A novel patient-specific model for predicting severe oliguria; development and comparison with KDIGO acute kidney injury classification

Authors: Dr Samuel H. Howitt MBChB, FRCA [1,2], Mr Jordan Oakley[3] MMath, [1], Dr Camila Caiado PhD[3], Prof Michael Goldstein PhD[3], Dr Ignacio Malagon PhD, FRCA, FFICM [1,2], Professor Charles McCollum MD, FRCS, Mr Stuart W. Grant PhD, MRCS [1]

1. Division of Cardiovascular Sciences, University of Manchester, ERC, Manchester University Hospitals Foundation Trust, Manchester, M23 9LT, UK.
2. Department of Cardiothoracic Anaesthesia and Critical Care, Wythenshawe Hospital, Manchester University Hospitals Foundation Trust, Manchester, M23 9LT, UK.
3. Department of Statistics, Durham University, Durham, DH1 3LE, UK.

Study undertaken at Wythenshawe Hospital, Manchester University Hospital Foundation Trust, Manchester, M23 9LT, UK.

Corresponding author

Samuel Howitt, Division of Cardiovascular Sciences, University of Manchester, 2nd Floor ERC, Wythenshawe Hospital Manchester, UK. M23 9LT. Tel. (0044)161 291 5842. Email samuel.howitt@manchester.ac.uk

No reprint required

Conflict of interests

The authors declare that they have no conflict of interests.

Funding Statement

Funding from the British Heart Foundation grant number PG/16/80/32411 was used to support this research.

Keywords: acute kidney injury, urine, critical care

2852)/3000 Words (ex abstract)

ABSTRACT

Objective

The KDIGO urine output criteria for acute kidney injury (AKI) have been shown to lack specificity for identifying patients at risk of adverse renal outcomes. The objective was to develop an alternative system to analyse urine output and identify those at risk of developing severe oliguria.

Design

This was a retrospective cohort study utilising prospectively collected data.

Setting

A cardiac intensive care unit in the UK.

Patients

Patients undergoing cardiac surgery between January 2013 and November 2017

Measurement and main results

Patients were randomly assigned to development (n=981) and validation (n=2389) datasets. A patient-specific, dynamic Bayesian model was developed to predict future urine output. Model discrimination and calibration for predicting severe oliguria (<0.3ml/kg/hr for 6 hours) occurring within the next 12 hours were tested in the validation dataset at multiple time points. Patients with a high-risk (probability of severe oliguria >0.8) were identified and their outcomes were compared with those for low-risk patients and for patients who suffered AKI based on KDIGO urine output criteria.

Model discrimination was excellent at all time points (AUC >0.9 for all). Calibration of the model's predictions was also excellent. Multivariable logistic regression demonstrated that patients in the high-risk group were more likely to require renal replacement therapy (OR 10.4, 95%CI 5.9-18.1), suffer prolonged hospital stay (OR 4.4, 95% CI 3.0-6.4) and die in hospital (OR 6.4, 95%CI 2.8-14.0) ($p<0.001$ for all). Outcomes for those identified as high-risk by the model were significantly worse than those classified as suffering AKI based on KDIGO urine output criteria.

Conclusions

This novel, patient-specific model accurately identifies patients at increased risk of severe oliguria. Classification according to model predictions outperformed the KDIGO urine output criteria. As the new model identifies patients at risk before severe oliguria develops it could potentially facilitate intervention to improve patient outcomes.

300/300 words

Introduction

Acute kidney injury (AKI) is defined and stratified by the KDIGO AKI guidelines(1) and occurs in up to 75% of patients in general intensive care units(2, 3) and up to 30% of patients following cardiac surgery(4). The KDIGO guidelines stratify the severity of AKI based on serum creatinine concentration and urine output. Studies in both cardiac surgery and general ICU patients have shown that the guidelines' creatinine criteria successfully identify patients with increased risk of prolonged length of stay, short-term mortality and long term mortality.(3, 5-8) However, there is less agreement about the value of the guidelines' urine output criteria which define AKI as urine output below 0.5ml/kg/hr for more than 6 hours. Most large studies were unable to obtain enough urine output data to adequately assess the importance of the urine output criteria in the prediction of adverse outcomes.(3, 7, 8) Some smaller studies demonstrated that calibration of the KDIGO urine output thresholds may be inadequate by showing that patients diagnosed with AKI by urine output alone had relatively good outcomes compared with those who also met the guideline's serum creatinine criteria.(2, 9-11) Ralib *et al* demonstrated that a urine output threshold of 0.3ml/kg/hr for 6 hours (severe oliguria) was more closely associated with adverse outcomes in general ICU patients.(9) However, use of this threshold rather than the 0.5ml/kg/hr for 6 hours threshold specified in the KDIGO stage 1 definitions could lead to adverse patient outcomes related to the 6 hours of marked oliguria required to before risk stratification could occur. Dynamic Bayesian modelling(12, 13) has been used in related settings (14, 15) and could provide a solution to this problem by identifying those at greatest risk of severe oliguria early enough to allow treatment to be administered. The objective of this study was to develop and validate a patient-specific dynamic Bayesian model which could run in real time to predict the risk of developing severe oliguria. To confirm the clinical usefulness of the model, associations between those at a high predicted risk of severe oliguria and adverse outcomes were also investigated. Outcomes of the high-risk group were also compared with patients who met existing KDIGO urine output criteria.

Materials and Methods

Data

Prospectively collected data from adult patients admitted to the cardiac intensive care unit (CICU) following cardiac surgery between January 2013 and November 2017 were analysed. Patients receiving mechanical circulatory support (MCS) or cardiac transplantation were excluded. Patients who received renal replacement therapy (RRT) preoperatively were also excluded.

Hourly urine output values and the time of any decision to initiate RRT were extracted from the electronic patient record. Only urine output data recorded before the initiation of RRT was analysed. Outcome data was collected from the hospital's clinical governance database. All data were cleaned and stored in the Vascular Governance NorthWest (VGNW) database, handled according to the database's ethical approvals and anonymised prior to analysis. All data cleaning and analysis was performed using R Studio (R Foundation for statistical computing).(16)

Model development

Eligible patients were randomly assigned to either model development or model validation datasets in a ratio of 1:2.5 to ensure a development group of around 1000 patients. A dynamic linear model was developed using data the development dataset. The model analysed each patient's own previous hourly urine output values and then from the 6th hour on CICU predicted that individual's urine output for the next 6 hours. The model produced updated predictions on an hourly basis throughout the CICU stay as each new measurement became available. The probability of the next 6 hours' urine output being below 0.3ml/kg/hr was calculated using Bayesian forecasting. The model applied weightings to the contributions of urine output values according to how recent they were with the most recent values deemed the most relevant. This allowed the forecast to update quickly in response to changing trends. Model development is described in detail in the appendix.

Model validation (statistical analyses)

It was recognised that for a subgroup of patients the model could potentially provide inappropriate reassurance to clinicians. The model could predict a low risk of severe oliguria occurring within 6 hours and the patient could pass enough urine to prevent severe oliguria occurring within 6 hours but go on to suffer severe oliguria soon afterwards. In this scenario while statistically correct, the model's output could limit its clinical usefulness. The validation analyses therefore tested the model's ability to identify which patients would suffer severe oliguria (UO <0.3ml/kg/hr for 6 hours) within 12 hours of the prediction. Risk classifications made during the last 12 hours of a patient's admission were disregarded as it was not possible to confirm if severe oliguria subsequently occurred following discharge from CICU. Discrimination (the ability to distinguish those who would suffer severe oliguria from those who would not) was assessed using Receiver Operator Curve (ROC) analyses. The 95% confidence intervals for the area under the curves (AUC) were calculated using DeLong's method.(17) As the dataset was unbalanced (severe oliguria was relatively rare) precision recall curves were also used to test model performance.(18) Calibration (how well predicted risk matched observed outcomes) was assessed using the ratio of observed to expected outcomes (O:E ratio) and calibration plots.(19) The calibration plots show the observed and predicted risk of severe oliguria for patients grouped into twenty evenly sized groups according to their predicted risk. For completeness, performance of the model when predicting severe oliguria limited to the six hours following predictions was also assessed with full results in the appendix.

Although in clinical practice clinicians are likely to interpret the model's continually updated risk predictions rather than a binary risk classification, to allow comparison of the model's predictions with the existing categorical KDIGO classification, patients were assigned to either a high-risk or a low-risk group. Patients for whom the probability of severe oliguria reached >0.8 during their stay were arbitrarily classified as high-risk and those who did not were classified as low-risk. This relatively high threshold was selected *a priori* as the aim

was to produce a classification with a high specificity. Associations between this classification and postoperative RRT, prolonged length of stay (PLOS) and hospital mortality were tested using univariable and multivariable analyses. Outcomes for patients grouped according to classification by the model and the KDIGO criterion were also compared. PLOS was defined as a hospital stay >10 days. If RRT was initiated within three hours of CICU admission, the patient was excluded from the analyses as case note analyses revealed that all of these decisions to start RRT had been made during surgery before the patient arrived on CICU. If the decision to initiate RRT was made before a high-risk classification, the patient was assigned to the low-risk group and the RRT was considered to have been administered to a low risk patient. Univariable analyses were performed using the Chi Square test or Fisher's exact test in the event of sparse data. Multivariable logistic regression was used to adjust for the confounding effects of pre- and perioperative variables associated with adverse outcomes using the extensively validated logistic EuroSCORE model. (20, 21) Cardiopulmonary bypass (CPB) time was used as a surrogate marker to adjust for intra-operative procedure complexity.

The sensitivity, specificity, positive predictive value and negative predictive value of classification by the new model based on the arbitrary threshold of 0.8 for the identification of those at risk of subsequent RRT were calculated. These values were compared with equivalent values obtained when classifying patients according to i) the KDIGO UO criterion (UO <0.5ml/kg/hr for 6 hours) and ii) observed severe oliguria (UO<0.3ml/kg/hr for 6 hours).

Missing data

Where hourly urine output was recorded as "0" this value was used. Where hourly values were blank, the next recorded urine output was divided by the number of hours that had elapsed since the previous reading and this value was substituted for the blank values. Where this imputation resulted in urine output lower than the 0.5ml/kg for 6 hours the cases notes were examined and the urine output entries verified through entries in the nursing

notes. Where weight was missing, the value was imputed using the median weight for a patient of that gender.

Results

In total 3,602 patients were admitted to CICU following cardiac surgery, 228 were excluded as they underwent cardiac transplantation or received MCS and four patients were excluded as they received RRT preoperatively. Of the eligible 3370 patients, 981 were randomly assigned to the development group and the remaining 2389 patients were assigned to the validation group. The patient characteristics of each group are shown in Table 1. Patient weight was missing for 13 (1.3%) and 23 (1.0%) patients in the development and validation cohorts respectively.

In the validation cohort, 2088 (87.4%) patients suffered at least one hour of urine output below 0.3ml/kg/h. There were 197 (8.2%) patients who experienced severe oliguria and 89 (3.7%) patients who required RRT. In total, 4942 (2.8%) hourly urine output entries were missing and these values were imputed using the methods described in the previous section. A total of 19 (0.8%) patients received RRT within three hours of arrival on CICU and these patients were excluded from the RRT analyses. PLOS was observed in 589 (24.7%) patients and 36 (1.5%) died prior to hospital discharge. There were no missing outcome data.

Predicting severe oliguria

The Receiver operating characteristic curves for the prediction of severe oliguria within the next 12 hours for predictions made at 12, 24, 36, 48 and 72 hours are shown in Figure 1a. At each time point the AUC for the predictions was >0.9 representing excellent discrimination between those who did and did not go on to suffer severe oliguria within the next 12 hours. As illustrated by Figure 2 and the O:E ratios detailed in Table A1 of the appendix, calibration was also excellent.

The precision recall curves (Figures appendix 1) illustrate the trade-off between ensuring that every patient who will go on to suffer AKI is identified and that the number of false positives is minimised. As shown in Figure appendix , as recall (also known as sensitivity) approaches 1 the Precision (positive predictive value) falls. This effect was most pronounced for predictions made in the first 24 hours.

Table A2 of the appendix describes the model's performance when predicting severe oliguria occuring within 6 hours of prediction. Discriminination was consistently better than when predicting severe oliguira occuring within 12 hours following predictions but risk was consistenty overestimated. Across the 5 time points analysed there were 258 incidences where a patient developed severe oliguria within 12 hours of predictions, however on 109 occasions patients severe oliguria only developed between 7 and 12 hours after prediction.

Classification task

In the validation dataset 158 patients experienced a probability of severe oliguria >0.8 and were assigned to the high-risk group. The remaining 2231 patients were assigned to the low-risk group. Outcomes for these two groups are reported in Table A3. High-risk patients experienced increased rates of subsequent RRT, PLOS and hospital mortality compared with those classified as low-risk (P<0.001 for all outcomes). On multivariable analysis, high-risk classification was associated with increased risk of RRT (OR 10.4, 95%CI 5.9-18.1), PLOS (OR 4.4, 95% CI 3.0-6.4) and hospital mortality (OR 6.4, 95%CI 2.8-14.0) (p<0.001 for all outcomes). The multivariable models used for risk adjustment are shown in the Appendix (Tables A4-A6). The median (IQR) time from high-risk classification to the onset of severe oliguria of 3.0 (0.0-4.0) hours

The KDIGO urine output criterion identified 628 patients (26.3%) as suffering AKI by urine output. The outcomes for classification of risk using the new model and the KDIGO criterion are compared in Table 2. Outcomes for those classified as being at high risk by the model

and those meeting the KDIGO criteria could not be compared directly as some patients would have been include in multiple groups. Patients who met the KDIGO urine output criterion for AKI but were classified as low-risk by the model (n=506) experienced rates of RRT (3.6%), PLOS (34.8%) and mortality (2.4%) which were significantly lower than the risks for those classified as high-risk by the Bayesian model ($p<0.001$ for all). When used to predict future RRT requirement, the Bayesian model achieved greater specificity and positive predictive value (but lower sensitivity) than the KDIGO AKI criterion. The performance of the dynamic Bayesian model was almost identical to that achieved by classification according to actual observed oliguria. (Table 3)

Discussion

This patient-specific Dynamic Bayesian model was developed and validated in separate cohorts which together contained high quality, prospectively-gathered data for over 3000 patients. The model successfully identified patients at risk of severe oliguria demonstrating excellent discrimination and calibration at each time point. Outcomes were significantly worse for patients with a high-risk of severe oliguria than for those assigned to the low-risk group. Those identified as high risk by the model also suffered worse outcomes than those who only met the KDIGO urine output criterion for AKI.

The unbalanced nature of the data had the potential to make the AUC statistics seem overly impressive. Indeed, precision recall curve analyses showed that the excellent discrimination identified on ROC curve analyses of predictions made at 12 and 24 hours was partly due to the large proportion of patients who did not suffer severe oliguria and whom the model identified as being at low risk of oliguria. However, this effect was less significant for predictions made after this time.

During the validation of predictions made at hours 12, 24, 36, 48 and 72, 109 incidences were identified in which a patient suffered severe oliguria between 7 and 12 hours following predictions. In some cases it is likely that the model apprpriately predicted severe oliguira

would occur within 6 hours but an intervention such as a fluid challenge or a trial of diuretic therapy was impemented at this time. In some patients the response would be transitory, causing a the urine output to rise briefly above the 0.3ml/kg/hr threshold before falling back to a lower level. Such patients would therefore only meet the 0.3ml/kg/hour for 6 hours later. Nevertheless, these patients suffered an adverse event and the identifiction of such a significant number of incidences of severe oliguria occurring between hours 7 and 12 justies the selection of severe oliguira occurring with 12 hours of prediction as the outcome used when validating predictions.

In clinical practice, classification into high and low-risk groups based on an arbitrary threshold is unlikely to be necessary and significantly diminishes the usefulness of the model. Rather, patient monitoring software would analyse the individual's urine output data in real-time and display updated estimates of the absolute risk of developing severe oliguria. This information, together with the trend of risk for that patient would inevitably be much more useful to a treating clinician than knowledge of the patient's risk group.

In this study a threshold was used to dichotomise the patients purely to allow the comparison of outcomes observed in patients classified as high and low-risk by the model. The categorisation also allowed comparison of outcomes between patients classified as high-risk by the model and patients who met the existing KDIGO AKI criteria. The threshold used for the classification exercise was deliberately high at 0.8 to reduce the number of false positive high-risk classifications which are a weakness of the existing KDIGO AKI classification.(2, 9, 10) As a result a large subgroup (n=506) met the KDIGO AKI criterion but were classified as low-risk by the model. Outcomes for these patients were significantly better than for the group classified as high-risk by the model suggesting that for a large proportion of those who meet the KDIGO urine output criterion risk of adverse outcomes is actually relatively low.

The significant increase in risk of adverse outcomes found to be associated with a predicted or observed fall in urine output to < 0.3ml/kg/hr for 6 hours is similar to that found in general ICU patients(9) and justifies the selection of this threshold in this study. Risk stratification was not significantly improved when classification was made according to observed rather than predicted severe oliguria. The main advantage of using the dynamic Bayesian model is that it provides reliable, early warnings of impending severe oliguria before it occurs, allowing time to deliver treatments to prevent the severe oliguria and its consequences. Even if a warning were only raised when a probability of 0.8 for severe oliguria was reached - as in our classification exercise – this would allow interventions aimed at preserving renal function. In reality patients for whom risk of severe oliguria is increasing are likely to be reviewed before a probability of 0.8 is reached, affording even more time for intervention.

Clinical use of a urine output screening protocol which employs this dynamic Bayesian model is perfectly feasible because although mathematically complex, the model is computationally inexpensive and can run on standard computers or tablets available at the bedside. The model uses the trend of urine output rather than comparison of point values against arbitrary thresholds. The progressive decline in urine output towards the defined threshold of 0.3ml/kg is intuitively more relevant than the occurrence of a point value below an arbitrary “normal”. Indeed, over 85% of those classified as low-risk suffered at least one hour of urine output below 0.3ml/kg/hr but this group had excellent outcomes. As the only data required by the model are patient weight and hourly urine output values, the model should be transferrable across all patients on critical care units. In this study we chose to calculate the probability of urine output dropping below 0.3ml/kg/hr but this threshold could be altered to suit different patient cohorts. Under these circumstances the model could be useful across a range of settings, alerting clinicians to the risk of urine output dropping below a threshold they consider to be clinically significant.

While these results are encouraging, analyses of urine output alone cannot identify all patients at risk of adverse outcomes related to renal dysfunction. Indeed, 41 patients

received RRT despite being classified as low-risk because their urine output was maintained around or above 0.3ml/kg/hr. Analysis of the EPR for these patients, identified deranged biochemistry (elevated urea and/or creatinine concentrations, hyperkalaemia or metabolic acidosis) (n=31), fluid overload (n=16), hyperlactataemia (n=4) and sepsis (n=1) as the indications for RRT initiation. In addition, while the novel model accurately predicted severe oliguria, less than 20% of those who suffered severe oliguria went on to require RRT. Currently, creatinine concentration performs a key role in the identification of those at risk of adverse outcomes related to renal dysfunction. The existing KDIGO(1) creatinine criteria - which are shared by the AKIN and RIFLE guidelines (22, 23) - have been shown to stratify risk accurately in both cardiac surgery patients (24, 25) and the general inpatient population (26, 27). Similarly, recent advances in the use of biomarkers have been shown to enable the early identification of those at increased risk of adverse outcomes related to renal dysfunction (28-30). Moreover, the combination of biomarkers and serum creatinine analyses increases the accuracy of patient risk classification. (29, 30) Future work should focus on integrating the novel analysis of urine output described in this study with other physiological variables measured in real-time together with biomarker and serum creatinine results to optimise the early detection of deranged renal physiology.

Limitations

Most patients in this study received interventions with the intention of normalising urine output. A total of 488 (20.4%) patients received diuretics during their ICU admission. Data on the success of such interventions has not been investigated as part of this study but is likely to be of value as part of future work. The development of this model benefited from being conducted in a group of patients undergoing cardiac surgery in one institution where the risk of complications is well known but the single centre design could limit transferability across other health care settings. The methodology developed will therefore need to be validated in different patient groups and in different institutions. With appropriate development, it could easily be applicable to all intensive care unit patients. The ability of the model to improve

patient outcome through early recognition of impending severe oliguria should then be tested.

Conclusions

This dynamic Bayesian model, which analyses a patient's current urine output in the context of their previous urine output, can be used to accurately predict the risk of severe oliguria occurring within the next 12 hours. Classification according to the model's predictions was shown to outperform the current method for screening patient urine output; the KDIGO AKI criteria. Crucially, the use of dynamic Bayesian modelling allows those at high-risk to be identified before they suffer a prolonged period of severe oliguria and in time to offer treatment. The model requires no additional information other than hourly urine output values and the patient's weight, can be easily run by computers routinely available at the bedside and provides an output that is easily interpreted by the clinical team. Before widespread adoption, the model requires validation in a range of critical care units and across the full range of critical care patients.

References

Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group. KDIGO Clinical Practical Guideline for Acute Kidney Injury in *Kidney Inter* 2012. 1-138.

1. Group KDIGOKAKIW. KDIGO Clinical Practical Guideline for Acute Kidney Injury In: Kidney Inter; 2012. p. 1-138.
2. Kellum JA, Sileanu FE, Murugan R, et al. Classifying AKI by Urine Output versus Serum Creatinine Level. *Journal of the American Society of Nephrology* 2015;26:2231-2238.
3. Joannidis M, Metnitz B, Bauer P, et al. Acute kidney injury in critically ill patients classified by AKIN versus RIFLE using the SAPS 3 database. *Intensive Care Med* 2009;35(10):1692-1702.
4. Vives M, Wijeysundera D, Marczin N, et al. Cardiac surgery-associated acute kidney injury. *Interactive cardiovascular and thoracic surgery* 2014;18(5):637-645.
5. Hansen MK, Gammelager H, Jacobsen CJ, et al. Acute Kidney Injury and Long-term Risk of Cardiovascular Events After Cardiac Surgery: A Population-Based Cohort Study. *J Cardiothorac Vasc Anesth* 2015.
6. Machado MN, Nakazone MA, Maia LN. Prognostic value of acute kidney injury after cardiac surgery according to kidney disease: improving global outcomes definition and staging (KDIGO) criteria. *Plos One* 2014;9(5):e98028.
7. Dasta JF, Kane-Gill SL, Durtschi AJ, et al. Costs and outcomes of acute kidney injury (AKI) following cardiac surgery. *Nephrol Dial Transplant* 2008;23(6):1970-1974.
8. Thakar CV, Christianson A, Freyberg R, et al. Incidence and outcomes of acute kidney injury in intensive care units: a Veterans Administration study. *Crit Care Med* 2009;37(9):2552-2558.
9. Ralib A, Pickering JW, Shaw GM, et al. The urine output definition of acute kidney injury is too liberal. *Critical Care* 2013;17(3):R112.
10. Howitt SH, Grant SW, Caiado C, et al. The KDIGO acute kidney injury guidelines for cardiac surgery patients in critical care: a validation study. *BMC Nephrol* 2018;19(1):149.
11. Lagny MG, Jouret F, Koch JN, et al. Incidence and outcomes of acute kidney injury after cardiac surgery using either criteria of the RIFLE classification. *BMC Nephrol* 2015;16:76.
12. Pole A, West M, Harrison J. Applied Bayesian forecasting and time series analysis. New York: Chapman and Hall; 1994.
13. West M, Harrison PJ, Migon HS. Dynamic Generalized Linear Models and Bayesian Forecasting. *Journal of the American Statistical Association* 1985;80(389):73-83.
14. Hickey GL, Grant SW, Caiado C, et al. Dynamic prediction modeling approaches for cardiac surgery. *Circ Cardiovasc Qual Outcomes* 2013;6(6):649-658.
15. Smith AFM, West M. Monitoring Renal Transplants: An Application of the Multiprocess Kalman Filter. *Biometrics* 1983;39(4):867-878.
16. R Core Team. R: A Language and Environment for Statistical Computing. In. Vienna, Austria: R Foundation for Statistical Computing 2015.
17. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44(3):837-845.
18. Saito T, Rehmsmeier M. The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *Plos One* 2015;10(3):e0118432.
19. Crowson CS, Atkinson EJ, Therneau TM. Assessing Calibration of Prognostic Risk Scores. *Statistical methods in medical research* 2016;25(4):1692-1706.
20. Roques F, Michel P, Goldstone AR, et al. The logistic EuroSCORE. *Eur Heart J* 2003;24(9):881-882.

21. Hickey GL, Grant SW, Murphy GJ, et al. Dynamic trends in cardiac surgery: why the logistic EuroSCORE is no longer suitable for contemporary cardiac surgery and implications for future risk models. *Eur J Cardiothorac Surg* 2013;43(6):1146-1152.
22. Mehta RL, Kellum JA, Shah SV, et al. Acute Kidney Injury Network: report of an initiative to improve outcomes in acute kidney injury. *Crit Care* 2007;11.
23. Bellomo R, Ronco C, Kellum JA, et al. Acute renal failure–definition, outcome measures, animal models, fluid therapy and information technology needs: the Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group. *Crit Care* 2004;8.
24. Machado MDN, Miranda RC, Takakura IT, et al. Acute kidney injury after on-pump coronary artery bypass graft surgery [Spanish;Portuguese;English] Lesion renal aguda post-revascularizacion del miocardio con circulacion extracorporea. *Arquivos Brasileiros de Cardiologia* 2009;93(3):230-235+242-247+247-252.
25. Birnie K, Verheyden V, Pagano D, et al. Predictive models for kidney disease: improving global outcomes (KDIGO) defined acute kidney injury in UK cardiac surgery. *Critical Care* 2014;18(6):606.
26. Fujii T, Uchino S, Takinami M, et al. Validation of the Kidney Disease Improving Global Outcomes Criteria for AKI and Comparison of Three Criteria in Hospitalized Patients. *Clinical Journal of the American Society of Nephrology* 2014.
27. Cheng X, Wu B, Liu Y, et al. Incidence and diagnosis of Acute kidney injury in hospitalized adult patients: a retrospective observational study in a tertiary teaching Hospital in Southeast China. *BMC Nephrol* 2017;18(1):203.
28. Banos V, Betbese A, Garcia-Alvarez M, et al. Urinary neutrophil gelatinase-associated lipocalin as an early predictor of acute kidney injury in cardiac surgery patients. *European journal of anaesthesiology* 2013;30.
29. Kashani K, Al-Khafaji A, Ardiles T, et al. Discovery and validation of cell cycle arrest biomarkers in human acute kidney injury. *Critical Care* 2013;17(1):R25-R25.
30. Egal M, de Geus HR, Groeneveld AB. Neutrophil Gelatinase-Associated Lipocalin as a Diagnostic Marker for Acute Kidney Injury in Oliguric Critically Ill Patients: A Post-Hoc Analysis. *Nephron* 2016;134(2):81-88.

Figure Legends

Figure 1 Legend. Receiver operating characteristic curves for the prediction of severe oliguria (<0.3ml/kg/hr for 6 hours) during the next 12 hours following each predictions made by the model at 12, 24, 36, 48 and 72 hours.

Figure 2 Legend. Calibration plots for the Bayesian model's prediction of severe oliguria (0.3ml/kg/hr for 6 hours) during the next 12 hours at a)12 hours, b)24 hours, c)36 hours, d)48 hours and e)72 hours. Patients were grouped into deciles according to predicted risk. For each of the ten groups mean observed risk is plotted against mean predicted risk.

Figure 1

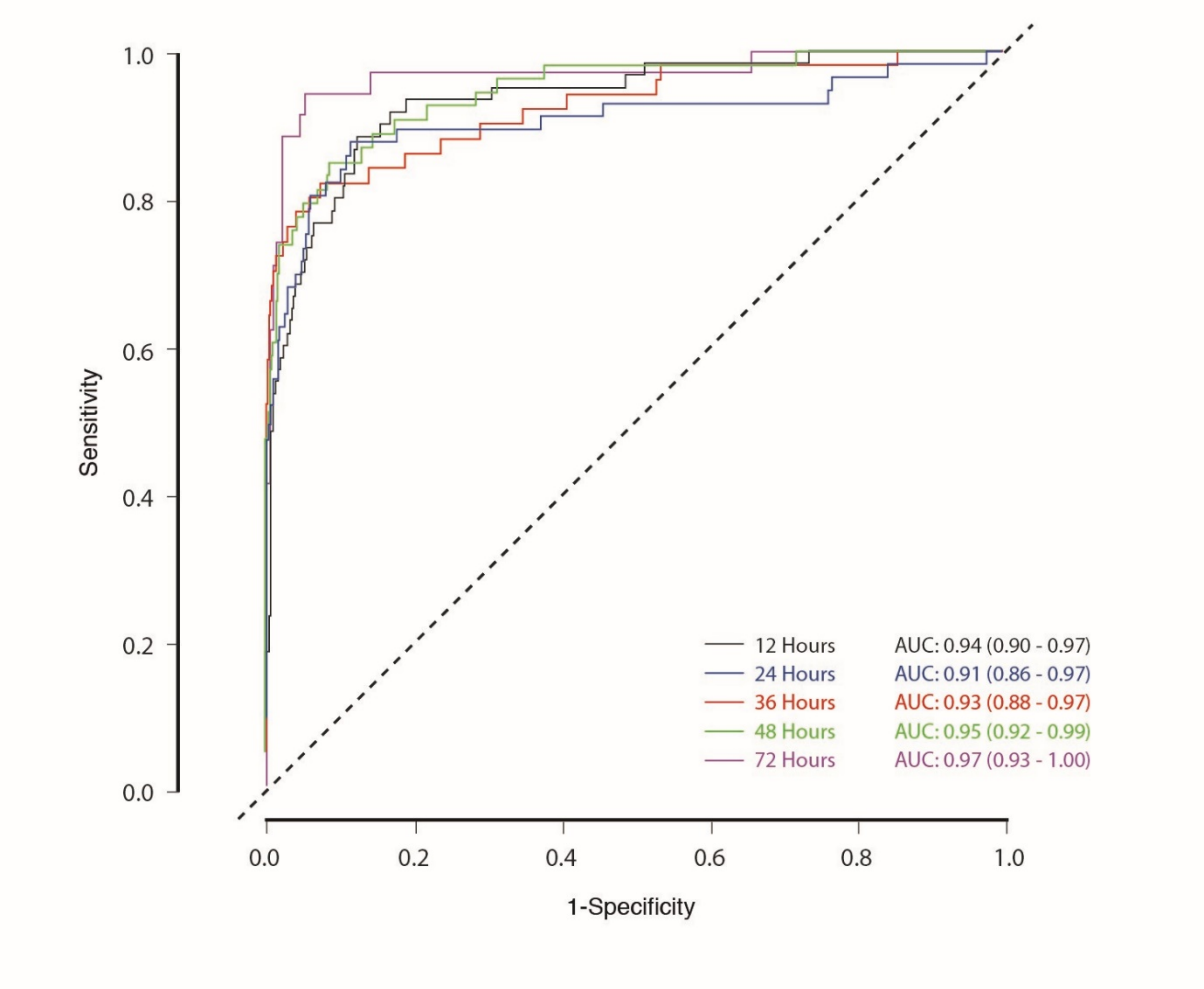
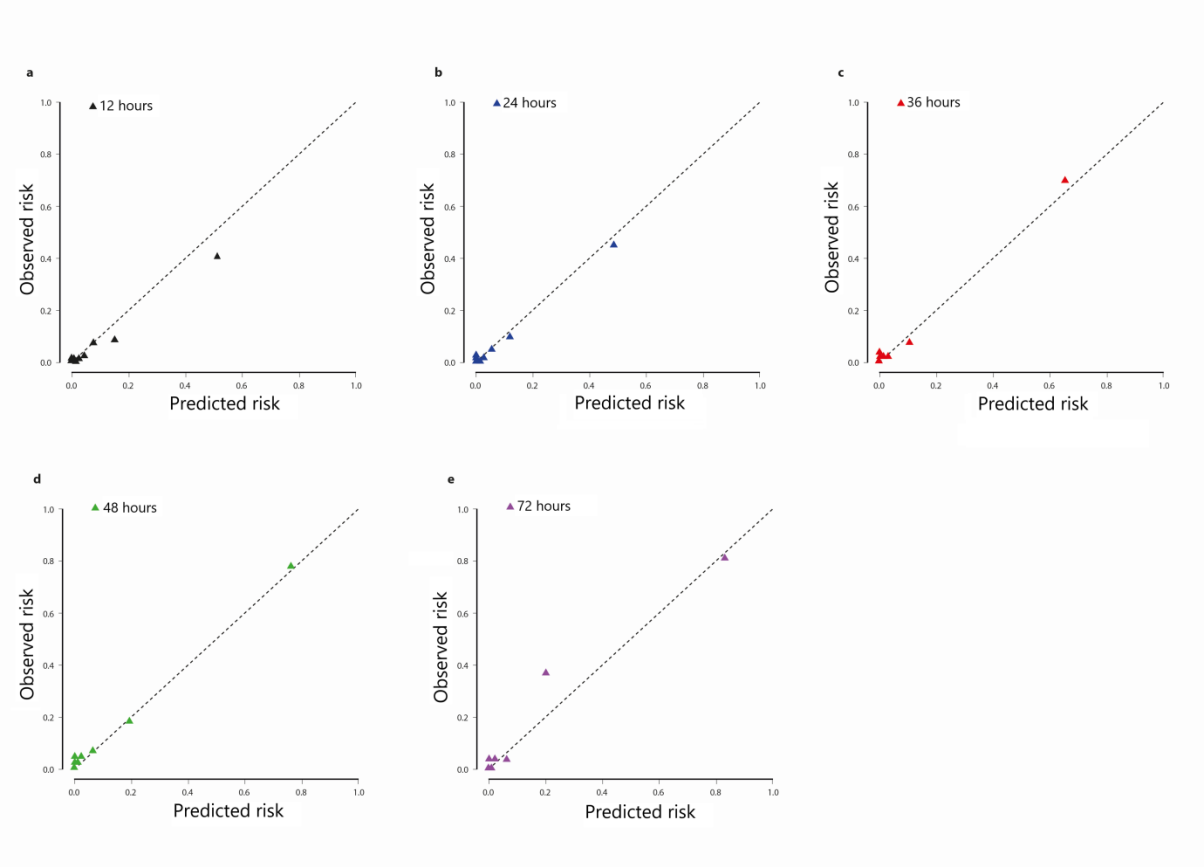

1.0
0.8
0.6
0.4
0.2
0.0
Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
1-Specificity
12 Hours AUC: 0.94 (0.90 - 0.97)
24 Hours AUC: 0.91 (0.86 - 0.97)
36 Hours AUC: 0.93 (0.88 - 0.97)
48 Hours AUC: 0.95 (0.92 - 0.99)
72 Hours AUC: 0.97 (0.93 - 1.00)

Figure 2


a
12 hours
Observed risk
Predicted risk
b
24 hours
Observed risk
Predicted risk
c
36 hours
Observed risk
Predicted risk
d
48 hours
Observed risk
Predicted risk
e
72 hours
Observed risk
Predicted risk

Tables

Table 1 – Patient Characteristics

Characteristic	Development group (n=981)	Validation group (=2389)
Age, mean (sd) , years	66.4 (11.2)	66.7 (10.9)
Female gender, %	279 (28.2)	660 (27.6)
Weight, mean (sd), Kg	82.2 (15.9)	81.8 (16.4)
Logistic EuroSCORE, median (Interquartile range)	3.8 (2.1-7.4)	3.7 (2.0-7.0)
Operation, n (%)		
CABG	544 (55.5)	1394 (58.4)
Valve	227 (23.1)	505 (21.1)
CABG and Valve	125 (12.7)	337 (14.1)
Aortic	65 (6.6)	118 (5.0)
Other – minor	3 (0.3)	5 (0.2)
Other – major	17 (1.7)	30 (1.3)
Urgency, n (%)		
Elective	574 (58.5)	1380 (57.8)
Urgent	395 (40.3)	958(40.1)
Emergency	9 (0.9)	44 (1.8)
Salvage	3 (0.3)	7 (0.3)
CPB time, median (Interquartile range), minutes	102.0 (81.0-129.0)	102.0 (82.0-129.0)

Table 2 – Outcomes for patients grouped according to risk level as determined by analysis of urine output by KDIGO-AKI guideline and the Bayesian model.

Group	N (% of total)*	RRT, N (%)	PLOS, N (%)	Hospital mortality, N(%)
Low-risk by model and no KDIGO AKI	1725 (72.2%)	15 (0.9)	320 (18.6)	10 (0.6)
Low-risk by model but KDIGO AKI	506 (21.2)	18 (3.6)	176 (34.8)	12 (2.4)
High-risk by model but no KDIGO AKI	36 (1.5)	3 (8.3)	30(83.3)	3 (8.3)
High-risk by model and KDIGO AKI	122 (5.1)	26 (21.3)	73 (59.8)	11 (9.0)

KDIGO =Kidney Disease Improving Global Outcomes, UO = urine output, AKI = Acute Kidney Injury, PLOS = prolonged length of stay in hospital, RRT = renal replacement therapy

Table 3 - Performance of the Bayesian model, existing KDIGO AKI-UO criterion and severe oliguria when identifying those at risk of RRT.

Classification Method	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value
AKI-UO	0.74	0.75	0.08	0.99
Model	0.41	0.94	0.18	0.98
Severe oliguria	0.41	0.94	0.18	0.98

Severe oliguria = observed UO <0.3ml/kg for 6 hours, AKI-UO = observed UO <0.5ml/kg for 6 hours, RRT = renal replacement therapy

Appendix

Table A1 – Comparison of observed outcomes and model's predictions for severe oliguria

Time point (number of patients still on CICU)	Observed severe oliguria within 12 hours, n(%)	Predicted severe oliguria within 12 hours, n(%)	O:E ratio
12 hours (1947)	61 (3.1)	82 (4.2)	0.74
24 hours (1694)	57 (3.4)	61 (3.6)	0.93
36 hours (1137)	51 (4.5)	44 (3.9)	1.16
48 hours (909)	54 (5.9)	48(5.3)	1.13
72 hours (545)	35 (6.4)	30 (5.6)	1.15

Table A2 - Performance of models when predicting severe oliguria occurring with the next 6 hours

Time point (number of patients still on CICU)	AUC (95% CI)	Observed severe oliguria within 6 hours	Predicted severe oliguria within 6 hours	O:E ratio
12 hours	0.98 (0.96-0.99)	21	90	0.23
24 hours	0.98 (0.97-0.99)	30	61	0.49
36 hours	0.99 (0.98-1.00)	34	49	0.69
48 hours	0.99 (0.98-1.00)	36	49	0.73
72 hours	0.99 (0.98-1.00)	38	31	0.92

Table A3 -Outcome of patients according to classification by the Bayesian model

Group	RRT (n,%)	PLOS (n,%)	Mortality (n,%)
High-risk (n=158)	29(18.4)*	93 (58.9) *	14 (8.9) *
Low-risk (n=2231)	41(1.8%)	496 (22.2)	22 (1.0)

p<0.001 when compared to low-risk classification by the model

RRT = renal replacement therapy, PLOS = prolonged length of stay in hospital

Logistic regression models used to adjust for confounders during multivariable analyses

Table A4 – Model for prediction of Renal replacement therapy

Variable	Beta coefficient	Odds ratio	95% CI for Odds ratio	P value
Intercept	-4.53			<0.001
Model high-risk classification	2.34	10.36	5.86-18.07	<0.001
Logistic EuroSCORE	0.04	1.03	1.01-1.06	<0.001
CPB (minutes)	0.00	1.00	1.00-1.01	0.34

Table A5 – Model for prediction of prolonged length of stay

Variable	Beta coefficient	Odds ratio	95% CI for Odds ratio	P value
Intercept	-2.41			<0.001
Model high-risk classification	1.48	4.38	2.99-6.44	<0.001
Logistic EuroSCORE	0.07	1.08	1.06-1.09	<0.001
CPB (minutes)	0.01	1.00	1.00-1.01	<0.001

Table A6 – Model for prediction of hospital mortality

Variable	Beta coefficient	Odds ratio	95% CI for Odds ratio	P value
Intercept	-6.13			<0.001
Model high-risk classification	1.86	6.44	2.82-13.98	<0.001
Logistic EuroSCORE	0.03	1.03	1.00-1.05	0.06
CPB (minutes)	0.01	1.01	1.00-1.01	<0.001