Estimating Baseline Kidney Function in Hospitalized Patients with Impaired Kidney Function

Edward D. Siew,* T. Alp Ikizler,* Michael E. Matheny,*‡§† Yaping Shi,* Jonathan S. Schildcrout,* Ioana Danciu,* Jamie P. Dwyer,* Manakan Srichai,* Adriana M. Hung,*** James P. Smith,†† and Josh F. Peterson‡§*

Summary

Background and objectives Inaccurate determination of baseline kidney function can misclassify acute kidney injury (AKI) and affect the study of AKI-related outcomes. No consensus exists on how to optimally determine baseline kidney function when multiple predmission creatinine measurements are available.

Design, setting, participants, & measurements The accuracy of commonly used methods for estimating baseline serum creatinine was compared with that of a reference standard adjudicated by a panel of board-certified nephrologists in 379 patients with AKI or CKD admitted to a tertiary referral center.

Results Agreement between estimating methods and the reference standard was highest when using creatinine values measured within a week of admission. The mean outpatient serum creatinine measured within a year of hospitalization most closely approximates nephrologist-adjudicated serum creatinine values.

Conclusions The mean outpatient serum creatinine measured within a year of hospitalization most closely approximates nephrologist-adjudicated serum creatinine values.

Introduction

Acute kidney injury (AKI) is common among hospitalized patients and is independently associated with morbidity and mortality. Although recently accepted diagnostic criteria rely on relative or fixed changes in serum creatinine, published studies have differed in their approach to determining baseline kidney function (1–9). We have previously demonstrated that the use of inpatient values or a single estimate of baseline kidney function results in bidirectional misclassification of AKI incidence and severity (10). In studies for which predmission serum creatinine data are available, an optimal approach to identifying baseline kidney function has not been identified.

In this study, we examined predmission serum creatinine measurements in a cohort of 379 high-risk patients admitted to a tertiary care academic medical center over a 12-month period. The study objective was to examine how commonly used methods for estimating baseline kidney function, including using the most recent, average, or nadir serum creatinine values, most closely approximated clinical nephrology opinion. We measured agreement between each of these commonly applied algorithms to a reference standard for baseline serum creatinine as determined by a blinded panel of board-certified nephrologists. We also examined how performance of each algorithm varied with the time interval chosen up to 2 years before admission.

Materials and Methods

Study Population

Adult patients (≥18 years of age) were eligible if admitted to Vanderbilt University Hospital for ≥24 continuous hours between October 1, 2007, and October 1, 2008. As the main adult inpatient facility within Vanderbilt University Medical Center (VUMC), Vanderbilt University Hospital is a large tertiary referral center serving middle Tennessee and surrounding areas. Because patients without predmission serum creatinine data will require a surrogate baseline value whereas those with only one predmission value do not require a specific selection algorithm (10, 11), we required study patients to have at least two creatinine values recorded in their electronic medical record within 24 months before hospitalization. Our aim was to focus on patients for whom determination of baseline kidney function might differ on the basis of the selection approach used (e.g., nadir, mean, most recent), thereby influencing AKI classification and staging.
We selected a population with evidence of impaired renal function on admission (estimated GFR < 60 ml/min per 1.73 m²) or AKI defined by Acute Kidney Injury Network diagnostic criteria (increase in serum creatinine of 0.3 mg/dl or 50% over any 48-hour interval). The study excluded patients with previous evidence of stage 5 CKD or ESRD according to prior International Classification of Diseases, Ninth Revision (ICD-9), code assignment of 585.6, 996.73, 996.68, 996.56, 792.5, or 458.21. For patients with multiple admissions, a single admission was chosen at random for evaluation. The VUMC Institutional Review Board approved the study protocol. Retrospective patient records were anonymized for analysis and the institutional review board approved waiver of consent.

Data Collection

General demographic information and laboratory data were collected from the institutional electronic medical record, which serves inpatients and outpatients at VUMC. Creatinine measurements were retrieved for each eligible patient for 2 years before admission. Information on comorbid conditions was identified using the corresponding ICD-9 codes for diabetes mellitus, hypertension, coronary artery disease, cerebrovascular disease, congestive heart failure, and peripheral vascular disease assigned before hospitalization (Appendix 1).

Adjudication of Baseline Kidney Function

Two board-certified nephrologists (M.S., J.P.D.) served as primary reviewers and were asked to determine a preadmission creatinine value that best reflected each patient’s baseline kidney function immediately before the acute illness that had prompted hospitalization. Reviewers were given access to the entire preadmission electronic medical records, including outpatient notes, problem lists, previous hospitalizations, emergency department or acute clinic visits, radiographic studies, consultations, any and all laboratory values measured at primary institution, and outside creatinine values obtained and scanned into the chart. The overarching goal was to reflect a similar clinical practice environment encountered by nephrologists asked to see a patient shortly after admission. We did not insist that a specific documented pre-existing value be used but rather allowed the adjudicators the flexibility to present a value that they thought best represented each patient’s baseline. For example, if a patient had several values of 0.8 mg/dl and 1.0 mg/dl, the adjudicator was allowed to enter 0.9 mg/dl if he or she thought it most representative of the patient’s baseline after examining the clinical scenario in which previous measurements were made.

Disagreement between reviewers was classified as a ≥0.2-mg/dl difference in baseline creatinine assessment if the lesser of two values chosen was ≤2.0 mg/dl, or as a >10% difference if the lesser of two values chosen was >2.0 mg/dl. The standard for disagreement was selected by considering the following: (1) the analytical coefficient of variation and the intra-individual coefficient of variation that contribute to an overall variation of 10%–20% (11, 12) and (2) the consensus opinion of a panel of local board-certified nephrologists. When the primary reviewers agreed, the final baseline creatinine value was the mean of the two chosen values. In cases of disagreement, the two primary reviewers and an independent board-certified nephrologist (A.M.H.) reviewed the case. To remain consistent with methods used for the primary reviewers, the final creatinine value was then determined using the mean between the third reviewer and the closer of the two chosen values from the primary reviewers.

Comparison between Baseline Estimation Algorithms and Clinical Nephrologist Opinion

After adjudicated determination of baseline creatinine as a reference standard, we compared the following algorithms for estimating baseline creatinine: (1) the mean outpatient value, (2) the most recent outpatient value, (3) the nadir outpatient value, and (4) the most recent inpatient or outpatient value. Three time intervals were also chosen for study: 7–365, 7–730, and 1–730 days before admission.

Statistical Analyses

Categorical and continuous variables were summarized using percentages and median and interquartile range (IQR), respectively. The intraclass correlation coefficient (ICC) with 95% confidence interval (CI) was used to measure the agreement between reviewers as well as between the final adjudicated serum creatinine values and surrogate estimation methods. Bland-Altman plots were used to illustrate agreement between surrogate estimation methods and the reference standards across the range of serum creatinine values. Pairwise comparisons of ICCs were conducted using the Konishi-Gupta modified Z test (13). Because overlap in creatinine value availability was incomplete between time intervals, only comparisons between estimation methods within the same intervals with identical availability were compared. P values from multiple comparisons were adjusted using the Bonferroni method. Analyses were performed using R 2.13.0 (R Foundation for Statistical Computing, Vienna, Austria) and SAS software, version 9.2 (SAS Institute, Inc., Cary, NC). A two-sided P value < 0.05 was considered to indicate a statistically significant result.

Results

Among 13,781 adult patients hospitalized with AKI or impaired GFR on admission, 400 were randomly selected for the study population. Twenty-one patients without a valid measurement of serum creatinine during hospitalization or with no serum creatinine value measured before admission were excluded. Table 1 lists the demographic, comorbidity, and hospital characteristics of the study population. A slight female predominance (53%) was observed, and most patients self-identified as white (83%). At index admission, patient comorbid conditions included hypertension (65%), diabetes mellitus (35%), coronary artery disease (43%), and congestive heart failure (32%).

Preadmission Serum Creatinine Availability

Figure 1 depicts the availability and counts of preadmission serum creatinine values grouped by time interval. A median of 8 (IQR, 3–18) total, 3 (IQR, 1–8) outpatient, and 3 (IQR, 0–9) inpatient creatinine measurements were recorded between 1 and 730 days before admission. A substantial portion of total creatinine measurements (median, 5 [IQR, 2–13]) occurred
Comparison of Adjudicated Baseline Serum Creatinine with Prespecified Estimation Methods

Overall agreement between the primary reviewers on baseline serum creatinine values was 88%, with an ICC of 0.93 (95% CI, 0.91–0.94). A third independent reviewer (A.M.H.) resolved 45 (12%) cases of disagreement. Median final adjudicated baseline creatinine and estimated GFR values were 1.4 mg/dl and 49 ml/kg per 1.73 m², respectively.

Table 2 shows the prespecified estimation methods studied, which included the most recent outpatient, the mean outpatient, the nadir outpatient, and the most recent inpatient or outpatient measure taken from the preadmission time intervals. The highest overall ICCs were noted for estimation methods using creatinine values between 7 and 365 days (ICC range, 0.83–0.91) before admission. Among outpatient values, an ICC of 0.91 (95% CI, 0.88–0.92) for the mean outpatient serum creatinine was higher than the ICCs for both the most recent outpatient creatinine (0.84 [95% CI, 0.80–0.88]; P<0.001) and the nadir outpatient serum creatinine (0.83 [95% CI, 0.76–0.87]; P<0.001). Allowance of the final creatinine measurement from a prior inpatient stay during the 7- to 365-day interval increased the ICC of the most recent outpatient value to 0.88 (95% CI, 0.85–0.91).

Extending the prehospitalization time interval increased the proportion of patients with baseline creatinine values available for assessment. In our study cohort, extending the time interval from 7–365 days to either 7–730 days or the entire 2-year interval (1–730 days) increased the proportion of patients with at least one outpatient creatinine value available from 81% (n=306) to 85% (n=321) or to 87% (n=329), respectively. Allowing the final inpatient creatinine to be included when an outpatient value was not available for most recent outpatient visit also increased the proportion of patients available to 93% (n=355). However, accuracy decreased as the time interval increased. For example, when the time interval was increased from 7–365 to 7–730 days before admission, the ICC for the mean outpatient value decreased from 0.91 (95% CI, 0.88–0.92) to 0.81 (95% CI, 0.77–0.84). Similarly, the nadir outpatient ICC was reduced from 0.83 (95% CI, 0.76–0.87) to 0.64 (95% CI, 0.46–0.75). Reductions in the most recent outpatient or most recent inpatient or outpatient method were more modest, at 0.84 (95% CI, 0.80–0.88) to 0.83 (95% CI, 0.78–0.86) and 0.88 (95% CI, 0.85–0.91) to 0.88 (95% CI, 0.85–0.91), respectively. Adding creatinine data from the week before (1–7 days) hospital admission resulted in poorer performance using mean outpatient, most recent outpatient, and most recent inpatient or outpatient methods.

Figure 2 illustrates the agreement between estimation methods for the 7- to 365-day time interval and adjudicated creatinine values. In addition, the proportion of estimated values decreasing at least 0.3 mg/dl above or below the adjudicated value is highlighted and drawn as an indicator of potential AKI misclassification. The mean outpatient serum creatinine, the most recent outpatient value, and most recent inpatient or outpatient methods had similar proportions of patients with decreases of at least 0.3 mg/dl above or below the adjudicated value at 11.1%, 13.7%, and 12.1%, respectively. In most of these cases, estimated baseline creatinine values were higher than the adjudicated value. In contrast, using the nadir outpatient serum creatinine resulted in 20.3% of estimated baseline values decreasing at least 0.3 mg/dl above or below the adjudicated serum creatinine value. In most of these cases, nadir serum creatinine was lower than the adjudicated creatinine value. In all estimation methods, the direction of the overall estimation error remained consistent as the time interval broadened while the proportion of patients falling outside of a 0.3-mg/dl deviation increased (data not shown).
Sensitivity Analyses

We performed two companion analyses. Because pre-existing CKD is a major challenge in assessing baseline kidney function, we enriched our study population for those with abnormal kidney function. Table 3 shows the ICCs for each estimation algorithm stratified according to baseline CKD status, defined as an eGFR derived from the adjudicated serum creatinine value of <60 ml/min per 1.73 m² for the 7- to 365-day interval before hospitalization. For patients with and those without CKD, mean outpatient serum creatinine represented a closer approximation of clinical nephrology opinion than did other algorithms tested.

Because the main study site involved a large academic tertiary referral center, the study population probably comprised patients from various practice settings and comorbidity backgrounds. To account for the effects of potential ascertainment bias of preadmission serum creatinine data,

Table 2. Agreement between estimation method and adjudicated reference standard

<table>
<thead>
<tr>
<th>Estimation Method</th>
<th>ICC (95% CI) per Days Before Admission</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7–365 d</td>
</tr>
<tr>
<td>Most recent outpatient</td>
<td>0.84 (0.80–0.88)(^a)</td>
</tr>
<tr>
<td>Mean outpatient</td>
<td>0.91 (0.88–0.92)(^a)</td>
</tr>
<tr>
<td>Nadir outpatient</td>
<td>0.83 (0.76–0.87)(^a)</td>
</tr>
<tr>
<td>Most recent inpatient or outpatient</td>
<td>0.88 (0.85–0.91)(^d)</td>
</tr>
</tbody>
</table>

All comparisons are made to the reference standard. ICC is calculated using a population of patients with sufficient creatinine values to yield an estimation method result. ICC, intraclass correlation coefficient; CI, confidence interval.

\(^a\)\(n=306\) patient (81% of cohort).
\(^b\)\(n=321\) patients (85% of cohort).
\(^c\)\(n=329\) patients (87% of cohort).
\(^d\)\(n=355\) patients (93% of cohort).
\(^e\)\(n=379\) patients (100% of cohort).
we further stratified our analysis by the number of creatinine measurements available during the 7- to 365-day preadmission time interval (Table 3). Patients were grouped into tertiles based on preadmission creatinine measurement availability (0–6, 6–14, and 14–93 values). ICCs for the mean outpatient algorithm were consistently higher than those for other algorithms tested within each tertile of creatinine availability.

Discussion

Accurate determination of baseline kidney function anchors the definition of AKI and establishes a longitudinal reference point in studies examining the long-term consequences of AKI (14). Variability in applied baselines can contribute to heterogeneity in reported AKI frequency (2,9,10,15), reduce fidelity during cross-validation, and confound examination of the longitudinal effect of AKI (14). We have shown that among high-risk patients, use of a mean outpatient serum creatinine between 7 and 365 days before admission most closely approximates clinical nephrology opinion for determining baseline kidney function. Higher ICCs yielded lower rates of potential AKI misclassification at baseline. The results also suggest that broadening the temporal detection window increases the number of patients with sufficient information to

Figure 2. | Bland-Altman plots comparing selected approaches for estimating baseline kidney function with expert adjudication during the 7- to 365-day time interval. (A) Most recent outpatient. (B) Mean outpatient. (C) Nadir outpatient. (D) Most recent inpatient or outpatient. x-axis represents the average of serum creatinine (SCr) value between the estimation method and adjudicated value. y-axis indicates the difference in serum creatinine values between the adjudicated value and the estimation method. The dotted lines represent a 0.3-mg/dl deviation in either direction. The right side of each figure shows a histogram of the difference between the adjudicated value and the estimation method. The percentage values show the potential bias in AKI diagnosis indicating the estimated values at least 0.3 mg/dl deviations above or below the adjudicated value. ICC, intraclass correlation coefficient.
establish a baseline at the expense of reduced accuracy in

Modern consensus definitions have helped to standardize how AKI is detected and compared across multiple settings (16,17). The Acute Kidney Injury Network definition currently relies on a rolling 48-hour window of detection for AKI (1). This specification helps to decrease the ambiguity in using functional markers for ensuring injury acuity and may be reasonable in patients with well preserved kidney function on presentation. However, the risks of not accounting for preadmission kidney function, particularly among those with abnormal kidney function on presentation, are considerable and include bidirectional misclassification of disease status (10). This potential source of error is amplified further with modest changes in serum creatinine included in modern definitions.

The optimal method for incorporating preadmission creatinine data into an accurate characterization of baseline is complicated by methodologic challenges involving data availability, variability, and quality. In the absence of a consensus approach, we created a reference standard that matches typical clinical practice by having adjudicators review all available clinical data. Because implementing manual adjudication in large-scale observational studies of AKI is not feasible, we sought to identify a quantitative rule that could be applied to a cohort of hospitalized patients to accurately classify changes in kidney function during hospitalization. The superiority of the mean outpatient creatinine in matching the adjudicated value may reflect a clinical recognition by the adjudicators that averaging multiple serum creatinine values, if available, reduces the inappropriate effect of outliers or the known effect of physiologic variation (18–20). This latter observation suggests that for a large proportion of patients at risk for AKI, baseline kidney function may be more accurately described as a range than as a single value. Averaging of serial data

Figure 2. | Continued
points during the recording of physiologic variables has already been found to enhance predictive power in other clinical settings. For example, both the American Heart Association and recent scientific literature have endorsed averaging multiple measurements in the recording of ambulatory BP for decision-making and quality reporting (21,22). This strategy is also commonly used to account for intraindividual variation in both establishing baseline and documenting changes in BP that occur in clinical trials of patients with hypertension (23).

Averaging of outpatient creatinine values may be less optimal in certain subsets of patients, especially if prevalent within a studied cohort. For example, the most recent serum creatinine may be most informative in patients with progressive CKD, especially among those with few data points (e.g., two or fewer). Nonaveraged values in this subgroup of patients may also be useful in detailing the premorbid trajectory of kidney function. Because a large proportion of patients in our study population had preexisting kidney impairment, this may partially explain why, despite lower ICCs, similar rates of potential AKI misclassification were observed between most recent serum creatinine values (Figure 1A, A and D) and mean outpatient values. It may also partially explain why the nadir value, even when restricted to outpatient values, resulted in the largest proportion of patients potentially misclassified (Figure 1C).

Our data also suggest that the defined time interval for capturing baseline creatinine data affects accuracy. Creatinine measurements proximal to the admission date (i.e., within a week) or remote (>1 year) often decrease accuracy. In patients in whom creatinine was recorded within a week of hospitalization, values tended to be higher than the adjudicated creatinine values (data not shown), suggesting that decreasing kidney function or increased creatinine turnover coincided with the onset of acute illness. Including more remote creatinine data (e.g., >1 year) may improve sample size and reduce selection bias; however, this benefit must be weighed against the inability to account for interim changes in kidney function and ensure that decrements in kidney function observed on presentation are truly acute.

The study is limited by the data available within the electronic medical record of our institution; some preadmission serum creatinine data measured elsewhere may be missing from the analysis. However, reviewers were allowed access to records scanned outside, and our requirement that patients have two preadmission serum creatinine measurements focused the analysis on patients typically followed within our institution. In addition, although not all adjudicators were trained at the same institution, the reference standard for baseline serum creatinine was determined within a single center, and we cannot rule out the contribution of potential institutional bias in the clinical interpretation of creatinine values. Finally, we enriched our population with patients who had abnormal kidney function during hospitalization, thereby increasing the possible effects of ascertainment bias. However, our sensitivity analysis demonstrating the consistently superior performance of mean averaged outpatient value across the spectrum of creatinine availability does not support this hypothesis. Nonetheless, these findings should be validated in other study settings.

In summary, using functional markers to define and study AKI and its consequences requires accurate quantification of baseline status. Use of averaged outpatient values between 7 and 365 days before admission yields a reliable rule for estimating baseline kidney function among high-risk patients in a tertiary care setting. Further work to validate this approach in other clinical settings, as well as to improve determination of baseline kidney function when preadmission data are not readily available, is warranted.

Acknowledgments

This work was funded by Grant R01 LM009965-01 from the National Library of Medicine. E.D.S. is supported by Grant K23 DK088964-01A1 from the National Institute of Diabetes and Digestive and Kidney Diseases. The work of M.E.M. is supported in full by the Career Development Program from the Department of Veterans Affairs (CDA-08-020 HSR&D). The work of A.M.H. is supported in full by the Career Development Program from the Department of Veterans Affairs (CDA [2-031-095] CSR&D).

Separate pilot data were presented in abstract form at the 2009 American Society of Nephrology Annual Scientific Meeting, October 27 to November 1, 2009, San Diego, California.

---

Table 3. Agreement between estimation method and adjudicated reference standard stratified by CKD status and number of preadmission serum creatinine values available

<table>
<thead>
<tr>
<th>Estimation Method</th>
<th>ICC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>eGFR &lt; 60 ml/min per 1.73 m² (n=259)</td>
</tr>
<tr>
<td>Most recent outpatient</td>
<td>0.80 (0.74–0.85)</td>
</tr>
<tr>
<td>Mean outpatient</td>
<td>0.87 (0.83–0.90)</td>
</tr>
<tr>
<td>Nadir outpatient</td>
<td>0.76 (0.66–0.83)</td>
</tr>
<tr>
<td>Most recent inpatient or outpatient</td>
<td>0.85 (0.81–0.88)</td>
</tr>
</tbody>
</table>

CKD status was defined using the baseline adjudicated creatinine value. ICC, intraclass correlation coefficient; CI, confidence interval; eGFR, estimated GFR.
Disclosures

None.

References


Received: October 19, 2011 Accepted: February 16, 2012

Published online ahead of print. Publication date available at www.cjasn.org.


Access to UpToDate on-line is available for additional clinical information at www.cjasn.org.