Article

Use of Multiple Imputation Method to Improve Estimation of Missing Baseline Serum Creatinine in Acute Kidney Injury Research

Edward D. Siew,* Josh F. Peterson,[†‡§¶] Svetlana K. Eden[‖] Karel G. Moons,[‖**] T. Alp Ikizler,* and Michael E. Matheny[†‡§¶]

Summary

Background and objectives Baseline creatinine (BCr) is frequently missing in AKI studies. Common surrogate estimates can misclassify AKI and adversely affect the study of related outcomes. This study examined whether multiple imputation improved accuracy of estimating missing BCr beyond current recommendations to apply assumed estimated GFR (eGFR) of 75 ml/min per 1.73 m² (eGFR 75).

Design, setting, participants, & measurements From 41,114 unique adult admissions (13,003 with and 28,111 without BCr data) at Vanderbilt University Hospital between 2006 and 2008, a propensity score model was developed to predict likelihood of missing BCr. Propensity scoring identified 6502 patients with highest likelihood of missing BCr among 13,003 patients with known BCr to simulate a "missing" data scenario while preserving actual reference BCr. Within this cohort (n=6502), the ability of various multiple-imputation approaches to estimate BCr and classify AKI were compared with that of eGFR 75.

Results All multiple-imputation methods except the basic one more closely approximated actual BCr than did eGFR 75. Total AKI misclassification was lower with multiple imputation (full multiple imputation + serum creatinine) (9.0%) than with eGFR 75 (12.3%; P<0.001). Improvements in misclassification were greater in patients with impaired kidney function (full multiple imputation + serum creatinine) (15.3%) versus eGFR 75 (40.5%; P<0.001). Multiple imputation improved specificity and positive predictive value for detecting AKI at the expense of modestly decreasing sensitivity relative to eGFR 75.

Conclusions Multiple imputation can improve accuracy in estimating missing BCr and reduce misclassification of AKI beyond currently proposed methods.

*Department of Medicine, Division of Nephrology, Vanderbilt University Medical Center, Nashville, Tennessee; †Health Services Research and Development, Tennessee Valley Healthcare System, Veterans Affairs Medical Center, Nashville, Tennessee; ‡Department of Biomedical Informatics, §Division of General Internal Medicine, and ‖Department of Biostatistics, Vanderbilt University Medical Center, Nashville, Tennessee; ¶Geriatric Research, Education and Clinical Center and **Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands

Correspondence:
Dr. Edward D. Siew, Department of Medicine, Division of Nephrology and Hypertension, Vanderbilt University Medical Center, 1161 21st Avenue South, MCN S3223, Nashville, TN 37232-2372. Email: edward.siew@vanderbilt.edu

Introduction

AKI commonly complicates hospitalization and is strongly associated with morbidity and mortality (1,2). The diagnosis of AKI relies on accurately quantifying changes in serum creatinine from a baseline value (3). A common limitation of studying AKI is the unavailability of baseline creatinine in hospital datasets, prompting the use of surrogate estimates. The Acute Dialysis Quality Initiative has proposed a solution: assigning a single estimated GFR (eGFR) value (*i.e.*, 75 ml/min per 1.73 m²) to all patients with missing data (eGFR 75). Although attractive for its simplicity and applicability, this technique disregards important patient information that can lead to poor estimation of kidney function, misclassify AKI, and adversely affect the study of associated outcomes (4,5).

Multiple imputation is a widely used approach that allows estimation of missing data in statistical analysis (6). Through leveraging of known patient characteristics and accounting for uncertainty in the multiple estimations of missing values, multiple imputation preserves sample size and reduces bias in modeling associations between variables (7–10). We hypothesized that multiple imputation would improve the accuracy of estimating baseline creatinine among hospitalized patients beyond imputing eGFR 75 (11). Using a large cohort of hospitalized adults with known baseline creatinine as a reference, we studied how various multiple-imputation approaches compared with eGFR 75 in estimating actual baseline creatinine values and accurately classifying AKI.

Methods

Study Population

Vanderbilt University Hospital (VUH), the main adult inpatient facility at Vanderbilt Medical Center, is an 832-bed tertiary referral center serving middle Tennessee and its surrounding regions. We identified 61,345 VUH admissions with a length of stay of ≥24 hours between October 1, 2006, and September 30, 2008, and ≥1 serum creatinine measurement during

the first 7 days of hospitalization. Admissions with a diagnosis of ESRD according to prior International Classification of Diseases, Ninth Revision (ICD-9), code assignments of 585.6, 996.73, 996.68, 996.56, 792.5, or 458.21 were excluded. For patients with multiple admissions, a single admission was chosen randomly, leaving 41,114 unique patient admissions. The Vanderbilt University Medical Center Institutional Review Board approved the study.

Demographic and laboratory data were collected from the institutional electronic medical record. Serum creatinine measurements were retrieved up to 2 years before admission. Comorbid conditions were identified using ICD-9 codes for diabetes mellitus, hypertension, coronary artery disease (CAD), cerebrovascular disease, congestive heart failure, and peripheral vascular disease assigned before admission (Appendix Table 1).

Missing Values Generation and Subcohort Selection

There are three general patterns of missing data (6,10). Data are missing completely at random if the probability of missing values does not depend on observed or unobserved variables, missing at random if the probability of missing values depends on observed variables, and missing not at random if the probability of missing values depends on some unobserved variables (6,10). Because the decision to measure serum creatinine is based on existing clinical information and is not completely at random, missing serum creatinine values are generally considered missing at random (12–14).

Our objective was to compare the accuracy of multiple imputation with that of eGFR 75 estimates of baseline creatinine against a known reference standard. To make this comparison, we required a dataset reflecting a missing-at-random pattern of serum creatinine values observed in clinical practice. We used a propensity score model in which the outcome of the model was whether baseline creatinine was missing (n=28,111 admissions) or not missing (n=13,003). We applied this model to the latter group of patients with known baseline creatinine data (n=13,003) to identify 50% of patients with the highest likelihood (n=6502 of 13,003) of missing baseline creatinine. Because each patient had a known baseline value, these initial steps provided a primary study cohort (n=6502) with known baseline creatinine data but labeled as "missing." It should be noted that these initial steps are not required to apply multiple imputation to a research cohort but were allowed for a reference standard reflecting a clinical missing-at-random pattern in this study.

The propensity model included the *a priori*–defined patient characteristics: sex, race, age, admission service type, congestive heart failure, hypertension, CKD, CAD, cerebrovascular disease, peripheral vascular disease, chronic liver disease, Charlson comorbidity index at discharge, distance of home residence from hospital, and length of stay. Clinical conditions were determined by inpatient and outpatient ICD-9 administrative codes. Age and length of stay were included as nonlinear terms using restricted cubic splines with four knots (15). Missing values of admission service type were imputed using multiple imputation, but these estimations were used only to generate propensity scores and were not retained for the subsequent primary multiple-imputation evaluation (6).

Multiple-Imputation Strategy

Four multiple-imputation strategies were studied:

1. Basic (sex, age, race).
2. Basic with serum creatinine (basic + minimum inpatient serum creatinine during the first 7 days of hospitalization).
3. Full multiple imputation (sex, race, dialysis, admission service, congestive heart failure, hypertension, CKD, diabetes mellitus, cerebrovascular disease, peripheral vascular disease, CAD, chronic liver disease, Charlson comorbidity index, distance of residence from hospital, length of stay).
4. Full multiple imputation with serum creatinine (full multiple imputation + minimum inpatient serum creatinine during the first 7 days of hospitalization).

Because potential users of multiple imputation may have access to varying amounts and quality of covariate data, the rationale for testing these different approaches was to provide performance information for users who may have to take such limitations into account and determine the incremental utility offered by more "data-intensive" multiple imputation.

Statistical Analyses

For each multiple-imputation strategy, values of baseline creatinine were imputed using linear regression adjusting for corresponding covariates mentioned in the previous section. All regressions used log-transformed creatinine. For each multiple-imputation strategy, serum creatinine was imputed 10 times using multiple imputation by chained equations (16), and values were averaged to obtain the final estimate.

The relative performance of each multiple imputation strategy and the eGFR 75 was calculated for 6502 patients marked as "missing baseline creatinine" but for whom baseline creatinine was actually available. We first compared the absolute mean difference of the actual and the imputed creatinine levels. For each method, the absolute difference between imputed and actual baseline creatinine was computed. Next, the mean and 95% nonparametric bootstrap confidence intervals (CIs), based on 1000 bootstrap replications (17), were compared with eGFR 75.

We also compared the ability of each method to accurately classify AKI. AKI was defined as a minimum serum creatinine of 0.3 mg/dl or a 50% increase in serum creatinine from the imputed baseline value to the peak serum creatinine value during the first 7 days after hospitalization. The reference standard AKI was defined as a minimum serum creatinine of 0.3 mg/dl or a 50% increase from the true baseline serum creatinine value to the peak serum creatinine value during the first 7 days after hospitalization. We further compared sensitivity, specificity, positive predictive value, negative predictive value, and misclassification rates across methods. Acute Kidney Injury Network (AKIN) staging was applied using creatinine criteria. Misclassification rate was calculated as the proportion of admissions that were incorrectly classified as AKI or non-AKI on the basis of the reference standard. Performance differences were also examined among different ranges of eGFR based on the minimum serum creatinine value during the first 7 days of hospitalization.

The rationale for doing this instead of calculating eGFR from the known baseline serum creatinine value was to provide performance data for potential users of this technique when baseline creatinine is unavailable. For each method, misclassification rates and 95% Wilson CIs were computed (18). Misclassification was compared using a McNemar test with Bonferroni correction. All statistical analyses were performed using statistical language R, version 2.12.1 (http://www.r-project.org/).

Results

Patient Characteristics

Our study population (n=6502) was selected from 13,003 adult patients with known baseline serum creatinine in order to simulate a liberal missing-data scenario of 50%. Admission data are listed in Table 1 and grouped by higher or lower propensity for missing baseline creatinine. The median age of patients in the study was 59 years (interquartile range, 46–69 years), with the highest reported comorbidity rates for hypertension (45%), CAD (24%), and diabetes mellitus (22%). In addition, 3377 (26%) patients had preadmission kidney dysfunction, as indicated by a baseline eGFR < 60 ml/min per 1.73 m^2. Overall, AKI was observed in 20.0% of the cohort, with 1843 (14%) patients experiencing AKIN stage I, 341 (3%) experiencing stage II,

and 356 (3%) experiencing stage III. Patients with a higher propensity for missing baseline creatinine were younger, were less likely to carry a reported diagnosis of all comorbid conditions studied, and experiencing AKI ($P<0.001$).

Comparison between Actual Baseline Creatinine Values with Multiple-Imputation Estimates

A total of 6502 patients with high propensity for missing baseline creatinine constituted the primary study cohort. The absolute differences between the estimated and known baseline creatinine and eGFR values with each estimation method are shown in Figure 1, A and B. Full multiple-imputation strategies resulted in smaller differences between estimated and actual baseline creatinine and eGFR than did either basic multiple-imputation strategies. All multiple-imputation strategies except the basic one resulted in smaller differences between estimated and actual creatinine and eGFR values than did those using the eGFR 75 approach.

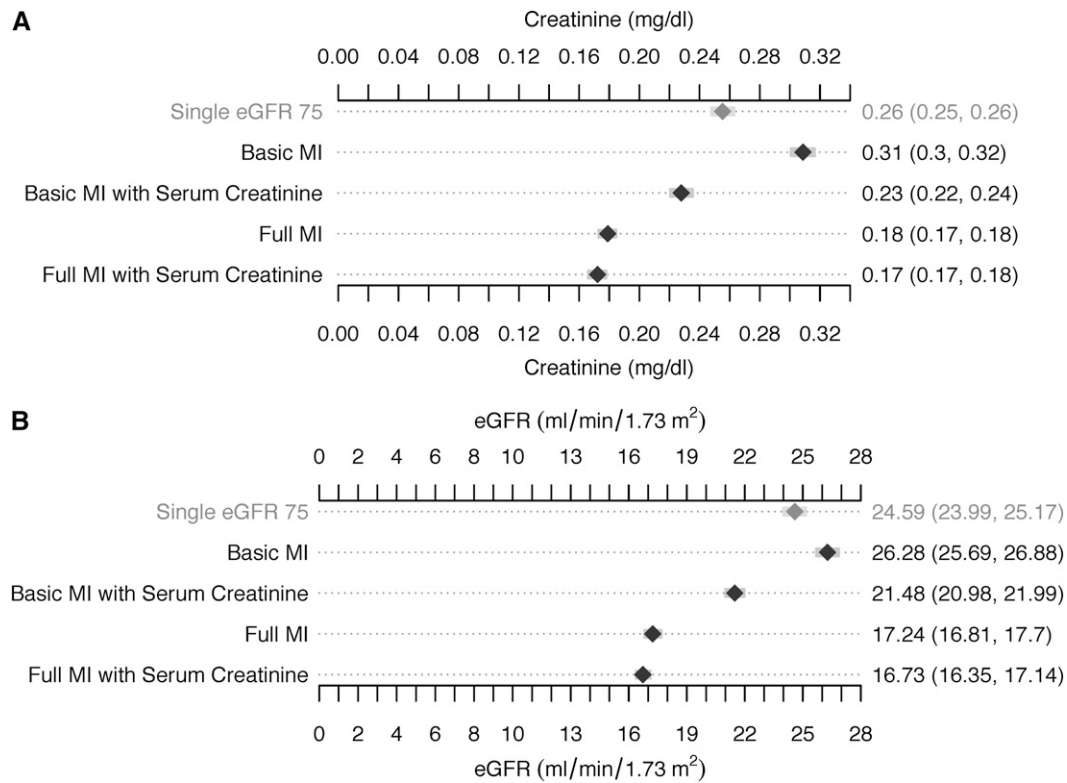### Relative Performances of Imputation Methods for Classification of AKI Status

Table 2 shows overall AKI misclassification rates observed with each strategy using a 0.3 mg/dl or 50% increase

Table 1. Demographic comparison of patients with higher versus lower likelihood of missing baseline serum creatinine data among those with known baseline creatinine data, as indicated by propensity score analysis

Baseline Characteristics	Higher Likelihood of Missing Baseline SCr (n=6502)	Lower Likelihood of Missing Baseline SCr (n=6501)	Total (n=13,003)
Median age (yr)	54 (40–67)	62 (53–71)	59 (46–69)
Men, n (%)	3250 (50)	2858 (44)	6108 (47)
Ethnicity, n (%)			
White	5388 (83)	5417 (83)	10,805 (83)
African-American	798 (12)	950 (15)	1,748 (13)
Other	316 (5)	134 (2)	450 (4)
Diabetes mellitus, n (%)	536 (8)	2334 (36)	2870 (22)
Hypertension, n (%)	1133 (17)	4721 (73)	5854 (45)
Coronary artery disease, n (%)	1070 (16)	2042 (31)	3112 (24)
Chronic heart failure, n (%)	403 (6)	1609 (25)	2012 (15)
Cerebrovascular disease, n (%)	587 (9)	1458 (22)	2045 (16)
Peripheral vascular disease, n (%)	162 (2)	887 (14)	1049 (8)
Chronic liver disease, n (%)	117 (2)	727 (11)	844 (6)
Charlson comorbidity index	0 (0–2)	6 (2–12)	2 (0–6)
Admission service, n (%)			
Medical	3339 (51)	4319 (66)	7658 (59)
Surgical	2853 (44)	1924 (30)	4777 (37)
Other/mixed	310 (5)	258 (4)	568 (4)
Home distance from hospital (miles)	40 (15–100)	25 (10–70)	35 (10–80)
Length of hospital stay (d)[a]	3.3 (2.1–6.0)	3.2 (2.1–5.3)	3.3 (2.1–5.7)
Baseline creatinine (mg/dl)	0.90 (0.76–1.10)	1.00 (0.80–1.30)	0.95 (0.80–1.20)
Baseline eGFR (ml/min per 1.73 m2)	83 (67–102)	70 (52–89)	77 (59–95)
AKI (%)	958 (15)	1582 (24)	2540 (20)
Died within 60 d (%)	259 (4)	658 (10.0)	917 (7)

All values are expressed as n (%) or median (interquartile range). $P<0.001$ for differences between missing baseline and available baseline serum creatinine groups for all characteristics listed except as noted. SCr, serum creatinine; eGFR, estimated GFR.
[a]P=0.05.

A

Creatinine (mg/dl)

0.00 0.04 0.08 0.12 0.16 0.20 0.24 0.28 0.32

Single eGFR 75	0.26 (0.25, 0.26)
Basic MI	0.31 (0.3, 0.32)
Basic MI with Serum Creatinine	0.23 (0.22, 0.24)
Full MI	0.18 (0.17, 0.18)
Full MI with Serum Creatinine	0.17 (0.17, 0.18)

0.00 0.04 0.08 0.12 0.16 0.20 0.24 0.28 0.32

Creatinine (mg/dl)

B

eGFR (ml/min/1.73 m^2)

0   2   4   6   8   10   13   16   19   22   25   28

Single eGFR 75	24.59 (23.99, 25.17)
Basic MI	26.28 (25.69, 26.88)
Basic MI with Serum Creatinine	21.48 (20.98, 21.99)
Full MI	17.24 (16.81, 17.7)
Full MI with Serum Creatinine	16.73 (16.35, 17.14)

0   2   4   6   8   10   13   16   19   22   25   28

eGFR (ml/min/1.73 m^2)

Figure 1. | Mean absolute difference between known baseline and estimated baseline serum creatinine (A) and estimated GFR (B) values. Values are shown as the mean and 95% nonparametric bootstrap confidence intervals, based on 1000 bootstrap replications. eGFR, estimated GFR; MI, multiple imputation.

Table 2.   Actual AKI misclassification rates using baseline serum creatinine estimation methods

Estimation Method[a]	Overall	eGFR < 60 ml/min per 1.73 m^2 (n=1118)	eGFR ≥ 60 ml/min per 1.73 m^2 (n=5384)
Single eGFR 75	12.3 (11.5–13.2)	40.5 (37.6–43.5)	6.5 (5.8–7.2)
Basic multiple imputation	10.0 (9.3–10.8)	19.1 (16.8–21.5)	8.1 (7.4–8.9)
Basic multiple imputation with SCr	9.3 (8.6–10.0)	13.9 (11.9–16.1)	8.3 (7.6–9.1)
Full multiple imputation	11.3 (10.6–12.1)	28.3 (25.7–31.0)	7.8 (7.1–8.6)
Full multiple imputation with SCr	9.0 (8.3–9.7)	15.3 (13.3–17.6)	7.7 (7.0–8.4)

Misclassification rates (%) with 95% Wilson confidence intervals for each method are shown for the overall group. Estimated GFR based on minimum serum creatinine during the first 7 days of hospitalization. All multiple-imputation groups were compared with single eGFR 75 as the reference group. All differences in misclassification were statistically significant ($P<0.001$) except for the full multiple imputation versus single eGFR 75 in the overall comparison ($P=0.06$). eGFR, estimated GFR; eGFR 75, eGFR = 75 ml/min per 1.73 m^2; SCr, serum creatinine.
[a]Basic multiple imputation = three-variable multiple imputation; basic multiple imputation with SCr = basic three-variable multiple imputation plus minimum inpatient serum creatinine; full multiple imputation = full-variable multiple imputation; full multiple imputation with SCr = full variable multiple imputation plus minimum inpatient serum creatinine.

---

between the estimated baseline and peak creatinine level during the first 7 days of hospitalization. Overall misclassification with multiple imputation was lower than with eGFR 75 and was lowest with multiple-imputation strategies that incorporated inpatient creatinine values. All comparisons were statistically significant ($P<0.001$) with the exception of full multiple imputation ($P=0.06$). Results were further stratified by eGFR calculated from baseline serum creatinine. Larger reductions in AKI misclassification

were notable for all multiple-imputation groups among patients with impaired eGFR (*i.e.*, <60 ml/min per 1.73 m^2) ($P<0.001$). In contrast, all multiple-imputation methods modestly increased misclassification compared with eGFR 75 in those with known baseline eGFR ≥ 60 ml/min per 1.73 m^2 ($P<0.001$).

Table 3 and Figure 2, A–H, demonstrate the test characteristics for diagnosing AKI with each estimation method. Overall, multiple imputation improved the specificity and

Table 3.   Performance characteristics for the diagnosis of AKI using estimation methods relative to known baseline: overall and in patients with eGFR < or ≥ 60 ml/min per 1.73 m^2

Estimation Method[a]	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value
Overall				
Single eGFR 75	0.70	0.91	0.57	0.95
Basic multiple imputation	0.57	0.96	0.70	0.93
Basic multiple imputation + SCr	0.55	0.97	0.76	0.93
Full multiple imputation	0.71	0.92	0.60	0.95
Full multiple imputation + SCr	0.57	0.97	0.76	0.93
Patients with eGFR < 60 ml/min per 1.73 m^2				
Single eGFR 75	1	0.49	0.34	1
Basic multiple imputation*	0.92	0.78	0.53	0.97
Basic multiple imputation + SCr	0.83	0.87	0.63	0.95
Full multiple imputation	0.95	0.66	0.42	0.98
Full multiple imputation + SCr	0.86	0.84	0.59	0.96
Patients with eGFR ≥ 60 ml/min per 1.73 m^2				
Single eGFR 75	0.61	0.99	0.87	0.94
Basic multiple imputation	0.46	0.99	0.88	0.92
Basic multiple imputation + SCr	0.45	0.99	0.87	0.92
Full multiple imputation	0.63	0.97	0.75	0.95
Full multiple imputation + SCr	0.47	0.99	0.91	0.92

Estimated GFR based on minimum serum creatinine during the first 7 days of hospitalization. eGFR, estimated GFR; eGFR 75, estimated GFR = 75 ml/min per 1.73 m^2; SCr, serum creatinine;.
[a]Basic multiple imputation = three-variable multiple imputation; basic multiple imputation with SCr = basic three-variable multiple imputation plus minimum inpatient serum creatinine; full multiple imputation = full-variable multiple imputation; full multiple imputation with SCr = full variable multiple imputation plus minimum inpatient serum creatinine.

positive predictive value of the estimated baseline creatinine value beyond eGFR 75. The largest improvements were seen with multiple-imputation strategies that incorporated minimum inpatient creatinine values. In contrast, use of multiple-imputation baselines tended to decrease sensitivity for AKI detection, particularly if minimal inpatient creatinine values were incorporated. Differences in the observed negative predictive value using multiple-imputation techniques were small.

Among patients with evidence of impaired eGFR during hospitalization (minimum eGFR < 60 ml/min per 1.73 m^2), eGFR 75 yielded the highest sensitivity (1.00) and negative predictive value (1.00) but the poorest specificity (0.49) and positive predictive value (0.34) for diagnosing AKI. In contrast, multiple-imputation approaches yielded sensitivities of 0.83–0.95, with higher specificities (0.66–0.87) and positive predictive values (0.42–0.63). Negative predictive values were high for all approaches. For patients with preserved eGFR during hospitalization (minimum inpatient eGFR ≥ 60 ml/min per 1.73 m^2), the sensitivity of all estimation methods for AKI was poor (≤0.63), while specificity and negative predictive values were uniformly ≥0.92. Positive predictive values were similar among estimation methods.

Supplemental Analysis

To examine the relative performance of multiple-imputation strategies across a varying range of missingness, we repeated our analysis for different proportions of patients with missing data (Table 4). The patients with missing baseline creatinine were selected using the same propensity score. In general, the accuracy of all methods decreased gradually as the amount of missing data increased, but the relative performance between methods was maintained across different proportions of missing data.

We also compared total misclassification of AKI stage between the actual and estimated baseline values (Supplemental Table 1). Total AKI stage misclassification rates for the eGFR 75 approach and full multiple imputation + serum creatinine were 15.2% (95% CI, 14.3%–16.0%) and 10.7% (95% CI, 10.0%–11.5%), respectively ($P<0.001$). If AKI was defined using more severe binary classification criteria (combined AKIN II and III versus AKIN 0 and I), the misclassification rates for eGFR and full multiple imputation + serum creatinine decreased to 3.5% (95% CI, 3.1%–4.0%) and 1.7% (95% CI, 1.4%–2.0%), respectively ($P<0.001$). In the subgroup of patients whose baseline eGFR was <60 ml/min per 1.73 m^2, misclassification of stage was higher using the eGFR 75 method (50.6% [95% CI, 47.7%–53.5%]) than with full multiple imputation + serum creatinine (17.1% [95% CI, 14.9%–19.2%]; $P<0.001$). When AKIN II and III were used to define injury, misclassification decreased but was still higher for eGFR 75 (14.0% [95% CI, 12.0%–16.1%]) than for full multiple imputation + serum creatinine (1.9% [95% CI, 1.1%–2.7%]; $P<0.001$).

Discussion

The diagnosis, staging, and study of AKI-related outcomes rely on quantifying changes in serum creatinine. Although widespread availability of electronic medical
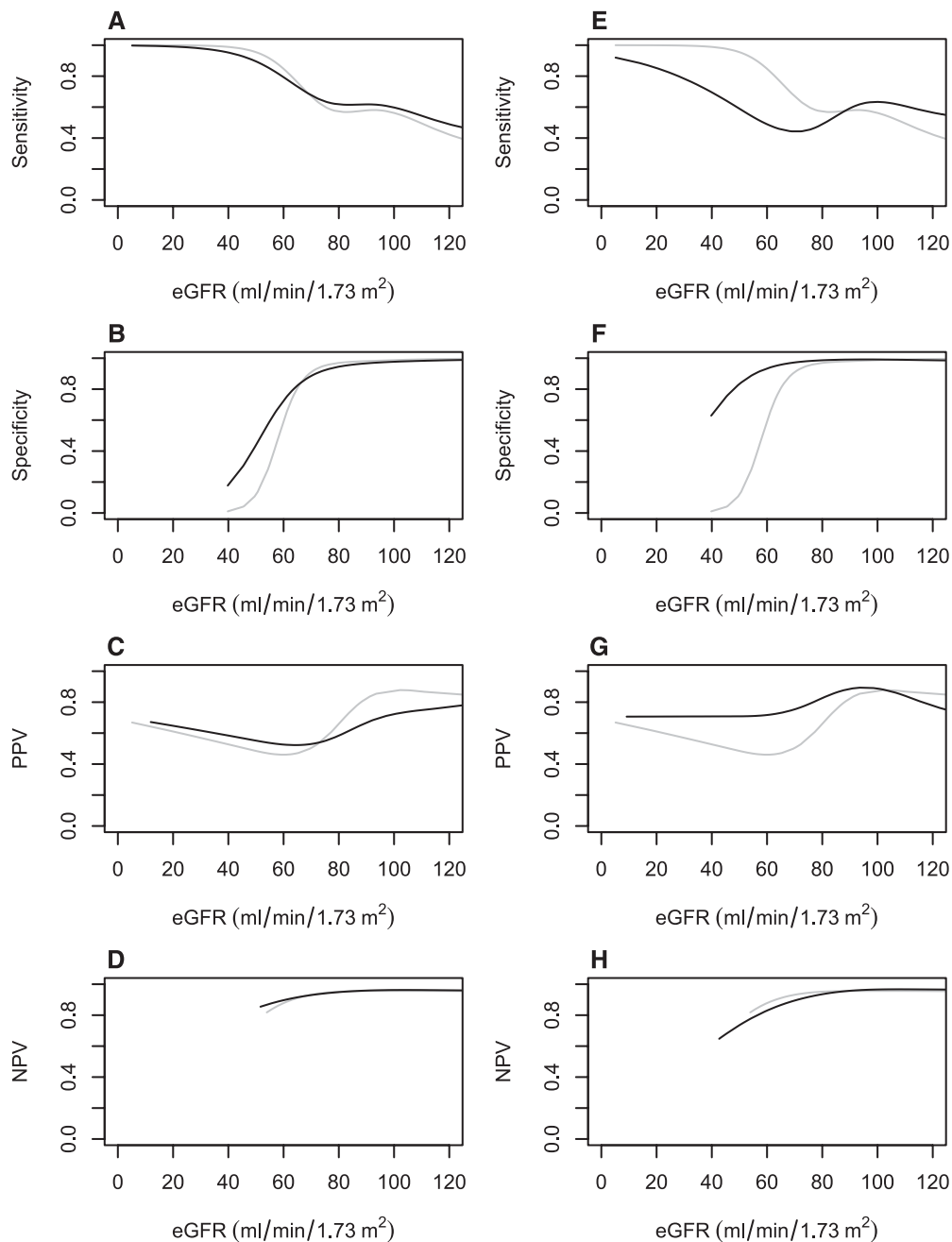
Figure 2. | **Sensitivity, specificity, and negative and positive predictive values.** (A–D) Full multiple imputation (dark) diagnosing AKI compared with the estimated GFR (eGFR) 75 ml/min per 1.73 m² approach (gray). Reference standard was calculated using each patient's known preadmission baseline serum creatinine. (E–H) Full multiple imputation + serum creatinine (dark) for diagnosing AKI compared with the eGFR 75 approach (gray). Reference standard was calculated using each patient's known preadmission baseline serum creatinine. Y-axes indicate sensitivity, specificity, positive predictive value, and negative predictive value. X-axes are based on the eGFR levels using the minimum inpatient serum creatinine during the first 7 days of hospitalization. NPV, negative predictive value; PPV, positive predictive value.

records has encouraged the study of this disease, preadmission kidney function is frequently missing, prompting the use of surrogate estimates. The accuracy of these estimates and the proportion of patients with missing data can significantly affect study quality (5). These results demonstrate that multiple imputation can outperform previously recommended strategies for estimating baseline kidney function and serve as a useful research tool.

Consensus definitions of AKI have standardized its study across settings (19,20). To reduce selection bias, replacing missing values with moderately preserved eGFR of 75 ml/min per 1.73 m² rather than excluding patients with missing data is reasonable (11). However, the advantages of simplicity and feasibility come at the risk for increasing misclassification, particularly among patients with CKD (*i.e.*, false-positive results), who are at highest

Table 4.   AKI misclassification rates of estimated GFR of 75 ml/min per 1.73 m² and full multiple imputation plus serum creatinine when proportion of patients with missing data is varied

Model	Proportion of Patients with Missing Data				
	10%	20%	40%	60%	80%
AKI misclassification					
eGFR 75	11.1(9.5–12.9)	10.7(9.5–12.0)	11.7(10.9–12.6)	12.7(12.0–13.5)	14.4(13.7–15.1)
Full multiple imputation + SCr	8.0(6.6–9.6)	8.7(7.6–9.8)	8.7(7.9–9.5)	9.2(8.6–9.8)	10.0(9.5–10.6)
AKI misclassification (eGFR < 60 ml/min per 1.73 m²)					
eGFR 75	44.5(36.4–53.0)	43.0(37.2–49.0)	40.8(37.4–44.3)	40.5(37.9–43.0)	41.4(39.3–43.2)
Full multiple imputation + SCr	14.4(9.3–21.4)	17.5(13.4–22.5)	14.9(12.5–17.5)	15.5(13.7–17.5)	16.1(14.6–17.7)
AKI misclassification (eGFR ≥ 60 ml/min per 1.73 m²)					
eGFR 75	6.9(5.5–8.5)	6.7(5.7–7.8)	6.4(5.7–7.1)	6.5(5.9–7.2)	7.0(6.4–7.5)
Full multiple imputation + SCr	7.2(5.8–8.9)	7.6(6.5–8.7)	7.5(6.8–8.4)	7.8(7.1–8.5)	8.3(7.8–9.0)

Rates shown are for within the proportion of patients with missing data. eGFR, estimated GFR; eGFR 75, estimated GFR = 75 ml/min per 1.73 m²; SCr, serum creatinine.

risk for AKI. Because studies focused on examining the predictors and outcomes of AKI may be enriched with these patients, unintended consequences may include overestimating disease incidence and severity while diluting or inflating associations between AKI and outcomes of interest (5).

A more insidious problem limiting the eGFR 75 approach is the assumption that missing baseline creatinine data are unrelated to observed patient characteristics. Our analysis identified substantial differences in characteristics among patients likely to be missing baseline serum creatinine data. The ability of multiple imputation to consider both the effect and uncertainty of these variables may explain its superior performance compared with eGFR 75. Further, the limitations of the latter may become more apparent when small changes in serum creatinine within current consensus definitions are used (21). Using a percentage change, or more severe injury stages, as our sensitivity analysis demonstrates, may help improve specificity and minimize false-positive AKI assessments. More recently, consensus definitions incorporated a rolling 48-hour window of detection for AKI. While helping to ensure the acuity of injury and "reduce the need for baseline" (3), this approach greatly diminishes the sensitivity to AKI occurring before admission or underestimating disease severity.

Our prespecified analysis assumed a fairly large proportion of missing data (*i.e.*, 50%) to reflect a liberal missing-data scenario. Worth noting is that the eGFR 75 approach has been applied to studies with larger ranges of missing data (22–24). Our sensitivity analyses revealed that multiple imputation consistently outperformed eGFR 75, particularly among those with baseline eGFR values < 60 ml/min per 1.73 m² across a broad spectrum of missing data. Further, only moderate decreases in performance were observed as the proportion of missing data increased,

consistent with previous simulations (25). However, this should not necessarily justify applying multiple imputation to any dataset with large proportions of missing values. Despite a relatively stable proportion of patients misclassified among those with missing data, the number of patients misclassified will, by definition, increase as this technique is applied to increasingly larger percentages of the population. Results may also differ considerably in cohorts in which the absolute number of patients with known data are small or if the reason for missing data stems from unobserved confounders for the outcome studied (10). Last, it should be mentioned that multiple imputation is by nature an estimation method for research use and has not been adapted for guiding clinical decision-making or trial enrollment on an individual-patient level.

Recognizing that no estimation method is failsafe, the optimal approach may depend on the underlying study priority and distribution of kidney function. Using an eGFR of 75 ml/min per 1.73 m², for example, assigns relatively preserved kidney function. Consequently, the higher sensitivity for detecting AKI compared with multiple imputation is not surprising, particularly among patients with a lower eGFR (*e.g.*, <60 ml/min per 1.73 m²). If a primary goal is to screen a population for all patients with potential AKI for evaluation, this approach may be reasonable because it tends to maximize sensitivity and negative predictive value with minimal effort. It may also be reasonable in low-risk groups, such as young cohorts without known comorbid conditions. In a subgroup of patients younger than age 50 years without known diabetes mellitus, hypertension, or cerebrovascular disease, eGFR 75 and multiple imputation resulted in similar AKI misclassification rates: 7.1% (95% CI, 6.1%–8.3%) and 7.8% (95% CI, 6.8%–9.1%), respectively. However, it is likely that any AKI-centered study would contain a substantial

Table A1. Diagnostic codes for comorbid conditions

Diagnosis	Definition
Hypertension	ICD-9 outpatient or inpatient discharge codes: 401*, 402*, 403*, 404*, and 405*
Diabetes mellitus	ICD-9 outpatient or inpatient discharge codes: 250* (primary diabetes), 249* (secondary diabetes), 357.2* (neuropathy in diabetes), 362* (diabetic retinopathy), 366.41* (diabetic cataract), 648* (diabetes in pregnancy)
Coronary artery disease	ICD-9 outpatient or inpatient discharge codes: 440*, 410*, 411*, 413*, 414*
Peripheral vascular disease	ICD-9 outpatient or inpatient discharge codes: 443*, 440.2*, 440.3*, 440.4*
Dialysis	Inpatient ICD-9 procedure codes: 12.55, 39.95*, 54.98* Outpatient and fee basis ICD-9 diagnosis codes: 585.6, V45.1, V56.0, V56.2, V56.3, V56.31, V56.32, V56.8, V58.8 Outpatient CPT codes: 90935, 90937, 90947, 90989, 90993, 90921, 90925
Congestive heart failure	ICD-9 outpatient or inpatient discharge codes: 428*, 425*, 398.91, 402.01, 402.11, 402.91, 404.91, 404.93
Hospice	ICD-9 outpatient or inpatient discharge codes: V66.7* Outpatient CPT codes: 99377
Renal transplant	ICD-9 outpatient or inpatient discharge codes: 55.6, 55.61, 55.69, 996.81, V42.0 ICD-9 inpatient procedure codes: 00.91, 00.92, 00.93 Outpatient CPT codes: 50380, 50365, 50360

* Includes all subcodes (*e.g.*, 401.01, 401.02, etc.). ICD-9, International Classification of Diseases, Ninth Revision; CPT, Current Procedural Terminology.

proportion of at-risk patients and eventually require more accurate assessments of baseline kidney function to better phenotype the degree of injury with high specificity or positive predictive value. Because patients with abnormal eGFR are among those at highest risk for AKI, the improved performance of multiple imputation may provide a more favorable balance than does eGFR 75 in populations enriched with these patients.

Strengths of the study include a large sample size, broad inclusion criteria, and availability of patient-related health information. There are also limitations. The pattern of missing data on baseline serum creatinine in this large, single academic center may not be generalizable to non-academic or nontertiary referral centers. Some centers may not have access to comorbidity data for the full multiple-imputation model and could use the more basic model. Although both methods reduced AKI misclassification beyond eGFR 75, the incremental improvements in AKI misclassification between basic and full multiple-imputation approaches were modest in this study. The extent to which this reflects limits of the comorbidity data or a marginal incremental benefit is not clear. Further research is needed to determine what patient characteristics provide the greatest utility in estimating baseline creatinine and the relative benefit of more data-intensive strategies in and beyond AKI classification, where the dichotomization of creatinine may mask some benefit of multiple imputation. Finally, we acknowledge the potential for residual bias from using propensity scores to predict patients with missing data. However, similar distributions in age and notable comorbid conditions, including diabetes and hypertension, were observed between those with actual missing baseline creatinine values and the 50% of patients with nonmissing baseline creatinine values used as the primary cohort (data not shown).

In summary, the use of smaller changes in serum creatinine to identify and stage AKI has placed a premium on accurate characterization of baseline kidney function. Imputation of a fixed baseline eGFR is less accurate than using multiple imputation to estimate baseline kidney function in patients with missing data. Rigorous methods to define baseline creatinine values in studies of AKI are critical to obtaining the proper incidence rates, facilitate comparisons between settings, and anchor robust longitudinal studies. Multiple imputation may serve as a useful tool that increases the accuracy of the data beyond conventional surrogates.

Disclosures
None.

References

1. Hsu CY, McCulloch CE, Fan D, Ordoñez JD, Chertow GM, Go AS: Community-based incidence of acute renal failure. *Kidney Int* 72: 208–212, 2007
2. Xue JL, Daniels F, Star RA, Kimmel PL, Eggers PW, Molitoris BA, Himmelfarb J, Collins AJ: Incidence and mortality of acute renal failure in Medicare beneficiaries, 1992 to 2001. *J Am Soc Nephrol* 17: 1135–1142, 2006

3. Mehta RL, Kellum JA, Shah SV, Molitoris BA, Ronco C, Warnock DG, Levin A; Acute Kidney Injury Network; : Acute Kidney Injury Network: Report of an initiative to improve outcomes in acute kidney injury. *Crit Care* 11: R31, 2007

4. Bagshaw SM, Uchino S, Cruz D, Bellomo R, Morimatsu H, Morgera S, Schetz M, Tan I, Bouman C, Macedo E, Gibney N, Tolwani A, Oudemans-van Straaten HM, Ronco C, Kellum JA; Beginning and Ending Supportive Therapy for the Kidney (BEST Kidney) Investigators; : A comparison of observed versus estimated baseline creatinine for determination of RIFLE class in patients with acute kidney injury. *Nephrol Dial Transplant* 24: 2739–2744, 2009

5. Siew ED, Matheny ME, Ikizler TA, Lewis JB, Miller RA, Waitman LR, Go AS, Parikh CR, Peterson JF: Commonly used surrogates for baseline renal function affect the classification and prognosis of acute kidney injury. *Kidney Int* 77: 536–542, 2010

6. Little RJA, Rubin DB: *Statistical Analysis with Missing Data*, Hoboken, NJ, John Wiley and Sons, 2002

7. Moons KG, Donders RA, Stijnen T, Harrell FE Jr: Using the outcome for imputation of missing predictor values was preferred. *J Clin Epidemiol* 59: 1092–1101, 2006

8. van der Heijden GJ, Donders AR, Stijnen T, Moons KG: Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: A clinical example. *J Clin Epidemiol* 59: 1102–1109, 2006

9. de Groot JA, Janssen KJ, Zwinderman AH, Moons KG, Reitsma JB: Multiple imputation to correct for partial verification bias revisited. *Stat Med* 27: 5880–5889, 2008

10. Donders AR, van der Heijden GJ, Stijnen T, Moons KG: Review: A gentle introduction to imputation of missing values. *J Clin Epidemiol* 59: 1087–1091, 2006

11. Bellomo R, Ronco C, Kellum JA, Mehta RL, Palevsky P; Acute Dialysis Quality Initiative workgroup; : Acute renal failure - definition, outcome measures, animal models, fluid therapy and information technology needs: The Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group. *Crit Care* 8: R204–R212, 2004

12. Rubin DB: Multiple imputation after 18+ years. *J Am Stat Assoc* 91: 473–489, 1996

13. Collins LM, Schafer JL, Kam CM: A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychol Methods* 6: 330–351, 2001

14. Schafer JL: Multiple imputation in multivariate problems when the imputation and analysis models differ. *Stat Neerl* 57: 19–35, 2003

15. Harrell F: *Regression Modeling Strategies*, New York, New York, Springer-Verlag, 2001

16. van Buuren S, Oudshoorn CGM: Multivariate Imputation by Chained Equations: MICE V1.0 User's manual. TNO Report PG/VGZ/00.38, Leiden, the Netherlands, TNO Prevention and Health, 2000

17. Davidson AC, Hinkley DV: *Bootstrap Methods and their Application*, New York, NY, Cambridge University Press, 2005

18. Wilson EB: Probable inference, the law of succession, and statistical inference. *J Am Stat Assoc* 22: 209–212, 1927

19. Ricci Z, Cruz D, Ronco C: The RIFLE criteria and mortality in acute kidney injury: A systematic review. *Kidney Int* 73: 538–546, 2008

20. Cruz DN, Bagshaw SM, Ronco C, Ricci Z: Acute kidney injury: classification and staging. *Contrib Nephrol* 164: 24–32, 2010

21. Waikar SS, Betensky RA, Emerson SC, Bonventre JV: Imperfect gold standards for kidney injury biomarker evaluation. *J Am Soc Nephrol* 23: 13–21, 2012

22. Hoste EA, Clermont G, Kersten A, Venkataraman R, Angus DC, De Bacquer D, Kellum JA: RIFLE criteria for acute kidney injury are associated with hospital mortality in critically ill patients: A cohort analysis. *Crit Care* 10: R73, 2006

23. Uchino S, Bellomo R, Goldsmith D, Bates S, Ronco C: An assessment of the RIFLE criteria for acute renal failure in hospitalized patients. *Crit Care Med* 34: 1913–1917, 2006

24. Bagshaw SM, George C, Bellomo R; ANZICS Database Management Committe; : A comparison of the RIFLE and AKIN criteria for acute kidney injury in critically ill patients. *Nephrol Dial Transplant* 23: 1569–1574, 2008

25. Janssen KJ, Donders AR, Harrell FE Jr, Vergouwe Y, Chen Q, Grobbee DE, Moons KG: Missing covariate data in medical research: To impute is better than to ignore. *J Clin Epidemiol* 63: 721–727, 2010