



## Estimated glomerular filtration rate better predicts 30-day mortality after non-cardiac surgery than serum creatinine: a retrospective analysis of 92,888 patients

## Le taux de filtration glomérulaire estimé est supérieur à la créatinine sérique pour prédire la mortalité à 30 jours après chirurgie non cardiaque: une analyse rétrospective de 92 888 patients

Jacek B. Cywinski, MD · Edward J. Mascha, PhD · Andrea Kurz, MD · Daniel I. Sessler, MD

Received: 27 August 2014 / Accepted: 13 April 2015 / Published online: 29 April 2015  
© Canadian Anesthesiologists' Society 2015

### Abstract

**Background** Serum creatinine is the most commonly used indicator of renal function, but its derivative, estimated glomerular filtration rate (eGFR), has been shown to be superior in non-surgical settings. It remains unknown if eGFR better predicts postoperative mortality in non-cardiac surgical patients. We thus tested the hypothesis that eGFR predicts 30-day mortality after non-cardiac surgery better than serum creatinine.

**Methods** We evaluated patients who had inpatient non-cardiac surgery of at least one hour duration during January 2006 to December 2011 at the Cleveland Clinic Main Campus and whose preoperative serum creatinine was measured within 30 days before surgery. The eGFR

was calculated using the Chronic Kidney Disease-Epidemiology Collaboration equation. Preoperative eGFR was compared in a multivariable analysis with preoperative serum creatinine (both assessed as continuous variables) on the ability to predict 30-day mortality in all patients. Secondarily, the comparison was made within subgroups based on amount of blood loss, blood transfusion, and sex.

**Results** There were 92,888 patients included in the final analysis. The eGFR was a modestly better discriminator of 30-day mortality than serum creatinine, with an estimated *c*-statistic (95% confidence interval) of 0.67 (0.65 to 0.68) for eGFR vs 0.61 (0.59 to 0.63) for serum creatinine ( $P < 0.001$ ). Furthermore, the eGFR was consistently a better discriminator of 30-day mortality across blood loss, transfusion, and sex groups. Reclassification analyses suggested improved individual predictions of 30-day mortality using eGFR compared with serum creatinine. Nevertheless, a multivariable combination of baseline characteristics of American Society of Anesthesiologists physical status, age, and body mass index (all  $P < 0.001$ ) discriminated 30-day mortality with a *c*-statistic of 0.850. Adding eGFR to the model improved the *c*-statistic to only 0.851, while separately adding serum creatinine did not change the *c*-statistic.

**Conclusion** The eGFR is a modestly better predictor of 30-day mortality than serum creatinine in patients having inpatient non-cardiac surgery. Given that eGFR is often reported by clinical laboratories and is otherwise easy to calculate, it should generally be used in preference to creatinine alone.

This article is accompanied by an editorial. Please see Can J Anesth 2015; 62: this issue.

**Author contributions** Jacek B. Cywinski, Andrea Kurz, and Daniel I. Sessler helped design and conduct the study and write the manuscript. Edward J. Mascha helped design and conduct the study, analyze the data, and write the manuscript.

J. B. Cywinski, MD (✉)  
Department of General Anesthesiology, Cleveland Clinic, 9500  
Euclid Avenue/G31, Cleveland, OH, USA  
e-mail: cywinsj@ccf.org

E. J. Mascha, PhD  
Departments of Quantitative Health Sciences and Outcomes  
Research, Cleveland Clinic, Cleveland, OH, USA

E. J. Mascha, PhD · A. Kurz, MD · D. I. Sessler, MD  
Department of Outcomes Research, Cleveland Clinic, Cleveland,  
OH, USA

## Résumé

**Contexte** La créatininémie est l'indicateur de la fonction rénale le plus souvent utilisé, mais son dérivé, le taux de filtration glomérulaire estimé (EGFR) s'est avéré meilleur dans des contextes non chirurgicaux. Nous ne savons pas si l'EGFR prédit mieux la mortalité postopératoire de patients chirurgicaux non cardiaques. Nous avons donc testé l'hypothèse selon laquelle l'EGFR prédisait mieux la mortalité à 30 jours après chirurgie non cardiaque que la créatininémie.

**Méthodes** Nous avons évalué des patients hospitalisés qui avaient eu une chirurgie non cardiaque d'une durée d'au moins une heure entre janvier 2006 et décembre 2011 au Cleveland Clinic Main Campus et dont la créatinine sérique avait été dosée dans les 30 jours précédant la chirurgie. L'EGFR a été calculé au moyen de l'équation de la Chronic Kidney Disease-Epidemiology Collaboration. L'EGFR préopératoire a été comparé dans une analyse multifactorielle à la créatinine sérique préopératoire (tous deux évalués en tant que variables continues) quant à leur capacité de prédiction de la mortalité à 30 jours chez tous les patients. Une comparaison secondaire a été effectuée au sein de sous-groupes en fonction du volume des pertes sanguines, des transfusions sanguines et du sexe des patients.

**Résultats** L'analyse finale a inclus 92 888 patients. L'EGFR a permis une prédiction de la mortalité à 30 jours avec une efficacité modestement supérieure à celle de la créatinine sérique, avec une aire sous la courbe (statistique *c*) estimée (intervalle de confiance à 95 %) de 0,67 (0,65 à 0,68) pour l'EGFR contre 0,61 (0,59 à 0,63) pour la créatinine sérique ( $P < 0,001$ ). De plus, l'EGFR a été constamment un meilleur prédicteur de la mortalité à 30 jours dans les sous-groupes de pertes sanguines, transfusion et selon le sexe des patients. Les analyses de reclassement ont suggéré des prédictions individuelles améliorées de la mortalité à 30 jours avec l'utilisation de l'EGFR par rapport à la créatinine sérique. Néanmoins, une combinaison multifactorielle des caractéristiques à l'inclusion du statut physique, de l'âge et de l'indice de masse corporelle (tous:  $P < 0,001$ ) de l'American Society of Anesthesiologists a été un meilleur élément de prédiction de la mortalité à 30 jours avec une aire sous la courbe de 0,850. L'ajout de l'EGFR au modèle a amélioré l'aire sous la courbe à seulement 0,851, alors que l'ajout indépendant de la créatinine sérique ne l'a pas changée.

**Conclusion** L'EGFR est un outil prédictif modestement meilleur de la mortalité à 30 jours que la créatinine sérique chez des patients hospitalisés ayant subi une chirurgie non cardiaque. Considérant que l'EGFR est souvent fourni par les laboratoires d'analyses cliniques ou

qu'il est facile à calculer, son utilisation devrait être généralisée, de préférence à la créatinine seule.

Preoperative renal dysfunction is a well-established risk factor for postoperative complications.<sup>1-4</sup> Quantifying renal dysfunction is thus clinically important; consequently, preoperative serum creatinine is included in various risk stratification models, including the Revised Cardiac Risk Index and the Cleveland Clinic Risk Score.<sup>5-7</sup> A limitation is that serum creatinine concentration is an imprecise measure of renal function. It is an especially poor estimate of mild renal dysfunction; furthermore, comparable creatinine concentrations may represent different degrees of renal impairment in patients of various age, sex, and ethnicity. And finally, the relationship between serum creatinine and glomerular filtration rate (GRF) – a clinically more important measure – is complex and depends on various factors, including sex, ethnicity, recent dietary intake, drug use, muscle mass, and extra-renal creatinine clearance.<sup>8</sup>

Direct measurement of GFR using exogenous markers remains the gold standard but remains impractical in most clinical situations. In contrast, estimated GFR (eGFR) provides a reasonable compromise between accuracy and practicality.<sup>9,10</sup> The two most popular methods of calculating eGFR are the Modification of Diet in Renal Disease (MDRD) Study equation and the Chronic Kidney Disease-Epidemiology (CKD-EPI) Collaboration equation. Both have been validated against measured GFR and offer acceptably precise assessment of renal function.<sup>9-11</sup>

In both cardiac<sup>3,12-14</sup> and non-cardiac<sup>2,4</sup> surgery settings, preoperative renal dysfunction is strongly associated with adverse outcomes. Most studies were conducted on the basis of serum creatinine,<sup>5,6,15,16</sup> although some used eGFR.<sup>2-4,12-14</sup> For example, Shavit *et al.* showed that even relatively small preoperative increases in serum creatinine or decreases in eGFR are associated with significant risk of mortality and morbidity.<sup>3</sup> Specifically, mortality increased significantly with every  $15 \text{ mL} \cdot \text{min}^{-1} \cdot 1.73 \text{ m}^{-2}$  decrease in preoperative eGFR.

A meta-analysis of studies assessing the impact of preoperative renal dysfunction on postoperative outcomes confirms that both an elevation in serum creatinine and/or a decrease in eGFR are associated with worse outcomes.<sup>1</sup> Nevertheless, cardiac and vascular surgery patients were significantly overrepresented among the included studies. Cardiac and vascular surgery represents a high-risk intervention, with the majority of patients exposed to cardiopulmonary bypass, large fluid shifts, and a high risk for hemodynamic perturbations. On the other hand, the

general surgical population is exposed to a wide spectrum of procedures with invasiveness ranging from low to high. Accordingly, the association of preoperative renal dysfunction with postoperative outcomes may differ between cardiac and non-cardiac surgery patients. Hence, it remains unknown whether use of eGFR enhances clinicians' ability to predict adverse perioperative outcomes compared with serum creatinine alone in patients having non-cardiac surgery.

We therefore tested the hypothesis that preoperative eGFR, which was estimated using the CKD-EPI formula, predicts 30-day postoperative mortality better than preoperative creatinine in a large cohort of non-cardiac surgery patients.

## Methods

Data for this retrospective investigation were obtained from the Perioperative Health Documentation System (IRB #8167), and the need for consent was waived by the Cleveland Clinic Institutional Review Board (April 2013). We considered adults who had inpatient non-cardiac surgery lasting at least one hour during January 1, 2006 to December 31, 2011 at the Cleveland Clinic Main Campus. For patients who had multiple operations, only the first surgery was included in our investigation. Inclusion criteria comprised patients whose age, sex, race, and preoperative serum creatinine level within 30 days of the surgical procedure were available. Exclusion criteria were patients on dialysis preoperatively and patients with a creatinine level  $> 10 \text{ mg}\cdot\text{dL}^{-1}$  or an eGFR  $> 180$  or  $< 10 \text{ mL}\cdot\text{min}^{-1}$ .

Basic demographic data were retrieved, including: age, sex, race, height, weight, type of surgical procedure, medical history (coronary artery disease, hypertension, diabetes mellitus, cerebrovascular disease, peripheral vascular disease, chronic obstructive pulmonary disease), and surgical class of the procedure. We also retrieved preoperative hemoglobin concentrations along with the Charlson Comorbidity Index.<sup>17</sup>

The intraoperative data included type of anesthesia (general vs regional), intraoperative fluid administration (crystalloids, colloids, blood products), intraoperative estimated blood loss (EBL), and the duration of the surgery. Preoperative eGFR was calculated using the CKD-EPI Collaboration equation<sup>10</sup>:

$$\text{eGFR} = 141 \times \min(\text{Scr}/\kappa, 1)^\alpha \times \max(\text{Scr}/\kappa, 1)^{-1.209} \times 0.993^{\text{Age}} \times 1.018 [\text{if female}] \times 1.159 [\text{if black}],$$

where Scr is serum creatinine in  $\text{mg}\cdot\text{dL}^{-1}$ ,  $\kappa$  is 0.7 for females and 0.9 for males,  $\alpha$  is  $-0.329$  for females and  $-0.411$  for males, min indicates the minimum of  $(\text{Scr}/\kappa)$  or 1, and max indicates the maximum of  $(\text{Scr}/\kappa)$  or 1.

Patients were divided into four blood loss groups based on the intraoperative EBL: blood loss minimal (0-50 mL), blood loss 50-500 mL, blood loss 500-1,500 mL, and blood loss  $> 1,500$  mL. Separately, patients were dichotomized by need for intraoperative transfusion of red blood cells.

Preoperative eGFR was compared with preoperative serum creatinine (both assessed as continuous variables) on the ability to predict 30-day mortality, primarily in all patients and secondarily in each blood loss group and in males and females. The blood transfusion groups were used as a surrogate for invasiveness of the surgical procedure.

Baseline and intraoperative characteristics of the study population were described using standard summary statistics. For the primary outcome, 30-day mortality, we assessed the ability of preoperative serum creatinine vs preoperative eGFR to discriminate 30-day mortality using area under the receiver operating characteristic curve (i.e., c-statistic). The c-statistics for eGFR and creatinine were compared using the DeLong method of comparing dependent areas under the curve (AUCs).<sup>18</sup> We assessed the calibration of all models, i.e., how well the model fits the data, using the Hosmer-Lemeshow goodness of fit test.

For each eGFR and serum creatinine, we searched for a best cut-point that might be used in clinical decision-making, and we decided *a priori* that a clinically useful cut-point would have sensitivity and specificity  $> 0.75$  (corresponding to a positive likelihood ratio of  $\geq 3$  and a negative likelihood ratio of  $\leq 0.33$ ). With such a small incidence of the outcome (1.4% mortality), we did not expect the positive predictive value to be high – even with a sensitivity and specificity of 0.90, it would be only about 0.11.

We further assessed the discriminative ability of preoperative serum creatinine and preoperative eGFR within levels of blood loss for patients who received and did not receive transfusions, and for males and females. We again assessed calibration with Hosmer-Lemeshow goodness of fit tests.

In addition to the primary univariable analysis, we used multivariable logistic regression to assess the ability of each of eGFR and serum creatinine to improve discrimination of 30-day mortality (i.e., the c-statistic) beyond what is achieved using the baseline variables, i.e., sex, age, body mass index (BMI), and American Society of Anesthesiologists (ASA) physical status.

Finally, we calculated two complementary reclassification indices to assess the practical utility of using eGFR vs serum creatinine in predicting 30-day mortality – the net reclassification improvement (NRI) and the integrated discrimination improvement (IDI).<sup>19,20</sup> The NRI index compares upward and downward reclassification of risk for a new marker (or prediction model) vs a standard model between patients having the event (here, 30-day mortality)

and controls (here, alive at 30 days). A better predictor of the event would have more upward reclassification for patients with events than for those without events (controls). Specifically, the NRI index was calculated as twice the difference between cases and controls on the proportion of patients for whom the predicted risk was *higher* using eGFR than using serum creatinine minus the proportion for whom the predicted risk was *lower* using eGFR than using serum creatinine. A large NRI would indicate that eGFR improves risk categorization compared with serum creatinine. The corresponding formula is as follows:

$$\text{NRI} = 2\{[P(\text{up}) - P(\text{down})|\text{event}] - [P(\text{up}) - P(\text{down})|\text{non - event}]\}$$

where, for example, “P(up) | event” is the proportion of patients having the event who reclassified with higher (i.e., upward) risk with eGFR than with serum creatinine, and “P(down) | event” is the proportion of patients having the event who reclassified with lower (i.e., downward) risk with eGFR than with serum creatinine. Since the results are dependent on the size of the categories, we calculated the NRI with several methods of risk categorization: using the raw predictions for each patient and using categories of each 1%, 2%, and 4% of the predictions.

The IDI assesses the difference (typically, an increase) in the mean probability of an event using the new marker *vs* the old marker for cases (i.e., true events) minus the difference (typically, a decrease) in non-cases. Equivalently, it is the difference in discrimination slopes between the new and the old marker, where the discrimination slope is the absolute difference in the mean predicted probability of outcome between cases and controls for a marker. We also calculated the relative IDI, which is the ratio of discrimination slopes minus 1. The IDI can be considered a continuous version of the NRI.

The significance level was 0.05 for each hypothesis. Bonferroni correction was used to adjust for comparisons within subgroups. SAS<sup>®</sup> 9.3 statistical software (Carey, NC, USA) was used for all analyses.

## Results

We considered 93,012 patients for our investigation; however, 30-day mortality could not be determined in 124 cases, leaving 92,888 patients for analysis. Table 1 describes patient demographics on baseline, intraoperative factors, and in-hospital (1%) and 30-day (1.4%) mortality.

Estimated GFR (using the CKD-EPI formula) was a better discriminator of in-hospital mortality than serum creatinine, with an estimated c-statistic (95% confidence interval [CI]) of 0.67 (0.65 to 0.68) for eGFR *vs* 0.61 (0.59

to 0.63) for serum creatinine ( $P < 0.001$ ) (Table 2 and Figure). For comparison, the c-statistic (95% CI) using the MDRD formula for eGFR was only 0.62 (0.60 to 0.64). Neither single-predictor model had adequate calibration (Hosmer-Lemeshow goodness of fit,  $P < 0.001$  for each variable).

Furthermore, eGFR was consistently a better discriminator of 30-day mortality across blood loss, transfusion, and sex groups (Table 2). Univariably, the estimated odds ratio (95% CI) for 30-day mortality was 1.30 (1.27 to 1.32) for a ten-unit decrease in eGFR and 2.3 (2.1 to 2.5) for a doubling of creatinine (i.e., for an increase of 1 on the log-transformed base 2 scale).

No particular cut-off values of either eGFR or serum creatinine predicting 30-day mortality had sufficient diagnostic accuracy to be recommended for clinical decision-making. For example, a value of 76 for eGFR achieved estimated sensitivity and specificity of 0.66, negative predictive value (NPV) of 0.66, and positive predictive value (PPV) of 0.03. The positive likelihood ratio (probability of a positive test for diseased *vs* non-diseased) was 1.9, while the negative likelihood ratio (probability of a negative test for diseased *vs* non-diseased) was 0.52. A value of 0.94 mg·dL<sup>-1</sup> for serum creatinine achieved an estimated sensitivity and specificity of 0.62, NPV of 0.62, and PPV of 0.02. The positive likelihood ratio was 1.6, while the negative likelihood ratio was 0.61.

Reclassification analyses also suggested that eGFR is a better predictor than serum creatinine for individual patients. The net reclassification index was significantly greater than zero for all categorizations attempted, indicating improved prediction using eGFR compared with serum creatinine. Table 3 outlines the net reclassification index results, including the proportion of patients who were reclassified upwards and downwards for both cases and controls when using eGFR *vs* serum creatinine, as well as a sample calculation of the NRI. As expected, the NRI was highest when no categorizations were used (0.53; 95% CI 0.44 to 0.58). In that analysis, any increase (and separately, decrease) in the predicted probability of 30-day mortality, when using eGFR compared with serum creatinine, was considered reclassification. Using categories as large as each 4% for the predicted probabilities still showed a modest improvement using eGFR *vs* serum creatinine.

The estimated IDI index was also significantly greater than zero, indicating slightly higher on-average predictions for true events using eGFR *vs* using serum creatinine (Table 4). Although the estimated IDI is quite small (0.0037), the relative IDI of 0.567 indicates somewhat better prediction using eGFR *vs* serum creatinine. Due to the low incidence of mortality in the study, the mean probability of events was very small for both deaths (eGFR

**Table 1** Descriptive summary of patient population

Factor	Statistics
All, <i>n</i> %	92,888 (100%)
BMI, kg·m <sup>-2</sup> <sup>a</sup>	29 (8)
Height, cm <sup>a</sup>	170 (28)
Weight, kg <sup>b</sup>	85 (24)
Age, yr	57 (18-104) [47, 68]
Male, <i>n</i> %	43,761 (47%)
ASA status, <i>n</i> %	
I	3,948 (4%)
II	38,357 (41%)
III	43,860 (47%)
IV	6,723 (7%)
Black race, <i>n</i> %	11,578 (12%)
<i>Medical History</i>	
Coronary artery disease	9,486 (10%)
Hypertension	46,170 (47%)
Diabetes mellitus	17,775 (19%)
Peripheral vascular disease	7,973 (9%)
Chronic obstructive pulmonary disease	13,708 (15%)
Cerebrovascular disease	3,457 (4%)
<i>Surgical procedures – top 10</i> <sup>b</sup>	
Operations on the digestive system	20,832 (22.7%)
Operations on musculoskeletal system	18,245 (19.9%)
Operations on the nervous system	9,736 (10.6%)
Operations on the urinary system	9,702 (10.6%)
Operations on the cardiovascular system	6,586 (7.2%)
Operations on the female genital organs	6,416 (7.0%)
Operations on the integumentary system	5,124 (5.6%)
Operations on the male genital organs	4,909 (5.4%)
Operations on the endocrine system	3,476 (3.8%)
Operations on the nose; mouth; and pharynx	1,994 (2.2%)
<i>Charlson Comorbidity Index</i>	
0	42,141 (45%)
1-2	32,537 (35%)
3-4	10,003 (11%)
≥ 5	6,579 (9%)
Preoperative hemoglobin concentration (g·dL <sup>-1</sup> )	13 (2)
<i>Blood loss (mL), n %</i>	
<50	43,092 (46%)
50-499	39,679 (43%)
500-1,500	7,390 (8%)
>1,500	2,748 (3%)
Any RBC	7,775 (8%)
Duration of surgery, hr	3.4 [2.3, 4.8]
General anesthesia <sup>c</sup> , <i>n</i> (%)	81,744 (89%)
Total fluid <sup>d</sup> , mL	2,300 [1,300, 3,500]
Colloids, mL	0 [0, 500]
Crystalloids, mL	2,000 [1,200, 3,000]

**Table 1** continued

Factor	Statistics
eGFR, mL·min <sup>-1</sup> ·1.73 m <sup>-2</sup>	85 (25)
Creatinine, mg·dL <sup>-1</sup>	1.0 (0.5)
In-hospital mortality	955 (1.0%)
30-day mortality	1,326 (1.4%)

Data are mean (SD), median [quartiles], range (min–max) unless otherwise specified; ASA = American Society of Anesthesiologists; BMI = body mass index; eGFR = estimated glomerular filtration rate; RBC = red blood cell

<sup>a</sup>, <sup>b</sup>, and <sup>c</sup> missing 5%, 0.1%, and 0.9% of data points, respectively;

<sup>c</sup> Total fluid = colloids + crystalloids

0.024 vs serum creatinine 0.021) and non-deaths (eGFR 0.014 vs serum creatinine 0.014).

A multivariable combination of baseline characteristics ASA physical status, age, and BMI (all  $P < 0.001$ ) discriminated 30-day mortality with a c-statistic of 0.850. Adding eGFR to the model improved the c-statistic to only 0.851 (multivariable model, Table 5), while separately adding serum creatinine did not change the c-statistic. Nevertheless, in their separate multivariable models, both predictors were still significantly associated with 30-day mortality, with estimated odds ratio (95% CI) of 1.06 (1.04 to 1.09) for a ten-unit decrease in eGFR and 1.22 (1.11 to 1.34) for a doubling of creatinine.

In exploratory analyses, we found that the relationship between both log-transformed serum creatinine and eGFR with mortality was somewhat nonlinear and that calibration could be improved by adding quadratic and cubic forms of the predictor to each model. For eGFR, this did not affect the c-statistic, but for serum creatinine, the c-statistic improved to 0.67 (the same as eGFR). Reclassification was still marginally better with eGFR compared with serum creatinine in the more complex model. Nevertheless, since, in practice, clinicians would be more likely to use the raw serum creatinine or eGFR values and not a prediction with linear, quadratic, and cubed forms of the variables, we opted to report on the more simple prediction model for our main results.

## Discussion

Creatinine-based eGFR equations, which aim to circumvent some of the limitations of serum creatinine, are well validated.<sup>10,11</sup> Various equations have been proposed, but we chose the CKD-EPI Collaboration equation because it offers high precision, especially at higher GFR values.<sup>10,11</sup> Precision at a higher GFR is important when the equation is applied to overall surgical populations in which renal function is generally good.

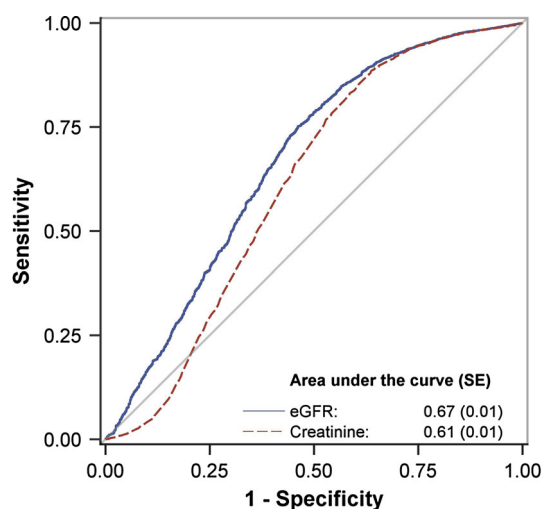
**Table 2** Comparing eGFR and creatinine on discrimination of 30-day vital status

Patients	n	c-statistic (95% CI)		Difference in c-statistic (SE)	Difference P value <sup>a, b</sup>
		eGFR	Creatinine		
All	92,888	0.67 (0.65 to 0.68)	0.61 (0.59 to 0.63)	0.06 (0.004)	< 0.001
Sex					
Male	43,761	0.65 (0.63 to 0.67)	0.59 (0.57 to 0.62)	0.06 (0.005)	< 0.001
Female	49,127	0.68 (0.66 to 0.71)	0.60 (0.57 to 0.63)	0.08 (0.006)	< 0.001
Blood loss (mL)					
<50	43,092	0.62 (0.60 to 0.65)	0.56 (0.52 to 0.59)	0.07 (0.008)	< 0.001
50-499	40,341	0.69 (0.66 to 0.71)	0.64 (0.61 to 0.66)	0.05 (0.007)	< 0.001
500-1,500	7,377	0.69 (0.65 to 0.74)	0.64 (0.59 to 0.69)	0.06 (0.010)	< 0.001
>1,500	2,740	0.65 (0.60 to 0.71)	0.62 (0.56 to 0.67)	0.04 (0.013)	0.005
Any RBC					
Yes	7,755	0.64 (0.61 to 0.66)	0.62 (0.59 to 0.65)	0.02 (0.005)	< 0.001
No	85,133	0.64 (0.61 to 0.66)	0.57 (0.55 to 0.60)	0.07 (0.006)	< .001

<sup>a</sup> Bonferroni correction was used to adjust for comparisons within subgroups, such that *P* values for other than the overall comparison were deemed significant if  $P < 0.05/8 = 0.0083$

<sup>b</sup> Delong method of comparing dependent areas under the curve (i.e., c-statistics)

CI = confidence interval; eGFR = estimated glomerular filtration rate; RBC = red blood cell; SE = standard error



**Figure** Comparative receiver operating characteristic curve plots of eGFR and creatinine as predictors of 30-day mortality. Area under the curve (AUC), or c-statistic, was significantly higher for eGFR, indicating better discrimination of 30-day vital status compared with serum creatinine alone ( $P < 0.001$ , Delong method). Diagonal line indicates an AUC of 0.5 corresponding to chance discrimination, to which the eGFR and creatinine AUCs are compared. eGFR = estimated glomerular filtration rate

Stevens *et al.*<sup>11</sup> compared precision of MDRD and CKD-EPI formulas across multiple ranges of GFR and found that the CKD-EPI equation can be used for an accurate estimate of GFR in patients with measured GFR  $> 60 \text{ mL}\cdot\text{min}^{-1}\cdot 1.73 \text{ m}^{-2}$ . On the other hand, the MDRD equation loses its accuracy if GFR is  $> 60 \text{ mL}\cdot\text{min}^{-1}\cdot 1.73 \text{ m}^{-2}$ . For that very reason, when eGFR exceeds  $> 60$

$\text{mL}\cdot\text{min}^{-1}\cdot 1.73 \text{ m}^{-2}$ , MDRD is reported as “eGFR  $> 60 \text{ mL}\cdot\text{min}^{-1}\cdot 1.73 \text{ m}^{-2}$ ” rather than as a discrete value. Therefore, mild renal dysfunction (in the range of  $60 \text{ mL}\cdot\text{min}^{-1}\cdot 1.73 \text{ m}^{-2}$  to normal value) cannot be quantified, and practically, based on eGFR calculated with the MDRD equation, it makes a patient with normal GFR indistinguishable from a patient with GFR just above  $60 \text{ mL}\cdot\text{min}^{-1}\cdot 1.73 \text{ m}^{-2}$  (clearly with some renal dysfunction), which would be a gross oversimplification. Thus replacing serum creatinine with the CKD-EPI (not MDRD) estimate of eGFR improves detection of mild preoperative renal dysfunction which otherwise often goes unnoticed but has been associated with worse outcomes.<sup>1,3,4</sup>

Our primary result is that eGFR is indeed a better discriminator of 30-day mortality than serum creatinine in a large cohort of patients having a wide variety of non-cardiac operations. The magnitude of the improvement, while highly statistically significant, was relatively small (c-statistic of 0.67 vs 0.61, respectively). Nonetheless, this represents an important improvement in discrimination, as eGFR is often reported by clinical laboratories and is otherwise easy to calculate. Although we did not find specific cut-points in eGFR or serum creatinine with adequate sensitivity and specificity to be reliable predictors of 30-day mortality, eGFR was clearly a better discriminator of mortality when used as a continuous variable.

Besides addressing discrimination and calibration, which are more global measures of prediction, we reported two reclassification indices to assess how well eGFR performs compared with serum creatinine in

**Table 3** Net reclassification improvement

NRI Categories	NRI (95% CI)	Events		Non-events		% Events correctly reclassified	% of Non-events correctly reclassified
		P(up)	P(down)	P(up)	P(down)		
None	0.531 (0.440 to 0.582)*	0.668	0.332	0.403	0.597	34%	19%
Each 1%	0.343 (0.308 to 0.378)*	0.397	0.112	0.170	0.228	29%	6%
Each 2%	0.131 (0.105 to 0.156)*	0.260	0.023	0.114	0.007	24%	11%
Each 4%	0.054 (0.037 to 0.071)*	0.085	0.017	0.018	0.004	7%	1%

\* *P* value < 0.0001; P (up) = proportion of patients reclassified up at least 1 category when using eGFR vs SCR; eGFR = estimated glomerular filtration rate; SCR = serum creatinine; P(down) = proportion of patients reclassified down at least 1 category when using eGFR vs SCR

CI = confidence interval; NRI = net reclassification improvement

NRI categories: groupings for predicted probability of 30-day mortality across all patients

Example: NRI using 2% categories of predicted probability of 30-day mortality

$$NRI = [P(\text{up}) - P(\text{down} | \text{event})] - [P(\text{up}) - P(\text{down} | \text{non-event})] = [0.260 - 0.023] - [0.114 - 0.007] = 0.237 - 0.107 = \sim 0.131$$

**Table 4** Integrated discrimination improvement

IDI (95% CI)	Mean Probability* for Events		Mean Probability* for Non-Events		Probability of change		Relative IDI
	Model eGFR	Model SCR	Model eGFR	Model SCR	Events	Non-events	
0.00374 (0.0033 to 0.0041)*	0.0244	0.0207	0.0141	0.0142	0.0036	-0.00005	0.567

CI = confidence interval; IDI = integrated discrimination improvement; \**P* < 0.001 vs IDI of zero; eGFR = estimated glomerular filtration rate; SCR = serum creatinine

\* Mean probability =  $\bar{p}$  = mean predicted probability of having the event (30-day mortality)

IDI = discrimination slope (eGFR) – discrimination slope (SCR)

$$= (\bar{p}_{\text{event, eGFR}} - \bar{p}_{\text{non-event, eGFR}}) - (\bar{p}_{\text{event, SCR}} - \bar{p}_{\text{non-event, SCR}}) = (0.0244 - 0.0141) - (0.0207 - 0.0142) = 0.00374$$

$$\text{Relative IDI} = \text{discrimination slope (eGFR)} / \text{discrimination slope (SCR)} - 1 = (0.0244 - 0.0141) / (0.0207 - 0.0142) - 1 = 0.567$$

**Table 5** Multivariable model\* for association with 30-day mortality

Factor	Units	Odds ratio (95% CI)	<i>P</i> value
eGFR, percent	10	0.94 (0.92 to 0.96)	< 0.001
Age, yr	10	1.20 (1.14 to 1.26)	< 0.001
BMI, kg·m <sup>-2</sup>	5	0.87 (0.84 to 0.91)	< 0.001
ASA physical status, level	1	7.6 (6.8 to 8.5)	< 0.001

\*Logistic regression model c-statistic = 0.851

Relationships between age, BMI, ASA, and mortality are almost identical to this model when using serum creatinine instead of eGFR

ASA = American Society of Anesthesiologists; BMI = body mass index; CI = confidence interval; eGFR = estimated glomerular filtration rate

predicting 30-day mortality for individual patients. Our results indicate that using eGFR would more appropriately predict 30-day mortality compared with serum creatinine, i.e., patients who would eventually die by 30 days would be more likely to have a higher prediction of mortality using eGFR, adjusting for those who were wrongly reclassified. The NRI has possible limitations: it is sensitive to the calibration of the model; it may be significantly different from zero for very subtle improvements in prediction, and

it is dependent on the number of reclassification categories used.<sup>21,22</sup> Nevertheless, in our study, the NRI showed a clear benefit of eGFR over serum creatinine – eGFR, on average, placed patients who would eventually die by 30 days into a higher risk category than would serum creatinine, independent of the number of risk categories used (i.e., up to 34% of deaths correctly reclassified).

The benefit of using eGFR rather than creatinine alone extended across blood loss and transfusion – both of which can be considered surrogates for the invasiveness of the surgical procedure – as well as across sex groups. We did separate analyses for males and females because, as opposed to the CKD-EPI formula, serum creatinine does not distinguish difference between sexes. Estimated GFR was thus a modestly superior predictor of mortality in each circumstance we considered. Our results strongly suggest that eGFR, as calculated using the CKD-EPI formula, should be used in preference to serum creatinine. For the sake of interest, we also calculated GFR using the MDRD formula and observed a c-statistic (standard error) of only 0.62 (0.01), indicating similar discrimination of 30-day mortality compared with creatinine alone.

As shown in many previous studies, even slight impairments in renal function were associated with substantial increases in 30-day mortality and/or cardiac events – whether renal function was evaluated on the basis of serum creatinine or eGFR.<sup>1,23</sup> Nevertheless, postoperative mortality is affected by multiple perioperative factors, with renal function being just one of them. It was therefore unlikely that preoperative eGFR alone would be a strong predictor of 30-day mortality in surgical patients. As one might thus expect, a multivariable combination of ASA physical status, age, and BMI predicted 30-day mortality far better than just eGFR (c-statistic of 0.85 compared with 0.67, respectively).

In summary, eGFR using the CKD-EPI formula is a modestly better predictor of 30-day mortality than serum creatinine in patients having non-cardiac inpatient surgery. Our study was based on a large and diverse non-cardiac surgical population. Furthermore, eGFR proved modestly superior to serum creatinine alone in various subgroups; our conclusions about the relative merits of eGFR and creatinine thus appear to be broadly applicable. Given that eGFR is often reported by clinical laboratories and is otherwise easy to calculate, it should generally be used in preference to serum creatinine alone.

**Conflicts of interest** None declared.

**Funding** Departmental funds.

## References

1. Mooney JF, Ranasinghe I, Chow CK, et al. Preoperative estimates of glomerular filtration rate as predictors of outcome after surgery: a systematic review and meta-analysis. *Anesthesiology* 2013; 118: 809-24.
2. Tollefson MK, Boorjian SA, Gettman MT, Rangel LJ, Bergstralh EJ, Karnes RJ. Preoperative estimated glomerular filtration rate predicts overall mortality in patients undergoing radical prostatectomy. *Urol Oncol* 2013; 31: 1483-8.
3. Shavit L, Tauber R, Lifschitz M, Bitran D, Slotki I, Fink D. Influence of minimal changes in preoperative renal function on outcomes of cardiac surgery. *Kidney Blood Press Res* 2012; 35: 400-6.
4. Ackland GL, Moran N, Cone S, Grocott MP, Mythen MG. Chronic kidney disease and postoperative morbidity after elective orthopedic surgery. *Anesth Analg* 2011; 112: 1375-81.
5. Lee TH, Marcantonio ER, Mangione CM, et al. Derivation and prospective validation of a simple index for prediction of cardiac risk of major noncardiac surgery. *Circulation* 1999; 100: 1043-9.
6. Higgins TL, Estafanous FG, Loop FD, Beck GJ, Blum JM, Paranandi L. Stratification of morbidity and mortality outcome by preoperative risk factors in coronary artery bypass patients. A clinical severity score. *JAMA* 1992; 267: 2344-8.
7. Davis C, Tait G, Carroll J, Wijeyesundera DN, Beattie WS. The Revised Cardiac Risk Index in the new millennium: a single-centre prospective cohort re-evaluation of the original variables in 9,519 consecutive elective surgical patients. *Can. J Anesth* 2013; 60: 855-63.
8. Lamb E. Assessment of kidney function in adults. *Medicine* 2011; 39: 306-11.
9. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. *Ann Intern Med* 1999; 130: 461-70.
10. Levey AS, Stevens LA, Schmid CH, CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration), et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009; 150: 604-12.
11. Stevens LA, Schmid CH, Greene T, et al. Comparative performance of the CKD Epidemiology Collaboration (CKD-EPI) and the Modification of Diet in Renal Disease (MDRD) Study equations for estimating GFR levels above 60 mL/min/1.73 m<sup>2</sup>. *Am J Kidney Dis* 2010; 56: 486-95.
12. Devbhandari MP, Duncan AJ, Grayson AD, North West Quality Improvement Programme in Cardiac Interventions, et al. Effect of risk-adjusted, non-dialysis-dependent renal dysfunction on mortality and morbidity following coronary artery bypass surgery: a multi-centre study. *Eur J Cardiothorac Surg* 2006; 29: 964-70.
13. Tolpin DA, Collard CD, Lee VV, et al. Subclinical changes in serum creatinine and mortality after coronary artery bypass grafting. *J Thorac Cardiovasc Surg* 2012; 143(682-8): e1.
14. Thourani VH, Keeling WB, Sarin EL, et al. Impact of preoperative renal dysfunction on long-term survival for patients undergoing aortic valve replacement. *Ann Thorac Surg* 2011; 91: 1798-806.
15. O'Brien MM, Gonzales R, Shroyer AL, et al. Modest serum creatinine elevation affects adverse outcome after general surgery. *Kidney Int* 2002; 62: 585-92.
16. Collins TC, Johnson M, Daley J, Henderson WG, Khuri SF, Gordon HS. Preoperative risk factors for 30-day mortality after elective surgery for vascular disease in Department of Veterans Affairs hospitals: is race important? *J Vasc Surg* 2001; 34: 634-40.
17. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987; 40: 373-83.
18. 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: 837-45.
19. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008; 27: 157-72.
20. 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.
21. Pencina MJ, D'Agostino RB, Pencina KM, Janssens AC, Greenland P. Interpreting incremental value of markers added to risk prediction models. *Am J Epidemiol* 2012; 176: 473-81.
22. Kerr K, Bansal A, Pepe M. Further insight into the incremental value of new markers: the interpretation of performance measures and the importance of clinical context. *Am J Epidemiol* 2012; 176: 482-7.
23. Gupta PK, Gupta H, Sundaram A, et al. Development and validation of a risk calculator for prediction of cardiac risk after surgery. *Circulation* 2011; 124: 381-7.