Derivation and Validation of a Simplified Predictive Index for Renal Replacement Therapy After Cardiac Surgery

Duminda N. Wijeysundera, MD
Keyvan Karkouti, MD, MSc
Jean-Yves Dupuis, MD
Vivek Rao, MD, PhD
Christopher T. Chan, MD
John T. Granton, MD
W. Scott Beattie, MD, PhD

Cardiac surgery remains a common surgical procedure, with approximately 646,000 open-heart operations performed annually in the United States alone.1 One of its most serious complications is acute renal failure. When severe enough to necessitate renal replacement therapy (RRT) with hemodialysis or continuous venovenous hemodiafiltration, postoperative acute renal failure is independently associated with mortality.2 Most clinicians believe that rates of RRT are likely to increase because patients undergoing cardiac surgery are increasingly older and have more comorbidity.3 Strategies that improve management of perioperative acute renal failure should, therefore, confer important benefits.

A simple risk index that uses readily available preoperative data to accurately predict RRT will improve clinical management and research design in this area of perioperative medicine. First, preoperative estimates of risk might help tailor use of existing, albeit limited, renal-protective strategies. Second, patients would receive accurate estimates of risk during pre-

Context A predictive index for renal replacement therapy (RRT; hemodialysis or continuous venovenous hemodiafiltration) after cardiac surgery may improve clinical decision making and research design.

Objectives To develop a predictive index for RRT using preoperative information.


Main Outcome Measure Postoperative RRT.

Results RRT rates in the derivation, Toronto validation, and Ottawa validation cohorts were 1.3%, 1.8%, and 2.2%, respectively. Multivariable predictors of RRT were preoperative estimated glomerular filtration rate, diabetes mellitus requiring medication, left ventricular ejection fraction, previous cardiac surgery, procedure, urgency of surgery, and preoperative intra-aortic balloon pump. The predictive index was scored from 0 to 8 points. An estimated glomerular filtration rate less than or equal to 30 mL/min was assigned 2 points; other components were assigned 1 point each: estimated glomerular filtration rate 31 to 60 mL/min, diabetes mellitus, ejection fraction less than or equal to 40%, previous cardiac surgery, procedure other than coronary artery bypass grafting, intra-aortic balloon pump, and nonelective case. Among the 53% of patients with low risk scores (≤1), the risk of RRT was 0.4%; by comparison, this risk was 10% among the 6% of patients with high-risk scores (≥4). The predictive index had areas under the receiver operating characteristic curve in the derivation, Toronto validation, and Ottawa validation cohorts of 0.81, 0.78, and 0.78, respectively. When these cohorts were stratified based on index scores, likelihood ratios for RRT were more concordant than observed RRT rates.

Conclusions RRT after cardiac surgery is predicted by readily available preoperative information. A simple predictive index based on this information discriminated well between low- and high-risk patients in derivation and validation cohorts. The index had improved generalizability when used to predict likelihood ratios for RRT.

JAMA. 2007;297:1801-1809

www.jama.com
operative counseling. Third, identification of high-risk patients could guide intensive care units in allocating resources for postoperative care. Although they comprise less than 2% of patients undergoing cardiac surgery, patients needing postoperative RRT use 12% of intensive care unit resources, as measured by length of stay.4 Finally, researchers might use this index to select intermediate- to high-risk participants for clinical trials of nephroprotective therapies, and to facilitate risk adjustment with multivariable regression or propensity score techniques.

To achieve widespread use, a predictive index for RRT after cardiac surgery should have specific characteristics. It should provide accurate prognostic information based on readily available, clinically sensible preoperative data. This prognostic information should be generalizable; specifically, predictive accuracy should remain stable when the index is externally validated in different data sets, geographic locales, and time periods.5 Where possible, the index should also be simple and easily scored. Predictive indices for perioperative cardiovascular risk demonstrate the feasibility of improving ease of use while preserving accuracy. Compared with older indices, the Revised Cardiac Risk Index provides comparably accurate prognostic information with a reduced number of equally weighted component variables.6

There have been few previous studies of predictive indices for postoperative RRT, all of which had important limitations.2,4,7 Aside from a clinical prediction rule by Chertow et al,2 these previous indices had relatively complex scoring schemes and underwent only limited external validation. The Chertow et al prediction rule, which was derived in a predominantly male (99%) cohort, demonstrated only fair discrimination when externally validated in a more diverse population.5 In view of these limitations, we undertook a retrospective cohort study to derive and validate a simple, accurate predictive index for RRT after cardiac surgery at 2 tertiary care hospitals in Ontario, Canada.

**METHODS**

**Data Sources**

After institutional research ethics approval, preoperative, intraoperative, and postoperative data on individuals undergoing cardiac surgery at the Toronto General Hospital (Toronto, Ontario) and University of Ottawa Heart Institute (Ottawa, Ontario) were prospectively collected in distinct hospital-specific registries. These databases have been previously described.9,10 Medical personnel (anesthesiologists, surgeons, perfusionists) collected preoperative and intraoperative data. Research personnel adjudicated all outcomes from medical records. Database accuracy at the Toronto General Hospital exceeded 95% when compared with 200 randomly selected medical records that had been reabstracted.11 Based on reabstraction of 175 randomly selected medical records, database accuracy at the Ottawa Heart Institute was 98%.9

**Patient Samples**

The study sample consisted of consecutive adults (≥18 years) who underwent cardiac surgery under cardiopulmonary bypass. Exclusion criteria included severe preoperative renal dysfunction (preoperative RRT dependence or creatinine concentration >3.4 mg/dL [300 µmol/L]) and infrequent procedures (eg, heart transplantation, ventricular assist device insertion). The derivation cohort consisted of eligible cases at the Toronto General Hospital between May 1999 and July 2004. This cohort was estimated at 10,000 patients, thereby permitting unbiased fitting of 10 or fewer predictor variables in multiple logistic regression (estimated 1% rate of postoperative RRT).12 We specified 2 distinct validation cohorts. One validation cohort consisted of eligible cases at the Toronto General Hospital between August 2004 and December 2005; this phase of external validation evaluated the generalizability of prognostic information when applied to a different time period.3 The other validation cohort consisted of eligible cases at the Ottawa Heart Institute between January 1999 and December 2003; this validation process evaluated generalizability based on a differing geographic locale and data collection method.3

**Variable Definitions**

Potential predictor variables were selected based on clinical sensibility, literature review, and availability before surgery. These variables included demographic characteristics (age, sex, weight), preoperative renal reserve, comorbid disease (left ventricular ejection fraction, hypertension, cerebrovascular disease, peripheral vascular disease, diabetes mellitus requiring medication, chronic obstructive pulmonary disease), operative characteristics (procedure type, previous cardiac surgery, preoperative intra-aortic balloon pump use), and urgency of surgery.7,11,13-16

Given that creatinine concentration has important limitations as a measure of preoperative renal function, we measured renal reserve using estimated glomerular filtration rate (GFR), based on the Cockcroft-Gault equation.11,17-20 New York Heart Association functional class was not included among potential predictor variables because its poor interrater reliability might adversely affect the reproducibility of the prediction index.21,22 Cardiopulmonary bypass time was not considered a potential predictor variable. Cardiopulmonary bypass time is associated with RRT, but it is not readily estimated before surgery.7,14,15 We calculated estimated GFR using serum creatinine concentrations and weights that are routinely measured within 30 days before surgery. If multiple measurements were present, the value closest to the date of surgery was used. All other preoperative variables were measured on the date of surgery.

The primary outcome was postoperative RRT (intermittent hemodialysis or continuous venovenous hemo-
All potential predictor variables were initially fitted into a multivariable logistic regression model that predicted postoperative RRT. This full model was subsequently reduced to include only variables with clinically sensible directions for their predictive effects. We assessed the discrimination of the reduced model using the area under the receiver operating characteristic (ROC) curve, and its calibration using an observed vs predicted plot and the Hosmer-Lemeshow goodness-of-fit statistic. The regression coefficients of variables in the reduced model were then used to develop component scores for a simplified renal index to predict RRT. Where possible, we tried to assign equal weights to each included component variable. The discrimination of the simplified renal index was described using the area under the ROC curve, while its calibration was assessed using an observed vs predicted plot and the Hosmer-Lemeshow statistic. We internally validated the index by computing bias-corrected estimates of discrimination and calibration by bootstrap resampling (n=200).

The simplified renal index was externally validated in the Ottawa and Toronto validation cohorts. Index discrimination was measured in each vali-
The derivation cohort consisted of 10 751 patients, while the Toronto validation and Ottawa validation cohorts consisted of 2566 and 6814 patients, respectively (N=20 131). The rates of RRT in the derivation cohort were 1.3% (n=139), while rates in the Toronto validation and Ottawa validation cohorts were 1.8% (n=45) and 2.2% (n=152), respectively.

### Bivariate and Multivariable Analyses

The perioperative characteristics of the derivation cohort are presented in Table 2. The bivariate associations of these potential predictor variables with postoperative RRT are also presented in Table 2. These bivariate associations guided further categorization of age (≤60 years, 61-80 years, >80 years); estimated GFR (>60 mL/min, 31-60 mL/min, ≤30 mL/min); weight (≤111 lb [<50 kg], 112-178 lb [51-80 kg], >178 lb [≥80 kg]); left ventricular ejection fraction (>40%, ≥40%); operative procedure (coronary artery bypass graft or atrial septal defect repair, other); and urgency of procedure (elective, nonelective).

While reducing the full multivariable regression model, we removed certain potential predictor variables (eg, age, weight) that were statistically significant in bivariate analyses. These eliminated variables had neither clinically sensible directions for their predictive effects nor statistically significant associations with RRT in the fitted regression model. The final multivari-
able model for predicting postoperative RRT included 8 predictor variables (Table 2). This model had good discrimination (area under ROC curve 0.83) and calibration (Hosmer-Lemeshow statistic $P = .48$).

**Predictive Index Derivation**

A simplified renal index (Table 3) was then developed using the regression coefficients of the multivariable model (Table 2). This simple predictive index contained 8 components, each of which was assigned 1 point, with the exception of estimated GFR of 30 mL/min or less (2 points). The proportions of the derivation cohort with scores of 0, 1, 2, 3, 4, and 5 or greater were 16% ($n = 1714$), 38% ($n = 4097$), 28% ($n = 3035$), 13% ($n = 1343$), 4% ($n = 449$), and 1% ($n = 113$), respectively.

The simplified renal index had excellent discrimination for postoperative RRT with an area under the ROC curve of 0.81 (bias-corrected 0.81; 95% confidence interval [CI], 0.78-0.84). The predictive index also had good calibration, based on its observed vs predicted plot and Hosmer-Lemeshow statistic ($P = .27$).

**External Validation**

The perioperative characteristics of the derivation, Toronto validation, and Ottawa validation cohorts are presented in Table 4. The cohorts differed with respect to several characteristics including estimated GFR, chronic obstructive pulmonary disease, cerebrovascular disease, left ventricular ejection fraction, operative procedure, and cardiopulmonary bypass time (Table 4).

The simplified renal index retained good discrimination in the Toronto and Ottawa validation cohorts. The area under the ROC curve in the Toronto cohort was 0.78 (95% CI, 0.72-0.84) and the area in the Ottawa cohort was 0.78 (95% CI, 0.74-0.81).

Nonetheless, when the simplified renal index score was used to estimate RRT rates, calibration varied among the cohorts (Figure 1). Specifically, patients' scores on the simplified renal index could discriminate between high- and low-risk individuals within the same cohort; however, when patients with the same score were compared between cohorts, observed RRT rates varied (Figure 1). Especially in risk strata defined by index scores of 2 or 3, the observed RRT rates were qualitatively different in the derivation and validation cohorts (Figure 2).

In contrast, the agreement between derivation and validation cohorts for likelihood ratios of postoperative RRT was better (Figure 2). These ratios described the probability of RRT relative to the mean risk in the cohort. Ratios greater than 1 represent increases in risk compared with the mean event rate in the specific cohort; values less than 1 represent decreases in risk compared with the mean event rate. The overlap in 95% CIs for likelihood ratios was consistent across all risk strata.$^{10}$

---

**Table 4. Perioperative Characteristics of the Derivation and Validation Cohorts**

<table>
<thead>
<tr>
<th></th>
<th>Derivation Cohort</th>
<th>Validation Cohorts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toronto ($n = 10751$)†</td>
<td></td>
<td>Ottawa ($n = 6814$)‡</td>
</tr>
<tr>
<td>Weight, mean (SD), lb</td>
<td>176 (38)</td>
<td>180 (36)</td>
</tr>
<tr>
<td>Women</td>
<td>2787 (26)</td>
<td>1891 (28)</td>
</tr>
<tr>
<td>Age, mean (SD), y</td>
<td>62 (12)</td>
<td>64 (12)</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peripheral vascular disease</td>
<td>1421 (13)</td>
<td>1119 (16)</td>
</tr>
<tr>
<td>Left ventricular ejection fraction, %</td>
<td>4864 (45)</td>
<td>3713 (56)</td>
</tr>
<tr>
<td>Surgical details</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative intra-aortic balloon pump</td>
<td>217 (2)</td>
<td>89 (1)</td>
</tr>
<tr>
<td>Preoperative intra-aortic balloon pump</td>
<td>7005 (65)</td>
<td>4752 (70)</td>
</tr>
<tr>
<td>Operative procedure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valve surgery alone</td>
<td>1755 (16)</td>
<td>799 (12)</td>
</tr>
<tr>
<td>Valve surgery alone</td>
<td>1755 (16)</td>
<td>799 (12)</td>
</tr>
<tr>
<td>Complex procedures</td>
<td>1991 (19)</td>
<td>263 (18)</td>
</tr>
<tr>
<td>Timing of surgery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elective</td>
<td>6392 (59)</td>
<td>3985 (58)</td>
</tr>
<tr>
<td>Emergency</td>
<td>4158 (39)</td>
<td>2499 (37)</td>
</tr>
</tbody>
</table>

©2007 American Medical Association. All rights reserved.

(Reprinted) JAMA, April 25, 2007—Vol 297, No. 16 1805
Comparison With Alternative Predictive Indices

The simplified renal index, its variant with exact weighting, the Thakar et al risk score, and the Mehta et al nomogram tool had qualitatively similar discrimination and widely overlapping 95% CIs (Table 5). In contrast, the Chertow et al prediction rule displayed only fair discrimination with an area under the ROC curve of 0.68 to 0.70 (Table 5).

COMMENT

Our study found that a predictive index with a simplified scoring system could use readily available, clinically sensible preoperative information to provide accurate prognostic information on RRT after cardiac surgery. The simplified renal index compares favorably with other perioperative risk indices with regard to discriminating between high- and low-risk patients.6,33,34

Comparison With Previous Indices

The simplified renal index warrants comparison against the few previous predictive indices for RRT after cardiac surgery. Compared with the more complex Thakar et al and Mehta et al predictive indices, our index demonstrated improved ease of use, because of a fewer number of component variables and simpler scoring system, with similar prognostic accuracy.4,7 In this respect, the simplified renal index mirrors the evolution of preoperative cardiac risk indices, where the newer Revised Cardiac Risk Index has comparable accuracy, but improved usability, compared with the older Detsky and Goldman indices.6,35,36

Potential explanations for the somewhat inferior performance of the Chertow et al prediction rule include its derivation cohort and its underlying statistical methods. As previously described, this prediction rule was derived in a predominantly male (99%) cohort.13 Its discrimination may, therefore, have deteriorated when externally validated in more diverse populations, as occurred in both a previous study by Fortescue et al (ROC area 0.72) and our present study (ROC area 0.68-0.70).8 In addition, Chertow et al derived their prediction rule using recursive partitioning. The latter statistical technique may have inferior discrimination compared with logistic regression, which was used to derive the other predictive indices.37

Strengths

The principal strength of the simplified renal index is its combination of simplicity and accuracy. It used only 8
clinically sensible components, all of which were equally weighted, to accurately predict a clinically relevant outcome, namely RRT. This simplicity will help promote its use in clinical care and research settings. It also differs from recent comparable indices by measuring preoperative renal function using estimated GFR. Perioperative studies have shown GFR to be superior to serum creatinine as a predictor of clinical outcomes, including mortality and postoperative renal function. Furthermore, guidelines recommend estimated GFR, as opposed to serum creatinine concentration, as the best overall estimate of kidney function. Despite its relative simplicity, the simplified renal index demonstrated good discrimination in derivation and validation cohorts. Furthermore, our external validation process evaluated performance in varying time periods, geographic locales, and data sets. It is important to emphasize that the Ottawa Heart Institute and Toronto General Hospital maintain distinct databases and data collection methods. The simplified renal index, therefore, likely has generalized discrimination in varied settings.

Our external validation process also demonstrated that calibration of the simplified risk index varied when used to predict RRT rates in different cohorts. There are several possible explanations for this observation. First, postoperative RRT is associated with intra- and post-operative factors that are not included in a preoperative risk index: cardiopulmonary bypass time, intraoperative hematocrit, low-cardiac output syndrome, postoperative glycemic control, available RRT resources, and thresholds for initiating RRT. Previous studies have demonstrated this combination of preserved discrimination but diminished calibration, with predictive indices used in similar contexts, namely other preoperative risk indices and critical care admission risk indices. Second, inter-laboratory variation in serum creatinine assays might have affected both the prevalence of diminished estimated GFR and its association with RRT. Finally, the calibration of the simplified renal index may have been influenced by unmeasured preoperative variables.

Interpretation of a predictive index using likelihood ratios may correct, in part, for influential factors that occur after index measurement. Likelihood ratios measure the probability of RRT relative to the average risk in the cohort. Our results suggest that likelihood ratios, by accounting for mean risk within a cohort, help preserve the calibration of a preoperative risk index in external validation.

### Implications

As a simple and accurate predictive index, the simplified renal index will improve preoperative risk stratification for RRT. Our index offers several options for risk stratification when prospectively applied in a new setting. One might simply extrapolate RRT rates observed in our derivation sample to the new cohort. In doing so, the simplified renal index will likely discriminate well between low- and high-risk patients; however, observed RRT rates may vary from predicted rates, especially if historical RRT rates in the new setting differ greatly from our derivation sample. This variation between observed and predicted RRT rates can be corrected, in part, by recalibrating the simplified renal index. This process of statistical recalibration entails both a preexisting large data set of cardiac surgery patients at the new setting and adequate statistical support. In the absence of these resources, clinicians and researchers might instead use our index to estimate likelihood ratios for RRT, which appear to be relatively stable when extrapolated to a new setting. If the historical RRT rate in the new setting is known, these likelihood ratios can, in turn, be used to calculate predicted RRT rates. Alternatively, clinicians and researchers can directly interpret the likelihood ratios stratify patients by risk. For example, patients with risk scores of 0 and 1 have risks for RRT that are only one tenth and one third, respectively, that for an average patient in the new setting. In contrast, patients with risk scores of 4 and 5 have risks of RRT that are 6-fold and 12-fold higher, respectively, than the average patient. Despite not providing absolute risks for RRT, this information is clinically useful. As has been observed with preoperative cardiac risk indices, clinicians often do not estimate absolute risks for adverse events; instead, they simply classify patients as low-, intermediate-, or high-risk. Given that likelihood ratios less than 0.2 represent at least moderate reductions in risk while values greater than 5 represent at least moderate increases in risk, patients might be conveniently classified into low-risk (≤1 point), intermediate-risk (2-3 points), and high-risk (≥4 points) categories based on their simplified renal index score.

Accurate preoperative risk stratification for RRT after cardiac surgery will, in turn, inform clinical decision making and research design. First, clini-
Cockcroft-Gault equation to be an acceptable method for determining estimates of risk during preoperative management. Intermediate-risk individuals might benefit from limited use of potential renal-protective interventions (eg, strict control of intraoperative hemocrit, off-pump coronary artery surgery, fenoldopam) and close postoperative surveillance with biomarkers of early renal injury (eg, urinary neutrophil gelatinase–associated lipocalin).10,44-46 High-risk individuals may benefit from more aggressive use of potential renal-protective interventions. Second, preoperative identification of high-risk patients will assist intensive care units in planning allocation of postoperative RRT resources. Third, patients would receive accurate estimates of risk during preoperative counseling. Finally, clinical trials might improve selection of intermediate- to high-risk participants by incorporating the simplified renal index into their inclusion criteria. Recent investigations of perioperative renal-protective therapies have specifically called for research to improve identification of patients at risk for renal events after cardiac surgery.47

Limitations

There are several limitations to be considered when interpreting our results. The association between estimated GFR and RRT observed in our present study should not be extrapolated to alternative prediction equations, such as the Modification of Diet in Renal Disease formula or cystatin-C-based equations.46 We did not apply the Modification of Diet in Renal Disease formula, which may have improved the accuracy of estimated GFR, because our prospective clinical registries did not capture required ethnicity data. Nonetheless, current guidelines consider the Cockcroft-Gault equation to be an acceptable method for determining estimated GFR in adults.18 Calculation of estimated GFR also introduces more complexity to the preoperative assessment process than measuring serum creatinine concentration alone. Nonetheless, the improved estimation of true renal reserve more than compensates for the slight additional complexity entailed by estimated GFR.18-20 Furthermore, estimated GFR calculation may be readily facilitated using existing computer or personal digital assistant software. In addition, our data were limited to in-hospital outcomes, therefore precluding any conclusions on long-term implications of postoperative RRT. Given that the initiation of RRT is based on clinical judgment, consulting nephrologists might have also modified their threshold for initiating in-hospital RRT based on patients’ preoperative characteristics. Finally, our data were derived from 2 tertiary-care hospitals in Ontario, Canada. Multisite external validation of the simplified renal index in other geographic regions remains needed to further characterize its generalizability.3

CONCLUSION

In summary, we found that a simple predictive index can use readily available preoperative information to accurately predict RRT after cardiac surgery. The calibration of this index across cohorts improved when used to estimate likelihood ratios for RRT as opposed to predicted event rates. This simple predictive index may facilitate preoperative risk stratification for RRT, and thereby improve clinical decision making, communication of perioperative risk, resource allocation, and research design.

Author Contributions: Dr Wijeysundera had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Wijeysundera, Karkouti, Rao, Chan. Acquisition of data: Wijeysundera, Karkouti, Dupuis, Beattie. Analysis and interpretation of data: Wijeysundera, Dupuis, Rao, Granton, Beattie. Drafting of the manuscript: Wijeysundera. Critical revision of the manuscript for important intellectual content: Wijeysundera, Karkouti, Dupuis, Rao, Chan, Granton, Beattie. Statistical analysis: Wijeysundera. Obtained funding: Dupuis, Chan. Administrative, technical, or material support: Wijeysundera, Karkouti, Dupuis, Chan, Beattie. Study supervision: Rao.


©2007 American Medical Association. All rights reserved.

(Reprinted) *JAMA*, April 25, 2007—Vol 297, No. 16 1809