Blind Men and Elephants and the Biological Markers of AKI

Edward D. Siew

Division of Nephrology and Hypertension, Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee

The ability to determine the onset, type, and location of injury is the holy grail of clinical nephrology, particularly acute kidney injury (AKI). Current evaluation limited to clinical intuition and functional markers such as serum creatinine belies a timely and full characterization of this disease. This uncertainty has challenged clinicians and investigators at fundamental levels, impeding the narrowing of differential diagnoses, the arrival of an effective consensus definition, and the development and testing of potential therapies. The emergence of novel biomarkers of injury with potential to add further insight has generated great interest within the nephrology community.

In this issue of JASN, Parikh and colleagues report the results of two large, integrated observational studies, which examine the ability of neutrophil gelatinase-associated lipocalin (NGAL) and IL-18 to predict severe AKI and provide prognostic information in parallel cohorts of adult and pediatric patients.[1,2] Both studies, part of the multi-centered National Heart, Lung, and Blood Institute-sponsored Translational Research Investigating Biomarker Endpoints in AKI (TRIBE-AKI) consortium, prospectively enrolled 1,219 adults and 311 children undergoing cardiac surgery from centers across the United States and Canada. The authors assessed whether postoperative levels of urine and plasma NGAL (uNGAL and pNGAL) and urine IL-18 (uIL-18) improve the ability to predict a doubling of baseline serum creatinine or the need for acute dialysis and other clinical endpoints.

As expected, patients at highest risk for developing AKI were those with higher comorbidity and lengthier, more complicated surgery. Consistent with earlier cardiac surgery studies, levels of biomarkers transiently peaked during early injury and remained elevated in patients with AKI relative to those without AKI. Diagnostic testing using the area under the receiver-operating curve (AUC), or C-statistic, revealed moderate performance in the adult study with AUCs of 0.67 (SE 0.04) and 0.70 (SE 0.04) for uNGAL and pNGAL, respectively, and 0.74 (SE 0.04) for uIL-18. Individual performance of each biomarker within the pediatric study was less robust than earlier reports with AUCs for uNGAL and pNGAL of 0.71 (SE 0.04) and 0.56 (SE 0.05), respectively, and 0.72 (SE 0.04) for uIL-18.

In both studies, the ability of these biomarkers to risk-stratify patients beyond a clinical model was also tested. After adjustment for demographic and intraoperative risk factors, adult patients within the highest quintiles of postoperative levels of pNGAL and uIL-18 had between a 5- to 7-fold incremental risk of developing severe AKI when compared with those in the lowest quintiles, with uNGAL not showing a statistically significant association. In pediatric patients, the highest quintiles of uNGAL and uIL-18 levels conferred an approximate 4- to 7-fold incremental risk relative to the lowest quintiles with pNGAL not demonstrating a significant association. Net reclassification and integrated discrimination indices, which quantify the global ability of biomarkers to correctly reclassify individuals with and without severe AKI to their respective higher or lower risk categories beyond a clinical model, indicate a 22 to 25% improvement in overall reclassification for uIL-18 in both studies, a 14 to 17% improvement for plasma and urine NGAL in the pediatric study, and an 18% improvement for pNGAL in the adult study. Finally, all biomarkers showed a significant ability to predict harder clinical endpoints including survival and duration of care (ventilation, ICU, and hospital days).

Investigators from both studies should be commended for their efforts and execution. In addition to an enriching sample size with higher-risk patients, meticulous attention was paid to standardizing serial data and biospecimen collection across centers; detailed quality control data for bioassays were reported; earnest attempts were made to reduce AKI misclassification through determination of baseline kidney function; indices of discrimination, calibration, and reclassification were examined for clinical utility; and finally, additional relevant endpoints (mortality and length of stay) were studied. As discussed previously,[3,4] many of these steps represent key elements of biomarker validation recommended by recent scientific consensus statements and clearly raise the bar for future biomarker studies in this field.

Unfortunately, these initial analyses suggest that considerable work remains in the quest to improve our ability to comprehensively phenotype this disease. Particularly disappointing were the diagnostic data suggesting only moderate levels of discrimination across all tested biomarkers (all AUCs <0.8). Even when applying more sensitive definitions of AKI used in previous studies indicating more robust performance,[5–7] the modest performance persisted. Although variations in case mix[8–10] and assay performance likely contributed partially toward these differences, further interpretation of these tests remind us of the inherent challenges faced when extracting information from these markers with conventional approaches.

From a diagnostic perspective, receiver operating characteristic (ROC) curves estimate the probability that the

levels of a given biomarker will correctly rank-order any two patients from the study population classified as having or not having AKI. The value of ROC curve analyses rests on the ability to make these classifications with high accuracy in both groups. Although often used for their ease of interpretation and ability to compare findings, ROC curves that use arbitrary creatinine cutoffs to perform this task are limited because the latter lends itself to misclassification of actual injury status. Use of milder creatinine cutoffs (an increase of 0.3 mg/dl or 50% from baseline) in previous biomarker studies, for example, has been criticized as it risks grouping patients without true injury as having AKI and potentially blunts important associations between biomarker and outcome.[11,12] The impressive recruitment in TRIBE, while allowing for a more specific and likely uncontested conventional definition of AKI (doubling of serum creatinine or dialysis), also results in a precarious assumption that anything less than a doubling of creatinine does not represent injury that would prompt a substantial rise in biomarker levels.

This approach may place potentially sensitive biomarker candidates in the position of being judged, for example, by their ability to distinguish a patient whose serum creatinine increases from 1.0 to 2.1 mg/dl from one whose creatinine only increases from 1.0 to 1.9 mg/dl. In addition, dichotomization of creatinine also complicates comparisons of performance between markers, because those with a *sweet spot* of detection in mild injury may perform less well relative to others when using more specific creatinine cutoffs.[12] In the end, restricting our focus to how much creatinine and injury markers agree rather than exploring novel methods to characterize their full relationship may ultimately mislead and hinder a fuller sense of the overall AKI picture.[13,14]

This current struggle recalls the parable of the blind men and the elephant, in which a group of blind men are challenged to describe what an elephant is by touch. As each feels and describes a different part of the beast, they soon find themselves to be in apparent disagreement. Lost among the ensuing conflict is that each has discovered important partial truths. Accepting the premise that AKI biomarkers may be providing additional information beyond creatinine will ultimately require lines of investigation extended in this direction. Importantly, the authors found a robust association between the levels of tested biomarkers and harder clinical endpoints including mortality and length of stay. This latter finding has been pivotal in the ascension of other injury markers, such as cardiac troponin, toward clinical usefulness.[15,16] However, because both markers tested in TRIBE have potential roles in both kidney and systemic inflammation, the relative contribution of the former remains unclear.

In addition to exploring methods to help determine the tissue origin of measured levels of biomarkers,[17,18] another potential avenue for investigation includes a detailed examination of the outcomes in patients in whom creatinine and biomarker data disagree on injury status. Indeed, a recent study by Haase *et al.*[19] examining data from several NGAL studies suggests that in the group of non-AKI patients, those with elevated levels of biomarker are at higher risk for dialysis and other poor outcomes than biomarker negative counterparts. Although limited by the pooling of data from different studies, further testing of this hypothesis is warranted.

Also notable is that despite their reported ability as markers of tubular injury, a conspicuous absence of data exists examining the relationship between biomarkers and indices of tubular function including concentrating or diluting ability (Fe_{H20} and Fe_{urea}), acidosis, electrolyte complications, sediment, volume overload, and novel imaging techniques.[20,21] Finally, whereas tissue diagnoses are not always feasible or necessary, establishing a multi-center registry of patients with urine and blood collected from patients with available biopsy data in kidney injury (acute or chronic) would help compare creatinine and biomarker data against a more informative standard as has been done in preclinical studies.[22,23]

In addition to diagnostic discrimination, the authors examined whether NGAL and IL-18 improve the ability to accurately assess the risk of severe AKI beyond current clinical predictors through prediction models and reclassification indices. Although the highest levels of biomarkers independently associated with severe AKI, it was notable that after inclusion of additional postoperative variables, including serum creatinine, only uIL-18 remained predictive in supplemental analyses for the adult study. The results suggest that caution may be warranted before applying these markers in a clinical setting. For example, although both NGAL and IL-18 are able to substantially improve the overall accuracy of risk categorization by net reclassification and integrated discrimination indices, the utility of these measures often depends on the risk categories chosen and the context in which they might be applied. The authors indicate that the choice of severe AKI was made on the basis of its potential as an endpoint used for future therapeutic trials. A closer look at the reclassification tables in the adult study reveal between 34 and 41% of patients with severe AKI were correctly reclassified to a higher level of risk relative to clinical predictors using pNGAL and IL-18, respectively. Although this reclassification is encouraging, it was also the case that up to 17% of patients without severe AKI were also incorrectly reclassified into a higher risk category using either biomarker. This latter finding would be important to consider, for example, before implementing them to guide clinical-trial enrollment when contemplating interventions with significant risk.

In summary, as the incidence of AKI continues to rise with marginal, if any, improvements in short-term mortality, the status quo of our ability to accurately recognize and treat this disease remains unacceptable.[24] These landmark investigations by Parikh and colleagues serves to highlight that although expectations are extremely high, significant caution is warranted before applying these markers in clinical care and research settings. Shedding further light on this beast will depend on our ability to determine what these markers are truly telling us. The latter will require collaborations using new and creative approaches that add to and integrate pieces of the puzzle already in place.

DISCLOSURES

REFERENCES

1. Parikh CR, Coca SG, Thiessen-Philbrook H, Shlipak MG, Koyner JL, Wang Z, Edelstein CL, Devarajan P, Patel UD, Zappitelli M, Krawczeski CD, Passik CS, Swaminathan M, Garg, AX: Postoperative biomarkers predict acute kidney injury and poor outcomes after adult cardiac surgery. *J Am Soc Nephrol* 22: 1748–1757, 2011

2. Parikh CR, Devarajan P, Zappitelli M, Sint K, Thiessen-Philbrook H, Li S, Kim RW, Koyner JL, Coca SG, Edelstein CL, Shlipak MG, Garg AX, Krawczeski, CD: Postoperative biomarkers predict acute kidney injury and poor outcomes after pediatric cardiac surgery. *J Am Soc Nephrol* 22: 1737–1747, 2011

3. Hlatky MA, Greenland P, Arnett DK, Ballantyne CM, Criqui MH, Elkind MS, Go AS, Harrell FE, Jr, Hong Y, Howard BV, Howard VJ, Hsue PY, Kramer CM, McConnell JP, Normand SL, O'Donnell CJ, Smith SC, Jr, Wilson PW: Criteria for evaluation of novel markers of cardiovascular risk: A scientific statement from the American Heart Association. *Circulation* 119: 2408–2416, 2009

4. Siew ED, Ware LB, Ikizler TA: Biological markers of acute kidney injury. *J Am Soc Nephrol* 22: 810–820, 2011

5. Mishra J, Dent C, Tarabishi R, Mitsnefes MM, Ma Q, Kelly C, Ruff SM, Zahedi K, Shao M, Bean J, Mori K, Barasch J, Devarajan P: Neutrophil gelatinase-associated lipocalin (NGAL) as a biomarker for acute renal injury after cardiac surgery. *Lancet* 365: 1231–1238, 2005

6. Haase-Fielitz A, Bellomo R, Devarajan P, Story D, Matalanis G, Dragun D, Haase M: Novel and conventional serum biomarkers predicting acute kidney injury in adult cardiac surgery: A prospective cohort study. *Crit Care Med* 37: 553–560, 2009

7. Parikh CR, Mishra J, Thiessen-Philbrook H, Dursun B, Ma Q, Kelly C, Dent C, Devarajan P, Edelstein CL: Urinary IL-18 is an early predictive biomarker of acute kidney injury after cardiac surgery. *Kidney Int* 70: 199–203, 2006

8. Cho K, Hsu CY: Quantifying severity of chronic kidney disease as a risk factor for acute kidney injury. *J Am Soc Nephrol* 21: 1602–1604, 2010

9. Grams ME, Astor BC, Bash LD, Matsushita K, Wang Y, Coresh J: Albuminuria and estimated glomerular filtration rate independently associate with acute kidney injury. *J Am Soc Nephrol* 21: 1757–1764, 2010

10. Huang TM, Wu VC, Young GH, Lin YF, Shiao CC, Wu PC, Li WY, Yu HY, Hu FC, Lin JW, Chen YS, Lin YH, Wang SS, Hsu RB, Chang FC, Chou NK, Chu TS, Yeh YC, Tsai PR, Huang JW, Lin SL, Chen YM, Ko WJ, Wu KD: Preoperative proteinuria predicts adverse renal outcomes after coronary artery bypass grafting. *J Am Soc Nephrol* 22: 156–163, 2011

11. Doi K, Yuen PS, Eisner C, Hu X, Leelahavanichkul A, Schnermann J, Star RA: Reduced production of creatinine limits its use as marker of kidney injury in sepsis. *J Am Soc Nephrol* 20: 1217–1221, 2009

12. Waikar SS, Bonventre JV: Creatinine kinetics and the definition of acute kidney injury. *J Am Soc Nephrol* 20: 672–679, 2009

13. Obuchowski, NA: An RO: C-type measure of diagnostic accuracy when the gold standard is continuous-scale. *Stat Med* 25: 481–493, 2006

14. Pickering JW, Frampton CM, Endre ZH: Evaluation of trial outcomes in acute kidney injury by creatinine modeling. *Clin J Am Soc Nephrol* 4: 1705–1715, 2009

15. Antman EM, Tanasijevic MJ, Thompson B, Schactman M, McCabe CH, Cannon CP, Fischer GA, Fung AY, Thompson C, Wybenga D, Braunwald E: Cardiac-specific troponin I levels to predict the risk of mortality in patients with acute coronary syndromes. *N Engl J Med* 335: 1342–1349, 1996

16. Hamm CW, Ravkilde J, Gerhardt W, Jorgensen P, Peheim E, Ljungdahl L, Goldmann B, Katus HA: The prognostic value of serum troponin T in unstable angina. *N Engl J Med* 327: 146–150, 1992

17. Munshi R, Johnson A, Siew ED, Ikizler TA, Ware LB, Wurfel MM, Himmelfarb J, Zager RA: MCP-1 gene activation marks acute kidney injury. *J Am Soc Nephrol* 22: 165–175, 2011

18. Gonzales PA, Pisitkun T, Hoffert JD, Tchapyjnikov D, Star RA, Kleta R, Wang NS, Knepper MA: Large-scale proteomics and phosphoproteomics of urinary exosomes. *J Am Soc Nephrol* 20: 363–379, 2009

19. Haase M, Devarajan P, Haase-Fielitz A, Bellomo R, Cruz DN, Wagener G, Krawczeski CD, Koyner JL, Murray P, Zappitelli M, Goldstein SL, Makris K, Ronco C, Martensson J, Martling CR, Venge P, Siew E, Ware LB, Ikizler TA, Mertens PR: The outcome of neutrophil gelatinase-associated lipocalin-positive subclinical acute kidney injury: A multicenter pooled analysis of prospective studies. *J Am Coll Cardiol* 57: 1752–1761, 2011

20. Molitoris BA, Sandoval RM: Multiphoton imaging techniques in acute kidney injury. *Contrib Nephrol* 165: 46–53, 2010

21. Sharfuddin AA, Sandoval RM, Molitoris BA: Imaging techniques in acute kidney injury. *Nephron Clin Pract* 109: c198–c204, 2008

22. Vaidya VS, Ozer JS, Dieterle F, Collings FB, Ramirez V, Troth S, Muniappa N, Thudium D, Gerhold D, Holder DJ, Bobadilla NA, Marrer E, Perentes E, Cordier A, Vonderscher J, Maurer G, Goering PL, Sistare FD, Bonventre JV: Kidney injury molecule-1 outperforms traditional biomarkers of kidney injury in preclinical biomarker qualification studies. *Nat Biotechnol* 28: 478–485, 2010

23. Negishi K, Noiri E, Doi K, Maeda-Mamiya R, Sugaya T, Portilla D, Fujita T: Monitoring of urinary L-type fatty acid-binding protein predicts histological severity of acute kidney injury. *Am J Pathol* 174: 1154–1159, 2009

24. Hsu CY: Where is the epidemic in kidney disease? *J Am Soc Nephrol* 21: 1607–1611, 2010

Screening for Proteinuria: A Tool for Predicting Rapid Declines in Kidney Function?

Amanda Hyre Anderson

Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania

The importance of chronic kidney disease (CKD) as an independent risk factor for cardiovascular disease (CVD) and mor-

Correspondence: Dr. Amanda Hyre Anderson, Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania Perelman School of Medicine, 929 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104. Phone: 215-573-7339; Fax: 215-898-0643; E-mail: andersah@mail.med.upenn.edu