Rule Out Acute Kidney Injury in the Emergency Department With a Urinary Dipstick

Check for updates

Jacob S. Stevens[1,7], Katherine Xu[1,7], Alexa Corker[1], Tejashree S. Gopal[1], Osman R. Sayan[2], Erin P. Geraghty[3], Andrew M. Yaeh[3], Yaagnik D. Kosuri[4], John R. Burton[4], Saul V. Lincoln[4], Miriam P. Callahan[4], Rebecca K. Breheney[4], Andrew S. Beenken[1], Juliana N. Gamino[2], Ariel E. Felman[1], Anjali Gehani[1], Hayley A. Giordano[2], Aileen Gozali[2], Eddie F. Guerrero Herrera[1], Britney A. Hatcher[2], Lena A. Kheir[2], Yuanji Li[1], Erika K. Mitsui[2], Jae I. Nha[1], Alexander T. Sayan[1], Samuel J. Spaiser[2], Siddarth Arumugam[5], Samuel K. Sia[5], Kristen L. King[1], Sumit Mohan[1,6] and Jonathan Barasch[1]

[1]Division of Nephrology, Columbia University Irving Medical Center, New York, New York, USA; [2]Emergency Department, Columbia University Irving Medical Center, New York, New York, USA; [3]Internal Medicine Residency, Columbia University Irving Medical Center, New York, New York, USA; [4]Vagelos College of Physicians and Surgeons, Columbia University, New York, New York, USA; [5]Biomedical Engineering, Columbia University, New York, New York, USA; and [6]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York, USA

Introduction: The identification of acute injury of the kidney relies on serum creatinine (SCr), a functional marker with poor temporal resolution as well as limited sensitivity and specificity for cellular injury. In contrast, urinary biomarkers of kidney injury have the potential to detect cellular stress and damage in real time.

Methods: To detect the response of the kidney to injury, we have tested a lateral flow dipstick that measures a urinary protein called neutrophil gelatinase-associated lipocalin (NGAL). Analysis of urine was performed in a prospective cohort of 479 patients (final cohort $N = 426$) entering an emergency department in New York City and subsequently admitted for inpatient care.

Results: Colorimetric development had high interrater reliability (88% concordance rate) and correlated with traditional enzyme-linked immunosorbent assay (ELISA) measurements ($\rho = 0.732$, $P < .0001$). Of the 14% of the cohort who met Acute Kidney Injury Network (AKIN) SCr criteria for acute kidney injury (AKI), 67% demonstrated transient (<2 days) and 33% demonstrated sustained (>2 days) elevation of SCr. Comparing the outcomes of patients with sustained versus transient or undetectable changes in SCr revealed that the urinary NGAL (uNGAL) dipstick had high specificity and negative predictive value (NPV) (high- vs. low-intermediate readings, sensitivity = 0.55, specificity = 0.91, positive predictive value = 0.24, NPV = 0.97, $\chi^2 = 20.39$, $P < 0.001$).

Conclusion: We show that the introduction of a bedside uNGAL dipstick permits accurate triage by identifying individuals who do not have tubular injury. In an era of shortening length of stay and rapid decisions based on isolated SCr measurements, real-time exclusion of kidney injury by a dipstick will be particularly useful to overcome the retrospective, insensitive, and nonspecific attributes of SCr.

The current diagnosis of kidney disease relies on SCr, a marker of excretory dysfunction. SCr is also thought to quantify kidney tubular injury. Yet, the intrinsic characteristics of SCr preclude it from being an immediate and proximate assay of damaged kidney cells.

A delay of >1 to 2 days between the onset of excretory dysfunction and the accumulation of SCr to a diagnostic threshold (e.g., Risk, Injury, Failure, Loss, and End-stage renal disease; AKIN; and Kidney Disease: Improving Global Outcomes "stages") limits SCr to the retrospective detection of renal dysfunction. The delay constitutes a major barrier to real-time diagnoses, which impacts clinical care, particularly in the setting

of rapid patient turnover. In addition, changes in SCr can only be detected when a substantial number of kidney tubules have been injured, given the presence of redundancy and compensation called "renal reserve."[1] This phenomenon suppresses changes in SCr despite the presence of injury, a major limitation in diagnostic sensitivity called "subclinical AKI."[2,3] Moreover, low muscle mass[4,5] and concurrent illness further reduce the test's sensitivity because SCr is a muscle metabolite subject to metabolic modulation.[6] These problems underscore the difficulties of using SCr to detect patchy and focal injuries typical of human kidney damage.[7,8] Finally, because SCr does not derive from damaged tubular cells, it cannot be used as a sole test to distinguish intrarenal from extrarenal defects, a major limitation in diagnostic specificity.[9,10]

Repeated measurements of SCr may be a useful surrogate of tubular damage.[11–13] This is because tubular injury might prolong the duration of elevated SCr due to cell cycle arrest and tubular obstruction.[14–18] In contrast, commonly observed extrarenal hemodynamic causes of elevated SCr may rapidly reverse.[13,19,20] However, repeated measurements are not practical in acute patient triage.

Rather than rely solely on SCr as a surrogate of tubular injury, investigators have identified a series of urinary proteins[21] that correlate with tubular damage called the "biomarkers of kidney injury." Here we examine uNGAL (23–25 kDa "monomer")[7] because it is rapidly transcribed and secreted by the nephron,[22,23] as demonstrated by RNA *in situ* in different types of acute injury in humans (T. Shen, K. Xu, and J. Barasch, unpublished data, 2020) and mice[22] and by gene knockouts in mice. In addition, uNGAL correlates with the timing and severity of the renal stressor.[22,24–28] Moreover, ischemic, toxic, and septic stimuli rather than hemodynamic deficiencies (e.g., volume depletion, heart and liver failure, and diuretics) induce NGAL monomer expression,[27–33] probably through nuclear factor κB, nuclear factor erythroid 2–related factor 2, and hypoxia-inducible factor "damage" signaling. In contrast, less is known about the serum form (molecular mass > 100 kDa, cross-linked partners),[34,35] and measurement probably requires different point-of-care measurement tools.

Despite the many useful characteristics of NGAL and the other biomarkers, their clinical application has lagged because of the lengthy turnaround time, burdensome and costly materials, and difficulty positioning the test in oversubscribed or resource-limited settings. Recently, RenaStick (kidney injury molecule-1)[36] and Nephrocheck (Astute, San Diego, CA; tissue inhibitor of metalloproteinases-2 and insulin-like growth factor–binding protein 7)[37–40] have been
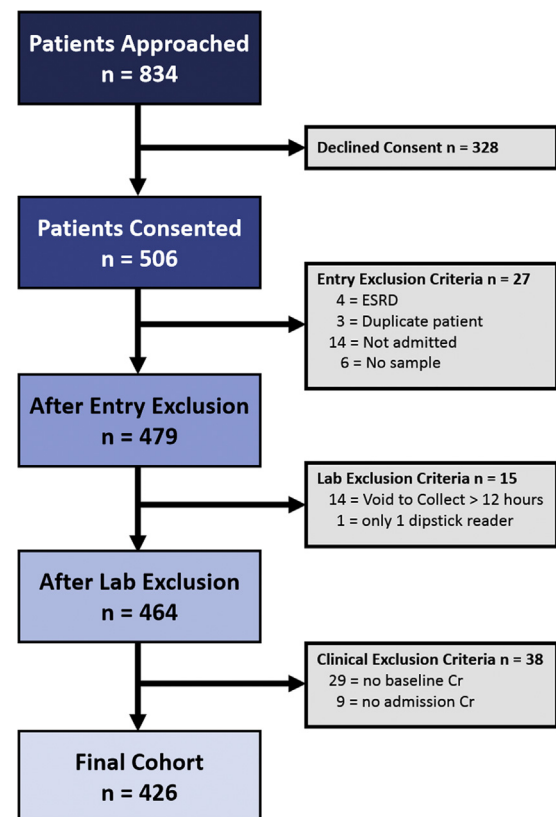

Figure 1. Patient enrollment from the emergency department. Eight hundred thirty-four patients were initially approached in the emergency department, but approximately 50% of these patients were not part of the final cohort for the following reasons: 39% declined to participate, whereas an additional 10% were excluded because they did not meet entry criteria or they failed to meet minimum necessary laboratory or clinical data standards. The final cohort included 426 patients. Cr, creatinine; ESRD, end-stage renal disease.

proposed as bedside point-of-care tools, whereas the US Army highlighted the use of a lateral-flow NGAL dipstick.[21] Here we test the ability of an uNGAL dipstick to help acutely triage patients in the emergency department (ED), a setting that does not have the luxury of time for repeated SCr measures.

METHODS

Enrollment

Recruiters identified all patients ≥18 years old in the ED processed for admission to New York Presbyterian Hospital at the Columbia University Irving Medical Center from June 2017 to January 2019. Informed consent was obtained from all enrolled participants. We excluded patients who (i) could not consent, (ii) were anuric at presentation, or (iii) had end-stage renal disease (Figure 1). Urine samples were collected within 24 hours of arrival in the ED and processed within 12 hours of voiding in order to correlate uNGAL at presentation with prospective changes in SCr over the

subsequent days. The research protocol was approved by the Columbia University Insitutional Review Board.

Baseline SCr Determination

A patient's baseline SCr was determined in order of preference as follows:

1. Median SCr from 365 to 7 days before presentation ($n = 285$, 66.9%). If none available, then
2. Minimum SCr from 7 days before presentation to the day of presentation ($n = 10$, 2.4%). If none available, then
3. Minimum SCr from presentation to discharge ("admission nadir," $n = 131$, 30.8%).
4. If no SCr or only a single SCr was measure\d during the index hospitalization without prior measurements, these patients were labeled as having an unknown baseline SCr ($n = 38$, excluded from the study).

The baseline estimated glomerular filtration rate (eGFR) was calculated using the Chronic Kidney Disease Epidemiology Collaboration equation.[41] Patients with a baseline eGFR < 30 ml/min per 1.73 m^2 (baseline chronic kidney disease [CKD], stage 4–5) were excluded in the secondary analyses.[34,42]

AKI Stratification

Because urine output was not available for the majority of our cohort, loss of excretory function was determined solely by SCr kinetics according to AKIN definitions,[43] interpreted as an absolute increase in SCr ≥ 0.3 mg/dl or a ≥50% increase from baseline, and staged as follows:

1. AKIN stage 1: ≥ 0.3-mg/dl increase in SCr within a 48-hour window or 1.5- to 2-fold increase in SCr compared with baseline
2. AKIN stage 2: >2- to 3-fold increase in SCr compared with baseline
3. AKIN stage 3: ≥ 0.5-mg/dl increase in SCr within a 48-hour window when SCr ≥ 4.0 mg/dl or >3-fold increase in SCr compared with baseline

Patients without measurements of SCr during hospitalization were excluded from the analysis.

The first SCr of each 24-hour period was used in our analysis. In select cases, the day 1 AKIN score was imputed when the preceding and subsequent AKIN scores were identical. Further categorization was based on the duration of an elevated SCr as follows:

1. "No AKI": when a patient did not meet AKIN criteria within 2 days of presentation (must have SCr values for both days)
2. "Transient AKI" (tAKI): when a patient met AKIN criteria on day 0 or day 1 of presentation but

normalized below AKIN detection thresholds within 2 days after first detection
3. "Sustained AKI" (sAKI): when a patient met AKIN criteria within 2 days of presentation but normalized below the AKIN detection thresholds >2 days from the first detection
4. "Unknown": when a patient was either discharged with insufficient measurements to determine SCr kinetics or had missing measurements on day 0 or day 1 that could not be imputed because of discrepant AKIN scores

Development of an Automated AKIN Scoring Algorithm

To ensure consistent temporal relationships between biomarkers of injury and excretory dysfunction, we developed a computer-based algorithm that determined baseline SCr and identified and classified changes in SCr by AKIN stages. The tool used variable detection windows and color-coded readouts to display SCr kinetics and AKIN scoring (Figure 2a).[28]

Urinary Tract Infection

Urinary tract infection (UTI) was defined as either urinalysis leukocyte esterase (LE) 2+ or 3+ and any colony-forming units or LE 1+ and >10,000 colony-forming units.[27,44] Subjects were considered to be uninfected if LE was negative or trace, regardless of the colony-forming units, or if LE = 1+ but colony-forming units were <10,000. Subjects were considered "unknown" if LE was not measured or LE = 1+ without urine cultures.

Laboratory Measurements

The lateral flow dipstick recognized the human uNGAL monomer (Figure 2b) in a graded fashion (∼6 ng/ml to ∼1000 ng/ml, Figure 2c).

uNGAL Dipstick Measurements

Urine was centrifuged (12,000 rpm, 10 minutes) and 10 μl applied to NGAL gRAD dipsticks (BioPorto Diagnostics, Gentofte, Denmark). Color development over 15 minutes was compared with the manufacturer's test line by 2 independent, blinded readers and associated with low, inconclusive, moderate, and high risk of kidney damage (Supplementary Figure S1A). These scores were transformed into 3 risk categories: low, intermediate, and high (Figure 2c and Supplementary Figure S1B), which approximated published values (low = 60 ± 147 ng/ml, intermediate = 114 ± 96 ng/ml, and high = 963 ± 1083 ng/ml; mean ± SD; Figure 2d) measured by ELISA.

Validation of uNGAL

The uNGAL dipstick was validated by (i) cross comparison with known quantities of NGAL in spiked
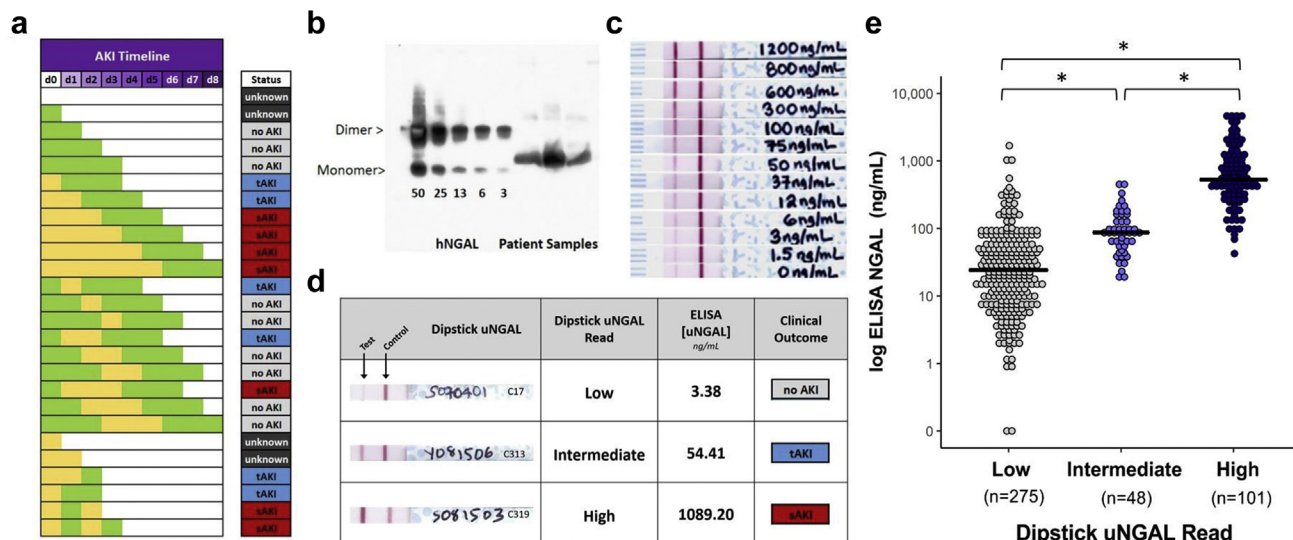
Figure 2. Development of analytical tools: clinical algorithm for characterization of creatinine kinetics and measurement of urinary neutrophil gelatinase-associated lipocalin (uNGAL) by dipstick. (a) Comparison of serum creatinine (SCr) kinetics and clinical outcomes. Yellow depicts days when the level of SCr met the Acute Kidney Injury Network (AKIN) criteria. Green depicts days when SCr levels did not meet the AKIN criteria. White depicts days when SCr was not measured. Urine was collected at presentation (day 0, d0), and, therefore, SCr trends between days 0 and 2 were critical to determine AKIN scoring (see Methods for AKI Stratification). The automated algorithm adjudicated these categories: no AKI (light gray), transient AKI resolving within 48 hours of detection (blue) (e.g., if met criteria on day 0, then normalized by day 2; and, if met criteria on day 1, then normalized by day 3), and sustained AKI persisting beyond 48 hours of detection (red) (e.g., if met criteria on day 0, then either did not normalize or only normalized ≥day 3; if met criteria on day 1, then either did not normalize or only normalized ≥day 4). Unknown diagnoses are represented in dark gray. tAKI, transient AKI; sAKI, sustained AKI; unknown=insufficient data to make categorization. (b) Anti-NGAL recognizes the glycosylated NGAL gene product (∼22 kDa) and the glycosylated dimer (∼44 kDa). Note nonglycosylated recombinant human NGAL (∼20KDa). (c) The same NGAL antibody was used for the lateral-flow dipstick. Human urine was spiked with increasing amounts of recombinant human NGAL. Note the increasing density of the test line. (d) Representative uNGAL dipsticks; correlation with ELISA measurements. (e) A summary of the dipstick and ELISA measurements in the final cohort. The mean ELISA uNGAL values correlated with low, intermediate, and high dipstick values (*<0.001, *n* = 424).[28]

urine; (ii) ELISA (KIT036; BioPorto Diagnostics, Hellerup, Denmark; Figures 2d and 3), and (iii) nonreducing immunoblot measuring the NGAL monomer (the gene product) with 4% to 15% sodium dodecylsulfate–polyacrylamide gel electrophoresis (Bio-Rad Laboratories, Hercules, CA), monoclonal anti-NGAL (1:1000; Enzo Life Sciences, Farmingdale, NY), and goat polyclonal secondary antibodies (1:5000; Jackson Immuno-Research, West Grove, PA). The nonreducing immunoblot rules out that NGAL measurements derive from high-molecular-weight species, which includes NGAL/matrix metalloproteinase-9.

Statistics

Continuous variables were compared using a 2-sample *t*-test or analysis of variance with pair-wise tests incorporating a Bonferroni multiple comparisons adjustment. The mean and SD are reported as mean ± SD. Non-normally distributed continuous variables were compared using the Wilcoxon rank-sum test for 2 groups or the Kruskal-Wallis test with Dunn's test using Bonferroni multiple comparison adjustment for more than 2 groups and are presented as the median (interquartile range). Categorical variables were compared using the χ^2 or Fisher's exact test. AKIN stage, uNGAL dipstick category, and CKD stage were analyzed as ordinal variables using a nonparametric test for trend and ordinal logistic regression.[45] The odds of reaching the composite end point of in-hospital mortality or renal replacement therapy were compared by the uNGAL dipstick category using logistic regression. The null hypothesis was rejected at $P < 0.05$. Statistical analyses were performed with Stata MP 15.1 (StataCorp, College Station, TX) with the dunntest package.

RESULTS

Patient Enrollment and Characteristics

Of the 834 patients approached, a total of 506 patients consented to participate, and 426 were included in our final cohort (Figure 1). The baseline characteristics of the entire consented cohort, those excluded, the final study cohort, and the cohort stratified by AKI status are presented in Supplementary Tables S1 and S2, respectively. The clinical outcomes of our patient cohort, those excluded, and the final study cohort are presented in Supplementary Table S3.
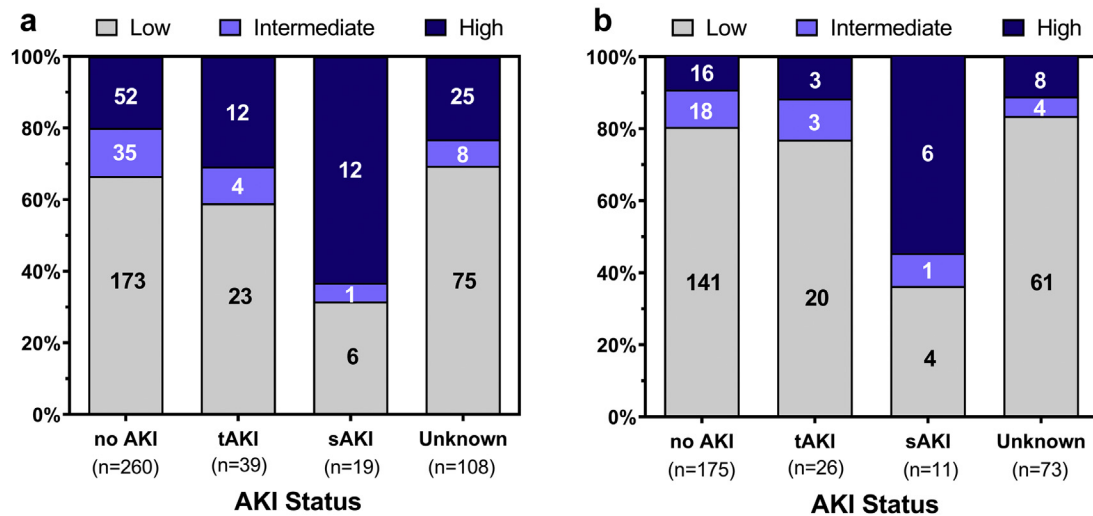
Figure 3. Urinary neutrophil gelatinase-associated lipocalin (uNGAL) measured by dipstick correlates with sustained elevations in creatinine. (a) High uNGAL readings in the emergency department correlate with sustained serum creatinine elevation after admission ($P = 0.004$, $n = 426$). (b) The relationship was even more evident once confounders (baseline estimated glomerular filtration rate < 30 or positive urinary tract infection or unknown urinary tract infection status) were excluded ($P = 0.008$, $n = 285$). Light gray = low uNGAL, dark gray = intermediate uNGAL, and black = high uNGAL. AKI, acute kidney injury; DPI, dots per inch; sAKI, sustained acute kidney injury; tAKI, transient acute kidney injury; unknown, insufficient data to make categorization.

The included cohort ($N = 426$) was diverse; 34% identified as white, 18% as black, 4% as other (4%), and 44% as unknown (44%, Table 1), including 24% identifying as Hispanic. The most prevalent comorbidities were hypertension (61%), diabetes mellitus (36%), CKD (stages 3–5, 31%), and congestive heart failure (24%), and the most common admission diagnoses were infection or fever (19%), dyspnea (12%), chest pain (7%), abdominal pain (7%), and electrolyte abnormalities (6%, Table 1). A majority of patients (91%) were admitted to the medicine service, and the most common admission locations were floor/ward (83%), intensive care unit (11%), and step-down units (6%). The median baseline SCr was 0.97 mg/dl (0.71–1.30 mg/dl) corresponding to a median baseline eGFR of 77 ml/min (55–99 ml/min, Table 2). The overall in-hospital mortality was 4.5%, and the 90-day mortality was 10%. Of the survivors to discharge, 22% (90/407) patients were readmitted within 30 days (Table 2).

Using the automated algorithm for AKIN determination by characterizing SCr kinetics (Figure 2a), 14% of the cohort met the SCr criteria for AKI. Of these, 67% were tAKI (transient elevation of SCr), and 33% were sAKI (sustained elevation in SCr). Most cases were AKIN stage 1 (50%), 26% were AKIN 2, and 24% were AKIN 3. Only 11% (47/426) had a nephrology consultation during their admission, and 2% required renal replacement therapy; 0.7% required renal replacement therapy at the time of discharge. The majority of patients (61%) in our study did not experience SCr criteria for AKI. An additional 25% had insufficient SCr data to determine their AKI status ("Unknown," Table 2).

Evaluating the Performance of the uNGAL Dipstick With SCr Kinetics

The uNGAL dipstick had high interrater reliability with a concordance rate of 88% between reader 1 and 2. NGAL concentrations measured by the dipstick also correlated with ELISA (Spearman correlation coefficient = 0.732, $P < 0.0001$; Figure 2e). The uNGAL-ELISA concentrations in the high dipstick group were significantly different from those in the intermediate and low dipstick groups (963 ± 1083 vs. 114 ± 96.3 or 60.3 ± 147.1 ng/ml, $P < 0.001$, respectively), whereas the concentrations in the intermediate and low groups were similar (114 ± 96.3 vs. 60.3 ± 147.1, $P = 1.0$; Figure 2e).

The exclusion of patients with intermediate dipstick values demonstrated that low and high uNGAL dipstick readings (Table 3, row A) distinguished patients with sustained SCr elevation (>2 days, sAKI) from patients with transient elevations (<2 days, tAKI) or no AKI with high NPV (0.97) and specificity (0.75). In fact, 63% of patients with sAKI had high uNGAL dipstick readings compared with 28% of patients with tAKI and 20% of patients without AKI (Figure 3a). Given the similarities in NGAL concentrations in the low and intermediate groups, they were combined, resulting in a slight decrease in sensitivity (0.63 vs. 0.67) but improved specificity (0.79 vs. 0.75; Table 3, row C).

Next, we examined potential confounders of test performance.[42,44] More severe CKD was associated with a significantly higher concentration of NGAL as well as a higher frequency of high uNGAL dipstick readings

Table 1. Patient characteristics of the final cohort

Patient characteristics	Patients	% Total
Age (yr)[a]	60.3	±17.9
Sex		
Female	182	42.7
Male	244	57.3
Race		
White	145	34.0
Black	77	18.1
Asian	11	2.6
American Indian	5	1.2
Hawaiian or Pacific Islander	2	0.5
Unknown	186	43.7
Ethnicity		
Hispanic	104	24.4
Non-Hispanic	94	22.1
Unknown	228	53.5
Admit diagnosis		
Infection or fever	76	17.8
Dyspnea	50	11.7
Abdominal pain	30	7.0
Chest pain	20	4.7
Electrolyte abnormality	21	4.9
Hematologic or oncologic	15	3.5
Pain	13	3.1
Syncope	9	2.1
Altered mental status	8	1.9
Acute kidney injury	8	1.9
Gastrointestinal bleed	7	1.6
Other	113	26.5
Missing	56	13.2
Admit location		
Floor	355	83.3
Step-down unit	24	5.6
Intensive care unit	47	11.0
Admit service		
Medicine	388	91.1
Surgery	18	4.2
Other	20	4.7
Hospitalization LOS[b]	4	(2–7)

LOS, length of stay.
[a]Presented as meant ± SD.
[b]Presented as median (interquartile range).
N = 426.

Table 2. Creatinine kinetics, acute kidney injury (AKI) definitions, and patient outcomes of the final cohort

Laboratory and diagnostic metrics	Patients	% Total
Baseline SCr[a]	0.97	(0.71–1.30)
Baseline eGFR[a]	77	(55–99)
Baseline SCr type		
−365 to −7 days	285	66.9
−7 days to admit	10	2.4
Admission nadir	131	30.8
No baseline	0	
Baseline CKD stage		
No CKD or CKD 1	151	35.5
CKD 2	130	30.5
CKD 3	105	24.7
CKD 4	32	7.5
CKD 5	8	1.9
unknown	0	0.0
Presentation SCr[a]	1.06	(0.78–1.53)
AKI status		
No AKI	260	61.0
tAKI	39	9.2
sAKI	19	4.5
Unknown	108	25.4
AKIN stage		
No AKI	260	61.0
AKIN 1	29	6.8
AKIN 2	15	3.5
AKIN 3	14	3.3
Unknown	108	25.4
AKI duration[a]	1.5	(1–2)
Renal consulted	47	11.0
RRT initiation	9	2.1
RRT at discharge	3	0.7
90-day mortality	42	9.9
In-hospital mortality	19	4.5
30-day readmission	90	22.1
ICU transfer	12	3.2

AKIN, Acute Kidney Injury Network; CKD, chronic kidney disease; eGFR, estimated glomerular filtration; ICU, intensive care unit; RRT, renal replacement therapy; sAKI, sustained acute kidney injury; SCr, serum creatinine; tAKI, transient acute kidney injury.
[a]Presented as median (interquartile range).
N = 426, except for AKI duration (n = 58), 30-day readmission (n = 407, excludes patients with in-hospital death), and ICU transfer (n = 379, excludes patients admitted to the ICU).

(Supplementary Figure S2A and B, respectively). Similarly, patients with a UTI had a higher concentration of NGAL (752.5 ± 953.4 vs. 169.5 ± 543.4 ng/ml, $P < 0.001$) and a higher frequency of high uNGAL dipstick readings compared with patients without a UTI (65% vs. 13%, odds ratio 14.34 [range 8.26–24.90], $P < 0.001$; Supplementary Figure S3A and B, respectively). When comparing low or intermediate to high uNGAL dipstick readings, the exclusion of individuals with evidence of CKD and/or UTI resulted in an improvement in the specificity (0.91), positive predictive value (0.24), and NPV (0.97) along with a decrease in the sensitivity (0.55; Table 3, row D).

To determine whether excluding confounders improved the power of NGAL to resolve no AKI,

tAKI, and sAKI, we compared uNGAL ELISA mean values. We found that both the tAKI and sAKI groups differed significantly from the no AKI group when all patients were included (Supplementary Figure S4A), but when confounders (baseline eGFR < 30 and/or +UTI or unknown UTI status) were excluded, the tAKI group no longer differed from the no AKI group, whereas the sAKI group differed from the no AKI and tAKI groups (Supplementary Figure S4B). In summary, the dichotomy of low or intermediate versus high uNGAL dipstick measurements produced the best performance in differentiating sAKI from no AKI and tAKI; the performance of the uNGAL dipstick was mildly improved when confounders were excluded (Figure 3b and Table 3, rows B and D; a full table of

Table 3. Test performance of the urinary neutrophil gelatinase-associated lipocalin dipstick in the final cohort

	Dipstick categories			AKI outcome			Confounder restrictions		n	Sensitivity	Specificity	PPV	NPV	Accuracy	Prevalence	+LR	−LR	χ^2	P value
							UTI status	Baseline eGFR											
A	Low	vs.	high	sAKI	vs.	no AKI or tAKI	Any	All	278	0.67	0.75	0.16	0.97	0.75	0.06	2.71	0.44	14.99	<.001
B	Low	vs.	high	sAKI	vs.	no AKI or tAKI	−UTI	eGFR > 30	190	0.60	0.89	0.24	0.98	0.88	0.05	5.68	0.45	20.27	<.001
C	Low or intermediate	vs.	high	sAKI	vs.	no AKI or tAKI	Any	All	318	0.63	0.79	0.16	0.97	0.78	0.06	2.95	0.47	17.12	<.001
D	Low or intermediate	vs.	high	sAKI	vs.	no AKI or tAKI	−UTI	eGFR > 30	212	0.55	0.91	0.24	0.97	0.89	0.05	5.77	0.50	20.39	<.001

AKI, acute kidney injury; eGFR, estimated glomerular filtration rate; +LR, positive likelihood ratio; −LR, negative likelihood ratio; NPV, negative predictive value; PPV, positive predictive value; sAKI, sustained acute kidney injury; tAKI, transient acute kidney injury; UTI, urinary tract infection.

available comparisons is available in Supplementary Table S4).

Comparing uNGAL Dipstick With Clinical Outcome

When we excluded the confounders of UTI and CKD, high uNGAL dipstick values correlated with the severity of AKI ($P = 0.039$; Figure 4a) and with composite outcomes of death and/or renal replacement therapy (high vs. low odds ratio 4.32 [range 1.19–15.65), $P = 0.026$; Figure 4b). Even among patients without SCr-defined AKI at presentation, high dipstick measurements were associated with an increased risk of either sAKI, in-hospital renal replacement therapy initiation, or death at 90 days compared with those with a low uNGAL dipstick value (odds ratio 2.68 [range 1.33–5.40], $P = 0.0057$).

DISCUSSION

Biomarkers of kidney damage provide an opportunity to identify kidney injury in real time rather than subsequent to SCr accumulation. Moreover, they may identify cases in which changes in SCr are suppressed by renal reserve. Additionally, in patients with elevated SCr, the biomarker may identify the subset with more severe tubular injury (e.g., sustained AKI) rather than hemodynamically reversible changes in SCr.[31,46,47]

The power and utility of the biomarkers might be enhanced by rapid assays at the bedside. Here we have evaluated a new bedside test for uNGAL that is able to provide a diagnosis within 15 minutes of obtaining a urine sample in a busy ED. We show that the bedside tool can provide insight into whether the patient's course will include a sustained elevation of SCr.

Using the uNGAL dipstick, validated by ELISA and immunoblot, we learned that (i) low uNGAL dipstick measurements were below previously established cut-offs ($\sim$104 ng/ml)[27] and were not associated with sustained changes in SCr (Figures 2e and 3, respectively); (ii) intermediate levels correlated with uNGAL ELISA values typically at or below the threshold for the diagnosis of sustained elevations in SCr (Figure 2e) and, consequently, could be grouped with low uNGAL-dipstick measurements; and (iii) high uNGAL dipstick measurements were associated with sustained elevations in SCr (Figure 3). Consequently, we recommend a simplified clinical algorithm (combining low and intermediate reads) to streamline the use of the uNGAL dipstick. Indeed, combining low and intermediate groups mirrored findings from a prior study by the US military.[21] Hence, although lateral-flow tests are subject to high interrater variability in reading categorical rather than continuous results, here we
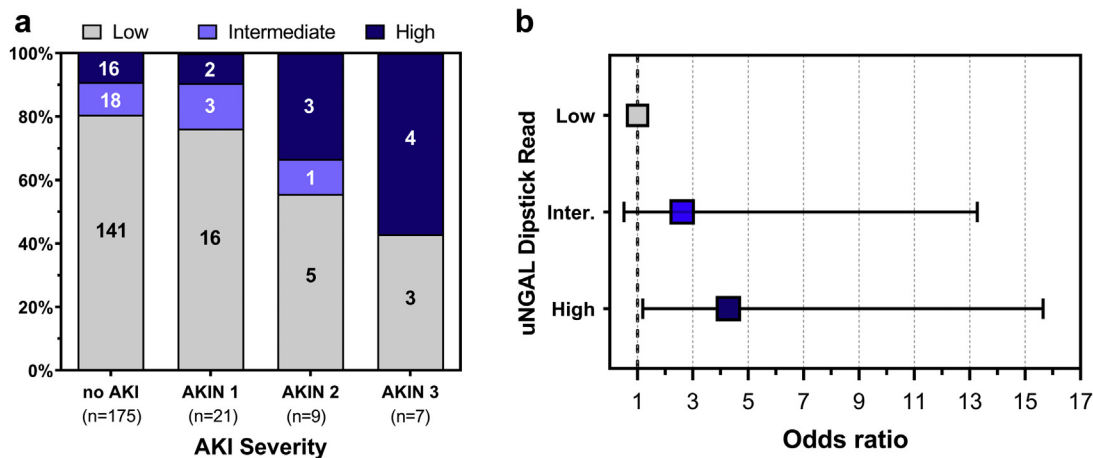
Figure 4. Urinary neutrophil gelatinase-associated lipocalin (uNGAL) dipstick measurements correlate with AKIN severity and predict the combined endpoint of in-hospital mortality and renal replacement therapy. (a) High uNGAL readings in the emergency department correlate with AKIN severity score ($P = 0.002$). Black = high uNGAL, dark gray = intermediate uNGAL; light gray = low uNGAL. (b) Patients with high uNGAL dipstick readings were significantly more likely to experience the composite outcome of death and/or renal replacement therapy during the index admission (high vs. low odds ratio = 4.32 [range 1.19–15.65], $P = 0.026$). AKIN, Acute Kidney Injury Network; DPI, dots per inch.

show that a binary readout (high vs. low or intermediate reads) facilitates its use.

We also note that despite substantial differences in patient characteristics and enrollment criteria, our test performance was similar to the US military study.[21] Using a similar cutoff, they found a sensitivity between 0.50 and 0.75 and a specificity of 0.82 to 0.97, which are comparable with our findings of 0.55 and 0.91, respectively. It is likely that differences in the positive predictive value and NPV are related to the differences in the prevalence of elevated SCr (14% in the current study vs. 37% in the Beyer *et al.* study[21]). In summary, the utility of the uNGAL dipstick is highlighted in 2 different clinical settings with different disease prevalence (the ED and intensive care unit).

In order to use the bedside tool accurately, it is necessary to understand the relation between different analytes that measure different aspects of kidney disease (biomarker [tubular injury] vs. SCr [decreased excretory function]). For example, transient increases in SCr may reflect compensatory responses rather than injury. As a result, analyses that fail to separate transient from more sustained elevations of SCr will result in misclassification of tAKI with low uNGAL as a false negative instead of a true negative. Similarly, the absence of a sustained rise of SCr despite a high uNGAL dipstick reading (e.g., in patients with limited muscle mass or in those individuals with adequate renal reserve) will result in misclassification as a false positive instead of a true positive. As a result, the test performance of the uNGAL dipstick is underestimated by a false-positive rate that is inflated by the poor temporal resolution and insensitivity of SCr as a

diagnostic test for acute tubular injury. The failure of SCr kinetics to identify all instances of tubular injury likely contributed to lower sensitivity in our study. In short, the inherent limitations of SCr as a gold standard adversely impacts the accurate assessment of biomarker performance. This limitation is likely reflected in patients who presented with biomarker evidence of injury (high uNGAL dipstick) but without SCr-based criteria of AKI. Indeed, 21% of these patients subsequently developed either sAKI, in-hospital renal replacement therapy initiation, or died by 90 days. Thus, in the absence of biomarker data, these individuals were at significant risk of inappropriate triage.

The bedside dipstick assay can also help clarify the clinical risk for patients with insufficient SCr kinetics (e.g., those patients who were hospitalized for <2 days). A number of these patients had high uNGAL dipstick readings at the time of ED presentation, representing a group at high risk of tubular injury. One could argue that the absence of serial SCr measurements in these patients may reflect the clinical judgment of the treating physicians (i.e., implying that the high uNGAL dipstick reading was a false positive). Nonetheless, individuals with high uNGAL dipstick readings who were discharged without serial SCr measurements had a higher 90-day mortality than patients with low uNGAL (30% vs. 2%, $P = 0.020$; Supplementary Table S5), suggesting that the high uNGAL dipstick readings were not false positives (please note the very limited number of events). Thus, the uNGAL dipstick provides a critical opportunity for improved triage of patients in the ED. Accordingly, we suggest that patients with low uNGAL but high SCr could be followed up in 1 to 2 days after

volume expansion, even as outpatients, whereas patients with high uNGAL but low SCr should be followed daily for the subsequent appearance of kidney dysfunction and electrolyte disorders.

Despite these strengths, this new point-of-care test is not without limitations. Although test performance characteristics were similar when confounders (advanced CKD and UTI) were ignored (Supplementary Table S3), the dipstick performed best when confounders were removed. Notably, although the presence or absence of a UTI can be easily resolved with a rapid urinalysis test in the ED, CKD may be difficult to immediately document in patients with inconsistent medical care. We also did not take into account whether the urine sediment may help in limiting the impact of confounders. In addition, this study was performed at a single site at an urban academic medical center, and only examined patients in the process of admission to the hospital, implying higher illness severity, which may not apply to all phases of care or geographies. A prospective, randomized clinical trial is needed for further evaluation for use in the clinical management of patients in the emergency room.

In summary, in the absence of confounders, 97% of individuals with low NGAL dipstick values did not have evidence of kidney injury based on SCr kinetics (true negatives). Furthermore, patients without SCr evidence of kidney injury were correctly identified by the NGAL dipstick in 91% of instances. Randomized clinical trials will highlight the potential of rapid bedside analyses in patient management. In the current era of shortened length of stay and triage based on isolated SCr measurements, the introduction of a dipstick allows for the accurate identification of individuals who do not have tubular injury. We suggest that these real-time data obtained safely at the bedside by dipstick technology will provide relief to our oversubscribed EDs, particularly during times of population-wide illness, by triaging patients with little risk of prolonged SCr elevation. The alternative, waiting 12 or 24 hours for the next SCr check, has always led to delays in treatments; it is no longer an option during a worldwide disaster.

DISCLOSURE

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

Supplementary File (PDF)
Table S1. Patient characteristics of all enrolled patients.
Table S2. Patient characteristics by AKI status.
Table S3. Creatinine kinetics, AKI definitions, and patient outcomes.
Table S4. Test performance characteristics of the uNGAL dipstick for all permutations.
Table S5. uNGAL dipstick and clinical outcomes in patients with insufficient SCr kinetics.
Figure S1. uNGAL detection by dipstick and categorization methods. (A) The color of the test line was compared with the manufacturer's scale by 2 independent readers who were blinded to clinical data. One percent (5/479 samples) had discordant readings >1 category; in these cases, independent laboratory personnel provided the tiebreaker. (B) These scores were transformed into 3 risk categories: low, intermediate, and high.
Figure S2. Comparison of baseline CKD stage with uNGAL measurements. (A) The mean uNGAL ELISA increases as the baseline eGFR decreases (*<0.001). (B) High uNGAL dipstick readings increase as the baseline eGFR decreases ($P < 0.001$). Light gray = low, dark gray = intermediate, and black = high.
Figure S3. UTI as a confounder of uNGAL. (A) The presence of a UTI raises the mean uNGAL ELISA (white = +UTI, light gray = unknown, and dark gray = no UTI; *<0.001) and (B) increases the uNGAL dipstick readings (light gray = low, dark gray = intermediate, and black = high; $P < 0.001$).
Figure S4. Comparison of uNGAL with clinical course. (A) Note that patients with a sustained elevation of SCr (sAKI = black) had the highest levels of uNGAL ELISA of all patients enrolled in the study (*<0.05). (B) After exclusion of the potential confounders (+UTI, unknown UTI status, and baseline eGFR < 30), patients with transient changes in SCr (tAKI = dark gray) had similar uNGAL dipstick values as patients without AKI (no AKI = white; *<0.01).
STROBE Statement.

REFERENCES

1. Sharma A, Mucino MJ, Ronco C. Renal functional reserve and renal recovery after acute kidney injury. *Nephron Clin Pract*. 2014;127:94–100.

2. Vanmassenhove J, Van Biesen W, Vanholder R, et al. Subclinical AKI: ready for primetime in clinical practice? *J Nephrol*. 2019;32:9–16.

3. Haase M, Kellum JA, Ronco C. Subclinical AKI–an emerging syndrome with important consequences. *Nat Rev Nephrol*. 2012;8:735–739.

4. Moretti C, Frajese GV, Guccione L, et al. Androgens and body composition in the aging male. *J Endocrinol Invest*. 2005;28:56–64.

5. Kimmel PL, Lew SQ, Bosch JP. Nutrition, ageing and GFR: is age-associated decline inevitable? *Nephrol Dial Transplant*. 1996;11(suppl 9):85–88.

6. Doi K, Yuen PS, Eisner C, et al. Reduced production of creatinine limits its use as marker of kidney injury in sepsis. *J Am Soc Nephrol*. 2009;20:1217–1221.

7. Desanti De Oliveira B, Xu K, Shen TH, et al. Molecular nephrology: types of acute tubular injury. *Nat Rev Nephrol*. 2019;15:599–612.

8. Molitoris BA. Therapeutic translation in acute kidney injury: the epithelial/endothelial axis. *J Clin Invest*. 2014;124:2355–2363.

9. Yuen PS, Jo SK, Holly MK, et al. Ischemic and nephrotoxic acute renal failure are distinguished by their broad transcriptomic responses. *Physiol Genomics*. 2006;25:375–386.

10. Xu K, Rosenstiel P, Paragas N, et al. Unique transcriptional programs identify subtypes of AKI. *J Am Soc Nephrol*. 2017;28:1729–1740.

11. Mehta S, Chauhan K, Patel A, et al. The prognostic importance of duration of AKI: a systematic review and meta-analysis. *BMC Nephrol*. 2018;19:91.

12. Coca SG, Nadkarni GN, Garg AX, et al. First post-operative urinary kidney injury biomarkers and association with the duration of AKI in the TRIBE-AKI Cohort. *PLoS One*. 2016;11:e0161098.

13. Coca SG, King JT Jr, Rosenthal RA, et al. The duration of postoperative acute kidney injury is an additional parameter predicting long-term survival in diabetic veterans. *Kidney Int*. 2010;78:926–933.

14. Yang L, Besschetnova TY, Brooks CR, et al. Epithelial cell cycle arrest in G2/M mediates kidney fibrosis after injury. *Nat Med*. 2010;16:535–543, 1p following 143.

15. Canaud G, Brooks CR, Kishi S, et al. Cyclin G1 and TASCC regulate kidney epithelial cell G2-M arrest and fibrotic maladaptive repair. *Sci Transl Med*. 2019;11, eaav4754.

16. Arai S, Kitada K, Yamazaki T, et al. Apoptosis inhibitor of macrophage protein enhances intraluminal debris clearance and ameliorates acute kidney injury in mice. *Nat Med*. 2016;22:183–193.

17. Yang L, Brooks CR, Xiao S, et al. KIM-1-mediated phagocytosis reduces acute injury to the kidney. *J Clin Invest*. 2015;125:1620–1636.

18. Zager RA, Johnson ACM. Acute kidney injury induces dramatic p21 upregulation via a novel, glucocorticoid-activated, pathway. *Am J Physiol Renal Physiol*. 2019;316:F674–F681.

19. Uchino S, Bellomo R, Bagshaw SM, et al. Transient azotaemia is associated with a high risk of death in hospitalized patients. *Nephrol Dial Transplant*. 2010;25:1833–1839.

20. Brown JR, Kramer RS, Coca SG, et al. Duration of acute kidney injury impacts long-term survival after cardiac surgery. *Ann Thorac Surg*. 2010;90:1142–1148.

21. Beyer CA, Burmeister DM, Gomez BI, et al. Point-of-care urinary biomarker testing for risk prediction in critically injured combat casualties. *J Am Coll Surg*. 2019;229:508–515.e501.

22. Paragas N, Qiu A, Zhang Q, et al. The Ngal reporter mouse detects the response of the kidney to injury in real time. *Nat Med*. 2011;17:216–222.

23. Kuwabara T, Mori K, Mukoyama M, et al. Urinary neutrophil gelatinase-associated lipocalin levels reflect damage to glomeruli, proximal tubules, and distal nephrons. *Kidney Int*. 2009;75:285–294.

24. Mishra J, Ma Q, Prada A, et al. Identification of neutrophil gelatinase-associated lipocalin as a novel early urinary biomarker for ischemic renal injury. *J Am Soc Nephrol*. 2003;14:2534–2543.

25. Mishra J, Dent C, Tarabishi R, et al. Neutrophil gelatinase-associated lipocalin (NGAL) as a biomarker for acute renal injury after cardiac surgery. *Lancet*. 2005;365:1231–1238.

26. Johnson ACM, Zager RA. Mechanisms underlying increased TIMP2 and IGFBP7 urinary excretion in experimental AKI. *J Am Soc Nephrol*. 2018;29:2157–2167.

27. Nickolas TL, Schmidt-Ott KM, Canetta P, et al. Diagnostic and prognostic stratification in the emergency department using urinary biomarkers of nephron damage: a multicenter prospective cohort study. *J Am Coll Cardiol*. 2012;59:246–255.

28. Nickolas TL, O'Rourke MJ, Yang J, et al. Sensitivity and specificity of a single emergency department measurement of urinary neutrophil gelatinase-associated lipocalin for diagnosing acute kidney injury. *Ann Intern Med*. 2008;148:810–819.

29. Singer E, Elger A, Elitok S, et al. Urinary neutrophil gelatinase-associated lipocalin distinguishes pre-renal from intrinsic renal failure and predicts outcomes. *Kidney Int*. 2011;80:405–414.

30. Mar D, Gharib SA, Zager RA, et al. Heterogeneity of epigenetic changes at ischemia/reperfusion- and endotoxin-induced acute kidney injury genes. *Kidney Int*. 2015;88:734–744.

31. Basu RK, Wong HR, Krawczeski CD, et al. Combining functional and tubular damage biomarkers improves diagnostic precision for acute kidney injury after cardiac surgery. *J Am Coll Cardiol*. 2014;64:2753–2762.

32. Ahmad T, Jackson K, Rao VS, et al. Worsening renal function in patients with acute heart failure undergoing aggressive diuresis is not associated with tubular injury. *Circulation*. 2018;137:2016–2028.

33. Belcher JM, Sanyal AJ, Peixoto AJ, et al. Kidney biomarkers and differential diagnosis of patients with cirrhosis and acute kidney injury. *Hepatology*. 2014;60:622–632.

34. Nickolas TL, Forster CS, Sise ME, et al. NGAL (Lcn2) monomer is associated with tubulointerstitial damage in chronic kidney disease. *Kidney Int*. 2012;82:718–722.

35. Kanda J, Mori K, Kawabata H, et al. An AKI biomarker lipocalin 2 in the blood derives from the kidney in renal injury but from neutrophils in normal and infected conditions. *Clin Exp Nephrol*. 2015;19:99–106.

36. Field M, Dronavalli V, Mistry P, et al. Urinary biomarkers of acute kidney injury in deceased organ donors–kidney injury molecule-1 as an adjunct to predicting outcome. *Clin Transplant*. 2014;28:808–815.

37. Kashani K, Al-Khafaji A, Ardiles T, et al. Discovery and validation of cell cycle arrest biomarkers in human acute kidney injury. *Crit Care*. 2013;17:R25.

38. Wetz AJ, Richardt EM, Wand S, et al. Quantification of urinary TIMP-2 and IGFBP-7: an adequate diagnostic test to predict acute kidney injury after cardiac surgery? *Crit Care*. 2015;19:3.

39. Oezkur M, Magyar A, Thomas P, et al. TIMP-2*IGFBP7 (Nephrocheck(R)) measurements at intensive care unit admission after cardiac surgery are predictive for acute kidney injury within 48 hours. *Kidney Blood Press Res*. 2017;42:456–467.

40. Edwards JK. Biomarkers. How precise is NephroCheck(R)? *Nat Rev Nephrol*. 2015;11:127.

41. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med*. 2009;150: 604–612.

42. Viau A, El Karoui K, Laouari D, et al. Lipocalin 2 is essential for chronic kidney disease progression in mice and humans. *J Clin Invest*. 2010;120:4065–4076.

43. Mehta RL, Kellum JA, Shah SV, et al. Acute Kidney Injury Network: report of an initiative to improve outcomes in acute kidney injury. *Crit Care*. 2007;11:R31.

44. Paragas N, Kulkarni R, Werth M, et al. alpha-Intercalated cells defend the urinary system from bacterial infection. *J Clin Invest*. 2014;124:2963–2976.

45. Cuzick J. A Wilcoxon-type test for trend. *Stat Med*. 1985;4:87–90.

46. Murray PT, Mehta RL, Shaw A, et al. Potential use of biomarkers in acute kidney injury: report and summary of recommendations from the 10th Acute Dialysis Quality Initiative consensus conference. *Kidney Int*. 2014;85:513–521.

47. Krawczeski CD, Goldstein SL, Woo JG, et al. Temporal relationship and predictive value of urinary acute kidney injury biomarkers after pediatric cardiopulmonary bypass. *J Am Coll Cardiol*. 2011;58:2301–2309.