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

Samira Bell,[1] Friedo W Dekker,[2] Thenmalar Vadiveloo,[3] Charis Marwick,[3] Harshal Deshmukh,[3] Peter T Donnan,[3] Merel Van Diepen[2]

[1]Renal Unit, Ninewells Hospital, NHS Tayside, Dundee DD1 9SY, UK

[2]Department of Clinical Epidemiology, Leiden University Medical Centre, Leiden, Netherlands

[3] Division of Population Health Sciences , School of Medicine, University of Dundee, Dundee, UK

Correspondence to: S Bell samira.bell@nhs.net

ABSTRACT

STUDY QUESTION

What is the predicted risk of acute kidney injury after orthopaedic surgery and does it affect short term and long term survival?

METHODS

The cohort comprised adults resident in the National Health Service Tayside region of Scotland who underwent orthopaedic surgery from 1 January 2005 to 31 December 2011. The model was developed in 6220 patients (two hospitals) and externally validated in 4395 patients from a third hospital. Several preoperative variables were selected for candidate predictors, based on literature, clinical expertise, and availability in the orthopaedic surgery setting. The main outcomes were the development of any severity of acute kidney injury (stages 1-3) within the first postoperative week, and 90 day, one year, and longer term survival.

STUDY ANSWER AND LIMITATIONS

Using logistic regression analysis, independent predictors of acute kidney injury were older age, male sex, diabetes, number of prescribed drugs, lower estimated glomerular filtration rate, use of angiotensin converting enzyme inhibitors or angiotensin receptor blockers, and American Society of Anesthesiologists grade. The model's predictive performance for discrimination was good (C statistic 0.74 in development cohort, 0.70 in validation cohort). Calibration was good in the development cohort and after recalibration in the validation cohort. Only the highest risks were over-predicted. Survival was worse in patients with acute kidney injury compared with those without (adjusted hazard ratio 1.53, 95% confidence interval 1.38 to 1.70). This was most noticeable in the short term (adjusted hazard ratio: 90 day 2.36, 1.94 to 2.87) and diminished over time (90 day-one year 1.40, 1.10 to 1.79; >1 year 1.28, 1.10 to 1.48). The model

used routinely collected data in the orthopaedic surgery setting therefore some variables that could potentially improve predictive performance were not available. However, the readily available predictors make the model easily applicable.

WHAT THIS STUDY ADDS

A preoperative risk prediction model consisting of seven predictors for acute kidney injury was developed, with good predictive performance in patients undergoing orthopaedic surgery. Survival was significantly poorer in patients even with mild (stage 1) postoperative acute kidney injury.

FUNDING, COMPETING INTERESTS, DATA SHARING

SB received grants from Tenovus Tayside, Chief Scientist Office, and the Royal College of Physicians and Surgeons of Glasgow; PT receives grants from Novo Nordisk, GlaxoSmithKline, and the New Drugs Committee of the Scottish Medicines Consortium. No additional data are available.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Acute kidney injury is common in patients undergoing surgery

No externally validated preoperative risk scores are in common use for patients undergoing non-cardiac surgery

WHAT THIS STUDY ADDS

A preoperative risk prediction model consisting of seven predictors for acute kidney injury was developed, with good predictive performance in patients undergoing orthopaedic surgery .

Survival was significantly poorer in patients with even mild (stage 1) postoperative acute kidney injury

Introduction

Acute kidney injury affects 1 in 5 people during hospital stay[1] and the condition is associated with considerably increased mortality. Increasing evidence shows that even mild, transient degrees of acute kidney injury are associated with both increased long term mortality and the future development of chronic kidney disease, independent of other factors.[2-5] In addition to these adverse health outcomes, acute kidney injury has a major economic impact. Increased costs result from increased length of hospital stay, higher number of investigations, admission to an intensive care unit, and renal replacement therapy. The annual cost of acute kidney injury to the National Health Service across the United Kingdom (not including cases in the community) is estimated at between £434m ($664m; €603m) and £620m.[6]

Surgery is an important cause of acute kidney injury, but the true incidence has been difficult to establish owing to the lack of a universal definition. A large US study showed that acute kidney injury affected 1% of patients undergoing non-cardiac surgery. This is likely to be an under-estimate as the definition used for acute kidney injury was an increase in serum creatinine levels of 176.8 μmol/L, corresponding to severe acute kidney injury.[7] Since the adoption of a universally accepted definition for acute kidney injury (Kidney Disease Improving Global Outcomes criteria),[8] we have shown using these criteria that the rates of acute kidney injury ranged from 6% to 12% in gastrointestinal

surgery and 23% to 25% in vascular surgery.[9] Several studies have described the risk factors for postoperative acute kidney injury.[10-13] To guide management, guidelines on acute kidney injury from both the Kidney Disease Improving Global Outcomes and National Institute for Health and Care Excellence highlight the importance of identifying patients at high risk for developing acute kidney injury.[8 14] Preoperative identification of high risk patients would allow for earlier intervention and optimal perioperative management thereby improving patient outcomes. Early identification and prevention of acute kidney injury is vital because once the condition is established, mortality is extremely high and the only treatment is supportive, necessitating renal replacement therapy in severe cases. Currently, however, no externally validated preoperative risk scores are in common usage for non-cardiac surgery.

We developed and validated a risk score to predict postoperative acute kidney injury in patients undergoing orthopaedic surgery and thereby identify people at high risk before surgery. To highlight the potential importance of the risk score, we also examined the impact of acute kidney injury on both short term and long term survival.

Methods

Study population

We used a development cohort and validation cohort for the risk prediction model. The development cohort included all adults aged more than 18 years who underwent orthopaedic surgical procedures (see supplementary appendix 1) at a large teaching hospital in Dundee (Ninewells Hospital), Scotland, or a neighbouring small hospital (Stracathro Hospital) from 1 January 2005 to 31 December 2011 and who were resident or died in the NHS Tayside region. For external validation we included patients from another hospital (Perth Royal Infirmary) in Scotland in the validation cohort. From both cohorts we excluded patients with missing preoperative or postoperative creatinine measurements and those receiving chronic renal replacement therapy. We combined both cohorts for the survival analysis. From the general register of deaths we extracted data on deaths until 30 June 2014.

Ethical statement

We carried out anonymised record linkage according to the standard operating procedure of the Health Informatics Centre, University of Dundee.[15] The Tayside research ethics committee does not require submission of individual studies that follow this standard operating procedure, which is Caldicott Guardian approved. Reporting and methods of the study adhered to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement.[16-18]

Patient involvement

No patients were involved in setting the research question or the outcome measures, nor were they involved in the design and implementation of the study. There are no plans to involve patients in dissemination of the results.

Data sources

Data were linked using the Health Informatics Centre, University of Dundee.[15] This centre enables anonymised linkage of health records from the population of Tayside (400 000), Scotland, using a unique identifying community health index number. We linked data between the following datasets: Scottish morbidity record of hospital admissions (SMR01); laboratory results, medicines dispensed by community pharmacies, Scottish index of multiple deprivation, the Scottish Care Initiative-Diabetes Collaboration, the Scottish renal registry, and the theatre system, which records information completed by the anaesthetist before surgery. From the operation theatre system we identified operations by the operation procedure codes (see supplementary appendix 1). The Scottish morbidity record of hospital admissions provides information on age, sex, and postcode, and admission and discharge dates. From the laboratory system we obtained creatinine measurements. We obtained the deprivation category from the Scottish index of multiple deprivation, and the Scottish Care Initiative-Diabetes Collaboration provided information on diabetes type and date of diagnosis. Using the Scottish renal registry, we identified patients receiving chronic dialysis or who had undergone renal transplantation. The operation theatre system provided information on type and urgency of operation in addition to the American Society of Anesthesiologists grade.

Outcomes

The outcome of interest for the prediction model was the development of any severity of acute kidney injury (stages 1-3) during the first postoperative week. Severity was defined using the Kidney Disease Improving Global Outcomes criteria.[8] We applied these criteria, using serum creatinine concentration at baseline (most recent before surgery) as preoperative measurement and maximal serum creatinine concentration during the first seven postoperative days as post-measurement. To be classified as having postoperative acute kidney injury, patients must have had at least mild acute kidney injury (stage 1), defined as an increase in serum creatinine level of at least 26.4 µmol/L or a postoperative serum creatinine level greater than 150% of baseline serum creatinine measurement. Because of inaccuracies in both measurement and recording of urine output, we did not apply the Kidney Disease Improving Global Outcomes urine ouput criteria for acute kidney injury to this variable.

The primary outcome for the survival analysis was postoperative all cause mortality. Secondary outcomes were short term mortality (up to 90 days after surgery), intermediate term mortality (up to one year after surgery), and long term mortality (to end of follow-up).

Candidate predictors

Based on previous literature and guidelines, the clinical knowledge of the study team, and the available data, we selected several preoperative variables for candidate predictors of acute kidney injury.[7 14] We obtained age to the nearest year in the year of surgery, sex from the community health index register, and social deprivation from the Scottish index of multiple deprivation (linked to

postcode from the community health index register). We obtained information on diabetes from the Scottish Care Initiative-Diabetes Collaboration, a Scottish national diabetes registry. From dispensed prescribing data we ascertained the number of dispensed prescriptions from community pharmacies before admission and use of drugs that predispose to renal impairment (non-steroidal anti-inflammatory drugs (NSAIDs), cyclo-oxygenase-2 selective inhibitors, angiotensin converting enzyme (ACE) inhibitors, and angiotensin II receptor antagonists). We classified regularly prescribed drugs as none, one or two drugs, or three or more drugs, and we combined use of ACE inhibitors or angiotensin II receptor antagonists into one variable as well as use of NSAIDs or cyclo-oxygenase-2 selective inhibitors. Owing to recent reports of the association between statins and acute kidney injury, we examined these drugs separately.[19] From the laboratory database we obtained baseline renal function, which was the most recent, preoperative serum creatinine measurement. This could include preoperative samples taken during the current admission for elective surgery. Patients undergoing emergency surgery, however, may have acute kidney injury on admission to hospital because of trauma. We therefore used the most recent serum creatinine concentration before admission for emergency patients to distinguish chronic kidney disease from preoperative acute kidney injury. The CKD-EPI equation was used to calculate an estimated glomerular filtration rate.[20] We then categorised the rates as greater than 60, 45-59, 30-44, and less than 29. Variables used from the theatre system included the American Society of Anesthesiologists grade, which categorises patients into five categories: 1=normal, healthy individual; 2=mild systemic disease; 3=severe systemic disease, not incapacitating; 4=incapacitating systemic disease that is a threat to life; and 5=moribund person who is not expected to survive without the operation,[21] and whether the operation was elective, expedited, or an emergency. As the emergency operations were mainly related to trauma we used this variable as a proxy for trauma.

Statistical analysis
For continuous variables we present the baseline characteristics as means and standard deviations or medians and interquartile ranges, depending on distribution. We present categorical variables as numbers and percentages.

Risk prediction
We entered 11 candidate predictors (age at operation, sex, baseline renal function, diabetes, number of prescribed drugs, treatment with ACE inhibitor or angiotensin receptor blocker, treatment with NSAID or cyclo-oxygenase-2 selective inhibitors, treatment with statin, urgency of operation, American Society of Anesthesiologists grade, and deprivation category) into a multivariable logistic regression analysis, with acute kidney injury as the dependent variable. We used backward selection to identify the most important independent predictors, and applied a conservative selection criterion of P<0.15 to limit chances of over-fitting. Repeating the analysis using forward selection to check

whether this resulted in the same list of predictors, ensured the model's stability. In order to calculate an individual patient's risk of acute kidney injury, we first calculated his or her prognostic index by multiplying the estimated coefficients with the values of the predictor variables of the patient and taking the sum of these multiplications and adding the intercept of the model. Using the prognostic index (PI) we then calculated the risk of acute kidney injury as exp(PI)/(1+exp(PI)).

As the apparent predictive performance (performance in the development cohort) usually over-estimates the performance in other patients, owing to over-fitting of the idiosyncrasies in the development cohort, we internally validated the model through boot-strapping[19 22] and then externally validated it in our validation cohort. To assess predictive performance, we estimated discrimination and calibration.

Discrimination indicates how well the model can distinguish patients with the outcome from those without the outcome, represented by Harrell's C statistic.[23] This is equal to the area under the receiver operator curve in a logistic regression model and lies between 0.5 and 1; values greater than 0.7 indicate good predictive performance and values greater than 0.8 indicate excellent predictive performance.

Calibration ascertains the concordance between the model's predictions and observed outcomes, which we evaluated using a calibration plot.[24] The plot displays the predicted risk versus the observed frequency of acute kidney injury in 10ths of increasing predicted risk, augmented by a smoothed (lowess) regression line.[25] Ideally the plot follows a 45° line, showing that the predicted risks are equal to the observed outcome frequencies. In addition, we assessed the difference in predicted and observed frequency in the total cohort, indicating the extent to which predictions are systematically too high or too low, referred to as calibration-in-the-large.[24] Finally, we assessed the calibration slope, reflecting the slope of the calibration plot and ideally equal to 1. A slope less than 1 indicates the degree of over-fitting of the model.[26]

When the prediction model is not well calibrated externally, predictions can be recalibrated in various ways.[27 28] We examined two methods of recalibration, requiring different levels of input from the new setting. The first method corrects for systematic over-prediction or under-prediction, generally caused by a difference in outcome incidence in the development and validation cohort. A correction factor, computed from the observed frequency of acute kidney injury and the predicted risk of acute kidney injury in the external cohort, is added to the intercept of the model.[29] The second method in addition also corrects for over-fitting—that is, for a calibration slope less than 1. For this method the intercept and regression coefficient of the logistic regression model with the prognostic index as the only predictor (recalibration model) are used to transform the prognostic index and compute recalibrated probabilities from this recalibrated prognostic index.[26] Although the second method is more accurate for calibration slopes less than 1, it is also less practical as it requires estimating the recalibration model in the new setting,

instead of merely adding an easily computable correction factor to the intercept.

To assess robustness of the model we performed several sensitivity analyses: using multiple imputation with fully conditional specification to impute missing data,[30-32] relaxing the backward selection removal criterion to 0.20 and restricting it to 0.10, and adding non-linear and interaction terms to the model.

Survival analysis

Using Kaplan-Meier plots and log rank analyses we calculated the cumulative survival of patients with compared with those without acute kidney injury. We used Cox regression analyses to calculate hazard ratios and 95% confidence intervals for patients with any acute kidney injury compared with those without acute kidney injury, adjusted for confounders. We selected the confounders age, sex, diabetes, baseline renal function, and number of prescribed drugs based on literature, guidelines, and clinical expertise.[8 14] We examined 90 day, one year, and survival until end of follow-up, where one year survival was conditional on surviving 90 days and long term survival was conditional on surviving one year.[33] Additionally, we analysed overall survival by stage of acute kidney injury. We performed bootstrap analysis using the rms package in R.[34] All other analyses were carried out in IBM SPSS (v21). For survival analyses we considered a P value of less than 0.05 to be statistically significant

Results

We identified 15 218 orthopaedic operations, as defined by the included operation procedure codes in the theatre operation system (see supplementary appendix 1), between 1 January 2005 and 31 December 2011. Both preoperative and postoperative creatinine values were available for 12 530 of these operations (17.7% missing data). We used the first operation for each patient and excluded subsequent operations in the same individual, leaving 10 615 patients in the analysis (6220 in the development cohort and 4395 in the validation cohort).

Baseline characteristics

Table 1 shows the baseline characteristics of the two cohorts. Mean age was similar between both (70.7 (SD 15.3) in the development cohort, 71.0 (12.3) in the validation cohort). Baseline estimated glomerular filtration rate was also similar between the groups (71 (SD 23) mL/min in the development cohort, 72 (20) mL/min) in the validation cohort). Postoperative acute kidney injury occurred in 672 (10.8%) of the patients in the development cohort and 295 (6.7%) in the validation cohort. Emergency surgery occurred most in the development cohort (28% v 13%).

Predictive variables for acute kidney injury

Logistic regression with backward selection resulted in the final model with seven predictors: age at operation, male sex, diabetes, lower estimated glomerular filtration rate, use of ACE inhibitors or angiotensin receptor blockers, number of prescribed drugs and American Society of Anesthesiologists grade. Table 2 presents the estimates. Forward selection led to the same results. Sensitivity

analyses also showed robustness of the model. Using multiple imputation for missing American Society of Anesthesiologists grade and deprivation category (926 (14.9%) missing American Society of Anesthesiologists grade, 80 (13%) missing Scottish index of multiple deprivation category) made no difference to the variables in the model and resulted in similar discrimination (C statistic 0.74, 95% confidence interval 0.73 to 0.75)). Relaxation and restriction of the removal criterion for backward selection to P<0.20 and P<0.10, respectively, did not change the final model. Adding age as a non-linear term, and age, sex, diabetes, type of surgery (elective, expedited, and emergency), chronic kidney disease category, and number of prescribed drugs as interaction terms with the other variables slightly altered the final predictor list but did not improve the model's performance.

Model validation

Calibration of the model was assessed by comparing observed to predicted risk in a calibration plot (fig 1). This plot shows, for example, that applying a cut-off of 10%, meaning that patients with a predicted risk higher than 10% are classified as at high risk for acute kidney injury, would affect about 40% of patients. Thus this plot can help in deciding the cut-off above which to target interventions. From these data, the inference is that using a cut-off of 10%, which is about the incidence of acute kidney injury in this cohort, more than 70% of acute kidney injury cases would be identified. Calibration was suboptimal in the validation cohort showing that the model over-predicts the risk of acute kidney injury in that cohort (fig 1). This was also shown by the calibration-in-the-large of 10.4% predicted risk compared with 6.6% observed risk and the calibration slope of 0.79 (table 3), and this could well be a consequence of the difference in incidence of acute kidney injury in each cohort. Recalibration using the first method improved calibration considerably (fig 2), resulting in a calibration-in-the-large of 6.8% compared with 6.6% (table 3). The second recalibration method using the full recalibration model to recalibrate prognostic indices (slope 0.79, intercept −0.93) resulted in a slightly better calibration-in-the-large than using the first method (6.6% v 6.6%) and a slope of 1 (by definition) (fig 2 and table 3). These results show that the difference between method 1 and method 2 in this case mainly stems from the highest 10th of predicted risk. After recalibration there was still some over-prediction at the highest end of the spectrum, but this would not further affect clinical decision making as these patients were already at high risk. We examined discrimination by calculating the C statistic using the receiver operator curve. The C statistic of the model was 0.74 (95% confidence interval 0.72 to 0.76), indicating reasonably good predictive performance. Performance was only slightly less in both the internal (0.73) and the external validation (0.70), indicating good external validity of the model (table 3). For further insight into the predictive value of the model, we assessed the incidence of acute kidney injury across 5% risk groups. Results in the validation cohort again showed that the model over-predicts the actual incidence of acute kidney injury, hence the need for recalibration (table 4).

Table 1 | Baseline characteristics of patients in the development and validation cohorts. Values are numbers (percentages) unless stated otherwise

Characteristics	Development cohort (n=6220)			Validation cohort (n=4395)		
	Acute kidney injury (n=672, 10.8%)	No acute kidney injury (n=5548)	Overall	Acute kidney injury (n=295, 6.7%)	No acute kidney injury (n=4100)	Overall
Mean (SD) age (years)	76.5 (11.1)	70.0 (15.6)	70.7 (15.3)	75.3 (10.2)	70.7 (12.3)	71.0 (12.3)
Men	314 (47)	1980 (36)	2294 (37)	144 (49)	1535 (37)	1679 (38)
Women	358 (53)	3568 (64)	3926 (63)	151 (51)	2565 (63)	2716 (62)
Median (interquartile range) length of hospital stay (days)	9 (5-15)	8 (5-12)	11 (5-12)	7 (4-10)	6 (5-8)	6 (5-9)
Baseline eGFR (mL/min)	61 (25)	72 (22)	71 (23)	65 (21)	72 (19)	72 (20)
Chronic kidney disease category (eGFR mL/min):						
>60	324 (48.2)	3846 (69.3)	4170 (67.0)	180 (61.0)	2963 (72.3)	3143 (71.5)
45-59	155 (23.1)	1077 (19.4)	1232 (19.8)	62 (21.0)	788 (19.2)	850 (19.3)
30-44	126 (18.8)	491 (8.9)	617 (9.9)	39 (30.2)	286 (7.0)	325 (7.4)
<29	67 (10.0)	134 (2.4)	201 (3.3)	14 (4.7)	63 (1.5)	77 (1.7)
SIMD deprivation category:						
1 (most deprived)	144 (21.4)	1026 (18.5)	1170 (18.8)	20 (6.8)	279 (6.8)	299 (6.8)
2	123 (18.3)	932 (16.8)	1055 (17.0)	50 (16.9)	471 (11.5)	521 (11.9)
3	105 (15.6)	992 (17.9)	1097 (17.6)	57 (19.3)	761 (18.6)	818 (18.6)
4	199 (29.6)	1671 (30.1)	1870 (30.1)	100 (33.9)	1613 (39.3)	1713 (39.0)
5 (least deprived)	89 (13.2)	859 (15.5)	948 (15.2)	68 (23.1)	919 (22.4)	987 (22.5)
Missing	12 (1.8)	12 (1.8)	80 (1.3)	0	57 (1.4)	57 (13.0)
No receiving NSAIDs	101 (15)	991 (17.9)	1092 (17.6)	56 (19.0)	1112 (27.1)	1168(26.6)
No receiving ACE inhibitor or angiotensin II receptor antagonist	322 (47.9)	1371 (24.7)	1693 (27.2)	157 (53.2)	1124 (27.4)	1281 (29.1)
No receiving statins	293 (43.6)	1595 (28.7)	1888 (30.4)	117 (39.7)	1138 (27.8)	1255 (28.6)
No of prescribed drugs:						
None	136 (20.2)	2090 (37.7)	2226 (35.8)	56 (19.0)	1301 (31.7)	1357 (30.9)
1 or 2	291 (43.3)	2423 (43.7)	2714 (43.6)	133 (45.1)	1950 (47.6)	2083 (47.4)
≥3	245 (36.5)	1035 (18.7)	1280 (20.6)	106 (35.9)	849 (20.7)	955 (21.7)
Operation urgency:						
Elective	312 (46.4)	2576 (46.4)	2888 (46.4)	212 (71.9)	2973 (72.5)	3185(72.5)
Expedited	174 (25.9)	1410 (25.4)	1584 (25.5)	50 (16.9)	588 (14.3)	638 (14.5)
Emergency	186 (27.7)	1562 (28.2)	1748 (28.1)	33 (11.2)	539 (13.1)	572 (13.0)
Diabetes	146 (21.7)	524 (9.4)	670 (10.8)	50 (16.9)	371 (9.0)	421 (9.6)
ASA grade:						
1	16 (2.4)	675 (12.2)	691 (11.1)	19 (6.4)	430 (10.5)	449 (10.2)
2	212 (31.5)	2070 (37.3)	2282 (36.7)	101 (34.2)	1810 (44.1)	1911 (43.5)
3	374 (40.8)	1609 (29.0)	1886 (30.3)	89 (30.2)	934 (22.8)	1023 (23.3)
4	70 (10.4)	368 (6.6)	438 (7.0)	31 (10.5)	221 (5.4)	252 (5.7)
Missing	100 (10.4)	826 (14.9)	926 (14.9)	55 (18.6)	705 (17.2)	760 (17.3)

eGFR=estimated glomerular filtration rate; SIMD=Scottish index of multiple deprivation; NSAIDs=non-steroidal anti-inflammatory drugs; ACE=angiotensin converting enzyme; ASA=American Society of Anesthesiologists.

Survival

A total of 3166 (30%) out of 10 615 patients died over a median follow-up of 4.58 years (interquartile range 2.93-6.69 years). Figure 3 shows the Kaplan-Meier plot of cumulative survival in patients with acute kidney injury compared with those without acute kidney injury. Fifty one (0.5%) out of the 10 615 patients died within seven days without developing acute kidney injury. Overall survival in patients with acute kidney injury was worse than in patients without acute kidney injury (crude hazard ratio 2.04, 95% confidence interval 1.84 to 2.26, P<0.001; adjusted hazard ratio 1.53, 1.38 to 1.70, P<0.001; table 5). The effect of acute kidney injury was most noticeable in the short term (90 day crude hazard ratio 3.24, 95% confidence interval 2.68 to 3.91, P<0.001; adjusted hazard ratio 2.36, 1.94 to 2.87, P<0.001) and diminished over time (90 day to one year crude hazard ratio 1.89, 1.49 to 2.41, P<0.001; adjusted hazard ratio 1.40, 1.10 to 1.79, P=0.007; >1 year crude hazard ratio 1.68, 1.45 to 1.94, P<0.001; adjusted hazard ratio 1.28, 1.10 to 1.48, P=0.001). Survival worsened with increasing severity of acute kidney injury (table 6). Nevertheless, even mild acute kidney injury resulted in a considerably worse overall survival (adjusted hazard ratio 1.46, 1.30 to 1.63, P<0.001).
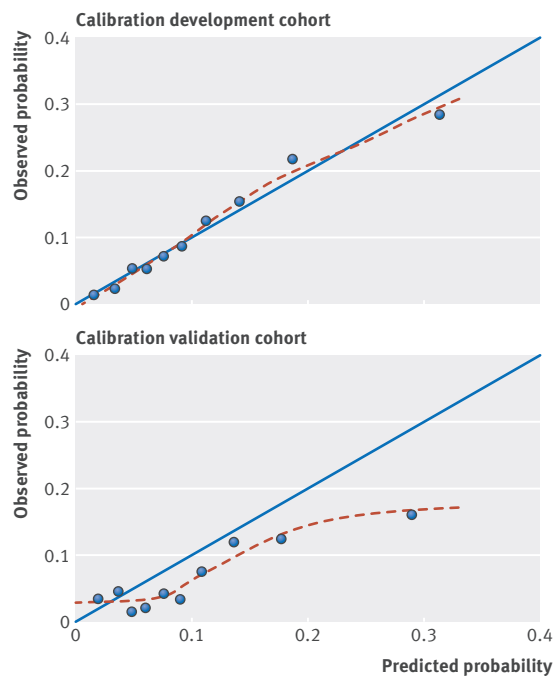
Discussion

We found that between 7% and 11% of patients undergoing orthopaedic surgery will experience acute kidney injury. In this study we developed a preoperative risk prediction model identifying seven predictors for postoperative acute kidney injury in patients undergoing orthopaedic surgery and validated the model in an external cohort of patients.

The predictors in the final model were age at operation, male sex, diabetes, lower estimated glomerular filtration rate, use of ACE inhibitors or angiotensin receptor blockers, number of prescribed drugs and American Society of Anesthesiologists grade. The importance of identifying patients at high risk, thereby potentially intervening to avoid acute kidney injury, was further emphasised by the effect of acute kidney injury on survival. We found that both short term and long

Table 2 | Predictive variables for acute kidney injury on multivariable logistic regression analysis of development cohort

Predictors	β	Odds ratio (95% CI)	P value
Sex (female)	−0.708	0.493 (0.407 to 0.596)	<0.001
Age at operation (years)	0.022	1.022 (1.013 to 1.031)	<0.001
Diabetes	0.427	1.532 (1.199 to 1.959)	0.001
No of prescribed drugs:			
None (reference)	—	—	0.94
1 or 2	0.130	1.139 (0.889 to 1.463)	0.308
≥3	0.347	1.415 (1.025 to 1.953)	0.035
ACE inhibitor or angiotensin receptor blocker	0.534	1.705 (1.343 to 2.164)	<0.001
Chronic kidney disease category (eGFR mL/min):			
>60	−1.417	0.242 (0.168 to 0.351)	<0.001
45-59	−1.108	0.330 (0.226 to.0483)	<0.001
30-44	−0.676	0.508 (0.434 to 0.753)	0.001
<29 (reference)	—	—	<0.001
ASA grade:			
1	−1.037	0.355 (0.194 to 0.647)	0.001
2	−0.141	0.869 (0.633 to 1.192)	0.384
3	0.014	1.014 (0.751 to 1.368)	0.929
4 (reference)	—	—	0.002

ACE=angiotensin converting enzyme; ASA=American Society of Anesthesiologists.
Intercept of model was −2.385.

Fig 2 | Calibration in validation cohort after recalibration using method 1 and method 2

Fig 1 | Calibration in development and validation cohorts

term survival was poorer in patients with acute kidney injury (defined as Kidney Disease Improving Global Outcomes stage 1 or more) after adjustment for age, sex, diabetes, baseline renal function, and number of prescribed drugs. In addition, survival was significantly poorer in patients with mild acute kidney injury (stage 1) compared with no acute kidney injury, and that it worsened with increased severity of acute kidney injury.

Comparison with other studies

Few risk scores exist for patients undergoing non-cardiac surgery. One study found blood transfusions, hepaticojejunostomy, and oliguria to be the strongest predictors for acute kidney injury after liver resection.[35] Furthermore, several risk scores are available for cardiac surgery.[36-38] Predictors in these risk scores are, however, generally specific to their particular specialties, rendering them inapplicable to an orthopaedic population. There is only one risk score in general surgical patients, developed by Kherterpal and colleagues from the United States.[7] These authors examined postoperative acute kidney injury in 75 952 patients undergoing general surgery. The independent risk factors for postoperative acute kidney injury they found were older age, male sex, emergency surgery, intraperitoneal surgery, diabetes mellitus necessitating oral treatment, diabetes mellitus necessitating insulin, active congestive heart failure, ascites, hypertension, mild preoperative renal insufficiency, and moderate preoperative renal insufficiency. Our study differs from that study in several ways. Firstly, the authors defined acute kidney injury as an increase in serum creatinine concentration of 176.8 µmol/L or more from the preoperative value over the first 30 postoperative days. Using those criteria, they found that acute kidney injury complicated 1% of their operations. In our cohorts, the rate of acute kidney injury was higher (around 10%) among patients undergoing orthopaedic surgery because we defined acute kidney injury using Kidney Disease Improving Global Outcomes criteria, which includes an increase of greater than 26.4 µmol/L in serum creatinine concentration, therefore including milder forms of acute kidney injury. Applying the same definition as that of Kherterpal and colleagues to our cohort would have identified similar rates of acute kidney injury but we thought it was important to include milder forms of acute kidney injury because of a notable difference in survival, even

Table 3 | Performance of prediction model in development and validation cohort

Model performance	Apparent validation	Internal validation	External validation		
	Development cohort	Development cohort (with bootstrapping)	Validation cohort	Validation cohort (recalibrated: method 1)	Validation cohort (recalibrated: method 2)
Calibration-in-the-large (%)	10.8 v 10.8*	10.8 v 10.8*	10.4 v 6.6	6.8 v 6.6	6.6 v 6.6
Calibration slope	1.00*	0.95	0.79	0.79	1.00*
Discrimination: C statistic	0.74	0.73	0.70	0.70	0.70

*Perfect values by definition.

Table 4 | Incidence of acute kidney injury across different risk categories in development and validation cohorts

Predicted risk	Development cohort		Validation cohort	
	No of patients	Incidence of acute kidney injury (No)	No of patients	Incidence of acute kidney injury (No)
<0.05	1377	0.03 (36)	972	0.03 (32)
0.05-0.10	1795	0.07 (123)	1247	0.03 (36)
0.10-0.15	951	0.14 (129)	658	0.10 (64)
0.15-0.20	481	0.19 (90)	363	0.12 (42)
0.20-0.25	292	0.25 (72)	170	0.17 (29)
>0.25	398	0.31 (122)	225	0.16 (37)

Table 5 | Comparison of short term and long term survival in patients with or without acute kidney injury

Survival variables	No (%) of deaths	Hazard ratio (95% CI)	P value
Overall survival (crude)	3166 (30)	2.04 (1.84 to 2.26)	<0.001
Overall survival (adjusted*)		1.53 (1.38 to 1.70)	<0.001
90 day (crude)	604 (5.7)	3.24 (2.68 to 3.91)	<0.001
90 day (adjusted)		2.36 (1.94 to 2.87)	<0.001
1 year (crude)	543 (5.4)	1.89 (1.49 to 2.41)	<0.001
1 year (adjusted)		1.40 (1.10 to 1.79)	0.007
Long term (crude) (end of follow-up)	2019 (21.3)	1.67 (1.45 to 1.94)	<0.001
Long term (adjusted) (end of follow-up)		1.23 (1.10 to 1.45)	0.001

*Adjusted for age, sex, diabetes, baseline renal function, and number of prescribed drugs.

Fig 3 | Kaplan Meier plot of overall survival in patients with acute kidney injury compared with no acute kidney injury

Table 6 | Overall adjusted survival according to severity of acute kidney injury

Stage*	No of patients	No (%) of deaths	Hazard ratio (95% CI)	P value
1 (mild)	826	362 (22.8)	1.46 (1.30 to 1.63)	<0.001
2	100	38 (38)	2.08 (1.51 to 2.87)	<0.001
3 (severe)	41	18 (43.9)	2.77 (1.74 to 4.41)	<0.001

*According to Kidney Disease Improving Global Outcomes criteria.

finally, risks and predictors may differ between American and European populations and healthcare systems.

Several studies have examined the relation between postoperative acute kidney injury and mortality. One study found that survival in a cohort of 10 518 patients undergoing major surgery who were admitted to a surgical intensive care unit was worse in patients with any acute kidney injury (as defined by Risk, Injury, Failure, Loss, End-stage Kidney (RIFLE) criteria) during their hospital stay, proportional to its severity (adjusted hazard ratio 1.18 for RIFLE risk category, 1.57 for RIFLE failure category). Acute kidney injury remained a risk factor for mortality even after adjustment for age, sex, race, type of surgery, comorbidities, other postoperative complications, discharge facility, and length of stay.[39] This study was limited to patients admitted to the surgical intensive care unit, therefore focusing on sicker patients with an expected poorer long term survival, whereas we included all patients undergoing orthopaedic surgery. More recently, in a cohort of more than 50 000 patients undergoing major surgery in the United States, the same authors found that both in-hospital and 90 day mortality were higher in patients with any degree of acute kidney injury, adjusted for age, sex, ethnicity, primary payer, Charlson comorbidity index score, surgery type, emergent surgery status, weekend admission, estimated glomerular filtration rate, and all postoperative complications.[40] Our results from a UK population are comparable to this US population. In addition, we also examined longer term follow-up, with a median follow-up of 4.58 years (interquartile range 2.93-6.69 years). Even though these observational studies adjusted for multiple baseline factors, it is difficult to determine if the excess mortality was due to the development of acute kidney injury; or whether sicker patients who have a higher mortality risk at baseline develop acute kidney injury.

Strengths and limitations of this study
This study has several limitations. In the absence of an agreed definition for baseline renal function, we used creatinine level at hospital admission, which is the most

with stage 1 severity. We were unable to assess how the authors' model performed in our population, as we did not have access to several variables used in the model. In addition, to increase the likelihood that the cause of acute kidney injury was associated specifically with surgery, we only included acute kidney injury occurring within the first seven postoperative days. Furthermore, we validated our prediction model externally and,

readily available in clinical practice. This could lead to underestimation of incidence of acute kidney injury.[41] We used routinely collected datasets in our model, therefore we were not able to include intraoperative factors such as hypotension or blood loss, which may improve the predictive performance of the model. However, this allowed us to develop a risk score that identifies patients who are at high risk, preoperatively. An advantage of using routinely collected data is that with the advent of electronic patients' records, this information can be used easily to flag patients at high risk to individual clinicians. We were unable to use direct data on comorbidity because of problems with under-reporting.[42] We therefore used number of prescribed drugs before admission, which has been shown to correlate well with comorbidity.[43] Our risk score was constructed and validated in patients undergoing orthopaedic surgery and so may not be applicable to other types of surgery. However, the increasing older population, higher expectations of patients, and improved technology have led to increasing rates of both elective and emergency orthopaedic operations, resulting in orthopaedic and trauma surgery being one of the largest surgical specialties, accounting for the greatest number of elective and unplanned surgery in the UK and other developed countries. Further work is required to validate this score in other surgical populations. The incidence of acute kidney injury was significantly lower in the validation cohort. As a result, calibration of our initial model was poor owing to over-prediction of acute kidney injury. The cause of this difference is not clear as it cannot be explained by available patient factors, although importantly we did not have data on smoking status. Observed differences in incidence may relate to factors pertaining to practice within the hospital. Recalibration to the incidence of acute kidney injury within a particular hospital can be a useful means of improving a model's predictive performance.[24 27 44 45] We have shown both a simple means of recalibration and a more complex method that is less practical to implement. In addition, our risk score utilised databases from the UK, and so before this score can be applied to other populations, validation is required using databases from other geographical areas. Other types of databases could affect the performance of the model owing to availability and reliability of recorded variables. The predictors included in our model are likely to be well documented in various other types of databases, especially in the (orthopaedic) surgery setting for which the model was developed and intended.

Other strengths of our study include the use of the Kidney Disease Improving Global Outcomes definition of acute kidney injury, allowing the inclusion of milder forms. Using a creatinine based definition may have led to the underestimation of acute kidney injury compared with using urine output.[46] Urine output, however, is poorly measured and recorded in most hospital settings other than intensive care units. We only included patients who developed acute kidney injury within the first seven postoperative days, thereby minimising the risk of including other causes of acute kidney injury. In contrast with many studies that examine in-hospital, 30

day, or 90 day mortality, our median follow-up was 4.58 years. Our study is also the first to use external validation of a risk score in this population. A UK report by the National Confidential Enquiry into Patient Outcome and Death (NCEPOD) found that a fifth of acute kidney injury cases were preventable and 50% were related to deficiencies in care.[44] A simple validated risk score to identify those patients undergoing non-cardiac surgery who are at high risk is currently lacking. Our risk score is simple to implement using routinely collected data. If the incidence of acute kidney injury is known to be substantially different from that in our development cohort, recalibrating the model to the new setting using the correction factor method[29] is advisable to help avoid systematic under-prediction or over-prediction. If the necessary data are available, recalibrating the model could be considered using the recalibration model method,[26] although in our external validation the gain in predictive performance was marginal. To illustrate the predictions of the model, consider a 70 year old man with diabetes and a baseline estimated glomerular filtration rate of 45 mL/min who takes three prescribed drugs in total (including ramipril) and is assigned an American Society of Anesthesiologists grade 2 (mild systemic disease). The model predicted a postoperative risk of 31% for acute kidney injury. The risk of acute postoperative kidney disease in the same patient 10 years older and with an estimated glomerular filtration rate of only 35 mL/min would be 47%. Appendix 2 provides details of how to calculate this risk. Appendix 3 shows our risk score translated into a Microsoft Office Excel spreadsheet, which could be further developed into a smartphone application. This would be an easy way for doctors to identify patients at high risk during the preoperative assessment stage and to guide perioperative management. Interventions could include increased frequency of monitoring renal function, avoidance of nephrotoxins in the perioperative period, and optimisation of fluid balance in high risk patients.[8] In addition, the renal team could be alerted before surgery to patients at very high risk. Also, this risk score could be used to target patients who would benefit the most from implementation of an intervention in the clinical trials setting.

Conclusions and policy implications

We have shown that acute kidney injury affects up to 11% of patients undergoing orthopaedic surgery, with increased longer term mortality even associated with milder forms of acute kidney injury. We developed and externally validated a preoperative risk prediction model, identifying seven predictors for acute kidney injury in patients undergoing orthopaedic surgery. These predictors are translated into a simple Excel spreadsheet (see appendix 3). Early identification and prevention of acute kidney injury has the potential to lead to improved patient outcomes worldwide. A simple risk score could be a low cost method of facilitating this, with important long term health and financial benefits.

1   Susantitaphong P, Cruz DN, Cerda J, et al. World incidence of AKI: a meta-analysis. *Clin J Am Soc Nephrol* 2013;8:1482-93.
2   Murugan R, Kellum JA. Acute kidney injury: what's the prognosis? Nature reviews. *Nephrology* 2011;7:209-17.
3   Bucaloiu ID, Kirchner HL, Norfolk ER, Hartle JE, 2nd, Perkins RM. Increased risk of death and de novo chronic kidney disease following reversible acute kidney injury. *Kidney Int* 2012;81:477-85.
4   Uchino S, Bellomo R, Bagshaw SM, Goldsmith D. Transient azotaemia is associated with a high risk of death in hospitalized patients. *Nephrol Dial Transplant* 2010;25:1833-9.
5   Coca SG, Singanamala S, Parikh CR. Chronic kidney disease after acute kidney injury: a systematic review and meta-analysis. *Kidney Int* 2012;81:442-8.
6   Kerr M, Bedford M, Matthews B, O'Donoghue D. The economic impact of acute kidney injury in England. *Nephrol Dial Transplant* 2014;29:1362-8.
7   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;110:505-15.
8   Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group. KDIGO Clinical Practice Guideline for Acute Kidney Injury. *Kidney Int Suppl* 2012;2:1-138.
9   Bell S, Davey P, Nathwani D, et al. Risk of AKI with gentamicin as surgical prophylaxis. *J Am Soc Nephrol* 2014;25:2625-32.
10   Calvert S, Shaw A. Perioperative acute kidney injury. *Periop Med* 2012;1:6.
11   Thakar CV, Kharat V, Blanck S, Leonard AC. Acute kidney injury after gastric bypass surgery. *Clin J Am Soc Nephrol* 2007;2:426-30.
12   Thakar CV, Liangos O, Yared JP, et al. ARF after open-heart surgery: influence of gender and race. *Am J Kidney Dis* 2003;41:742-51.
13   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;107:892-902.
14   National Institute for Health and Care Excellence. Acute kidney injury: prevention, detection and management of acute kidney injury up to the point of renal replacement therapy (Clinical guideline CG169). 2013. www.nice.org.uk/guidance/CG169.
15   University of Dundee. Health Informatics Centre. http://medicine. dundee.ac.uk/health-informatics-centre.
16   Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med* 2015;162:55-63.
17   Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 2015;350:g7594.
18   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:W1-73.
19   Efron B, Tibshirani RJ. An introduction to the bootstrap. Chapman & Hall/CRC, 1994.
20   Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009;150:604-12.
21   Dripps RD. New classification of physical status. *Anesthesiology* 1963;24:111.
22   Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* 2001;54:774-81.
23   Harrell FE, Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361-87.
24   Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010;21:128-38.
25   Austin PC, Steyerberg EW. Graphical assessment of internal and external calibration of logistic regression models by using loess smoothers. *Stat Med* 2014;33:517-35.
26   Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. Springer, 2009.
27   Toll DB, Janssen KJ, Vergouwe Y, Moons KG. Validation, updating and impact of clinical prediction rules: a review. *J Clin Epidemiol* 2008;61:1085-94.
28   Moons KG, Kengne AP, Grobbee DE, et al. Risk prediction models: II. External validation, model updating, and impact assessment. *Heart* 2012;98:691-8.
29   Janssen KJ, Vergouwe Y, Kalkman CJ, Grobbee DE, Moons KG. A simple method to adjust clinical prediction models to local circumstances. *Can J Anaesth* 2009;56:194-201.
30   De Goeij MC, van Diepen M, Jager KJ, Tripepi G, Zoccali C, Dekker FW. Multiple imputation: dealing with missing data. *Nephrol Dial Transplant* 2013;28:2415-20.
31   Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res* 1999;8:3-15.
32   Van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res* 2007;16:219-42.
33   Dekker FW, de Mutsert R, van Dijk PC, Zoccali C, Jager KJ. Survival analysis: time-dependent effects and time-varying risk factors. *Kidney Int* 2008;74:994-7.
34   Package 'rms' [program]. 4.3-0 version, 2015.
35   Slankamenac K, Beck-Schimmer B, Breitenstein S, Puhan MA, Clavien PA. Novel prediction score including pre- and intraoperative parameters best predicts acute kidney injury after liver surgery. *World J Surg* 2013;37:2618-28.
36   Candela-Toha A, Elías-Martín E, Abraira V, et al. Predicting acute renal failure after cardiac surgery: external validation of two new clinical scores. *Clin J Am Soc Nephrol* 2008;3:1260-5.
37   Thakar CV, Arrigain S, Worley S, Yared JP, Paganini EP. A clinical score to predict acute renal failure after cardiac surgery. *J Am Soc Nephrol* 2005;16:162-8.
38   Wijeysundera DN, Karkouti K, Dupuis JY, et al. Derivation and validation of a simplified predictive index for renal replacement therapy after cardiac surgery. *JAMA* 2007;297:1801-9.
39   Bihorac A, Yavas S, Subbiah S, et al. Long-term risk of mortality and acute kidney injury during hospitalization after major surgery. *Ann Surg* 2009;249:851-8.
40   Hobson C, Ozrazgat-Baslanti T, Kuxhausen A, et al. Cost and mortality associated with postoperative acute kidney injury. *Ann Surg* 2015;261:1207-14.
41   Siew ED, Matheny ME, Ikizler TA, et al. Commonly used surrogates for baseline renal function affect the classification and prognosis of acute kidney injury. *Kidney Int* 2010;77:536-42.
42   Quan H, Parsons GA, Ghali WA. Validity of information on comorbidity derived rom ICD-9-CCM administrative data. *Med Care* 2002;40:675-85.
43   Perkins AJ, Kroenke K, Unutzer J, et al. Common comorbidity scales were similar in their ability to predict health care costs and mortality. *J Clin Epidemiol* 2004;57:1040-8.
44   Ivanov J, Tu JV, Naylor CD. Ready-made, recalibrated, or remodeled? Issues in the use of risk indexes for assessing mortality after coronary artery bypass graft surgery. *Circulation* 1999;99:2098-104.
45   Krumholz HM. Mathematical models and the assessment of performance in cardiology. *Circulation* 1999;99:2067-9.
46   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;35:1692-702.

Appendices 1-3: operation procedure codes included in study; example of using model to calculate risk of acute kidney injury; Excel spreadsheet showing risk calculator