**
Heart rate[2]				x		x						**2**
Temperature						x						**1**
Respiratory rate[3]						x		x				**2**
O2 saturations						x						**1**
Consciousness[4]						x		x				**2**
Surgery, Other												
Type of surgery	x	x										**2**
Emergency	x	x										**2**
Time						x						**1**

Field	General Surgery		T&O	General					Heart Failure			Total
Study	Kheterpal 2007	Kheterpal 2009	Bell 2015	Drawz 2008	Matheny 2010	Koyner 2016	Bedford 2016	Forni 2013	Forman 2004	Breidthardt 2011	Wang 2013	
Labs, Diagnosis												
Primary diagnosis							x					**1**
SCr		x			x	x			x		x	**5**
Haemoglobin						x						**1**
↓Platelets[5]					x	x						**2**
CRP							x					**1**
↑WCC					x	x	x					**3**
Bacterial infection[6]					x							**1**
MI[7]					x		x					**2**
Rhabdomyolysis[8]					x							**2**
Hepatitis/AST[9]					x	x						**2**
Alk Phosphatase						x						**1**
Bilirubin						x						**1**
Pancreatitis[10]					x							**1**
Bicarbonate				x	x	x				x		**4**
Anion gap						x						**1**
BUN/Cr						x						**1**
Urea				x		x						**2**
↑Ca^{2+}/Ca^{2+}[11]					x	x						**2**
↑glucose					x	x						**2**
Magnesium							x					**1**
Potassium						x	x					**2**
↓Na^{+}/Na^{+}[12]						x					x	**2**
Albumin				x		x						**2**
Total protein						x						**1**
Proteinuria							x				x	**2**

eTable 11 – Handling of Serum Creatinine and Chronic Kidney Disease in the models

Author, year (n=derivation cases)	Kheterpal 2007 (n=14,066)	Kheterpal 2009 (n=57,080)	Bell 2015 (n=6,220)	Drawz 2008 (n=360)	Matheny 2010 (n=26,107)	Koyner 2016 (n=202,961)	Bedford 2016 (n=7,556)	Forni 2013 (n=1,867)	Forman 2004 (n=1,004)	Breidthardt 2011 (n=657)	Wang 2013 (n=1,010)
CKD defined			X				X	X	X	X	X
Admission SCr used as baseline		X		X		X			X		X
Omitted cases with reduced GFR	X				X						
Admission SCr assessed as predictor		X		X	X	X			X	X	X
CKD included in model			X				X	X		X	
Admission SCr included in model		X			X	X			X		X

CKD - Chronic Kidney Disease, SCr - Serum Creatinine. Red shading indicates concern over handling of SCr in the study.

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	Title page
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	Page 2-3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	5-6
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	7-8
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	NA
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	8, online supplementary eTable 1
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	7
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	online supplementary eTable 2-4
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	7-8
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	8-9, online supplementary eTable 5
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	8-9 online supplementary eTable 6
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	9
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	NA

PRISMA 2009 Checklist

Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	N/A

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	online supplementary eTables 6, 8
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	NA
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	10, Figure 1,
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	supplementary eTables 6
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	supplementary eTables 6, 8
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	supplementary eTable 6
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	NA
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	NA
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	NA
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	13-15
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	supplementary eTable 6, 8, pg.14
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	15-18
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	NA

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097 For more information, visit: www.prisma-statement.org.