

# Biomarker Enhanced Risk Prediction for Adverse Outcomes in Critically Ill Patients Receiving RRT

Francis Pike,<sup>\*†‡</sup> Raghavan Murugan,<sup>\*†</sup> Christopher Keener,<sup>\*†‡</sup> Paul M. Palevsky,<sup>\*§</sup> Anitha Vijayan,<sup>||</sup> Mark Unruh,<sup>†¶</sup> Kevin Finkel,<sup>\*\*</sup> Xiaoyan Wen,<sup>\*†</sup> and John A. Kellum<sup>\*†</sup> for the Biological Markers for Recovery of Kidney (BioMaRK) Study Investigators

## Abstract

**Background and objectives** Higher plasma concentrations of inflammatory and apoptosis markers in critically ill patients receiving RRT are associated with RRT dependence and death. This study objective was to examine whether plasma inflammatory (IL-6, -8, -10, and -18; macrophage migration inhibitory factor) and apoptosis (death receptor-5, tumor necrosis factor receptor I and II) biomarkers augment risk prediction of renal recovery and mortality compared with clinical models.

**Design, setting, participants, & measurements** The Biologic Markers of Recovery for the Kidney study ( $n=817$ ) was a prospective, nested, observational cohort study conducted as an ancillary to the Veterans Affairs/National Institutes of Health Acute renal failure Trial Network study, a randomized trial of intensive versus less intensive RRT in critically ill patients with AKI conducted between November 2003 and July 2007 at 27 Veterans Affairs- and university-affiliated centers. Primary outcomes of interest were renal recovery and mortality at day 60.

**Results** A parsimonious clinical model consisting of only four variables (age, mean arterial pressure, mechanical ventilation, and bilirubin) predicted renal recovery (area under the receiver-operating characteristic curve [AUROC], 0.73; 95% confidence interval [95% CI], 0.68 to 0.78) and mortality (AUROC, 0.74; 95% CI, 0.69 to 0.78). By contrast, individual biomarkers were only modestly predictive of renal recovery (AUROC range, 0.55–0.63) and mortality (AUROC range, 0.54–0.68). Adding plasma IL-8 to a parsimonious model augmented prediction of recovery (AUROC, 0.76; 95% CI, 0.71 to 0.81;  $P=0.04$ ) and mortality (AUROC, 0.78; 95% CI, 0.73 to 0.82;  $P<0.01$ ) compared with the clinical model alone.

**Conclusions** This study suggests that a simple four-variable clinical model with plasma IL-8 had predictive value for renal recovery and mortality. These findings require external validation but could easily be used by clinicians.

*Clin J Am Soc Nephrol* 10: 1332–1339, 2015. doi: 10.2215/CJN.09911014

## Introduction

Severely ill patients with AKI requiring RRT are at a higher risk of dialysis dependence and death. More than half of patients receiving RRT die, and only a third are alive and independent of RRT by 2 months after acute illness (1). Despite decades of research, no specific treatment has improved outcomes from severe AKI. One important reason might be the heterogeneous nature of patients with severe AKI (2). While some patients with severe AKI may recover without interventions, other patients require specific interventions to improve outcomes.

Early risk stratification has important therapeutic implications, such as to determine the prognosis, ascertain the timing of initiation of interventions, and enroll a homogenous group of patients in clinical trials of AKI. Thus, it is important to distinguish patients at risk for death and nonrecovery of kidney function from those who are likely to recover early on. Emerging biomarkers of AKI alone or in combination with clinical variables could aid in risk stratification.

We previously found that higher circulating concentrations of plasma inflammatory (IL-6, -8, -10, and -18 and macrophage migration inhibitory factor) and apoptosis (tumor necrosis factor receptor (TNFR) I and II and death receptor-5) biomarkers were strongly associated with RRT dependence and death using the Biologic Markers of Recovery for the Kidney (BioMaRK) study cohort (1). BioMaRK was as an ancillary study to the Veterans Affairs (VA)/National Institutes of Health (NIH) Acute renal failure Trial Network (ATN) study (3).

Prior work using the ATN cohort also found that a model consisting of 21 clinical variables predicted 60-day mortality and outperformed existing severity of illness scoring systems (4). However, it is not clear whether a more parsimonious clinical model, which can be easily used by clinicians at the bedside, with or without biomarkers, can be used to predict outcomes. In this study, we examined the predictive value of various clinical models along with plasma biomarkers for mortality and renal recovery using the BioMaRK cohort.

\*The Center for Critical Care Nephrology, Department of Critical Care Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania; †The CRISMA (Clinical Research, Investigation, and Systems Modeling of Acute illness) Center, Department of Critical Care Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania; ‡Department of Biostatistics, University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania; §Renal Section, Veterans Affairs Pittsburgh Healthcare System, Pittsburgh, Pennsylvania; ||Division of Renal Diseases, Washington University, St. Louis, Missouri; ¶Division of Nephrology, Department of Internal Medicine, University of New Mexico, Albuquerque, New Mexico; and \*\*Division of Renal Diseases and Hypertension, University of Texas Medical School at Houston, Houston, Texas

**Correspondence:** Dr. John A. Kellum, Center for Critical Care Nephrology, University of Pittsburgh, 604 Scaife Hall, 3550 Terrace Street, Pittsburgh, PA 15261. Email: kellumja@ccm.upmc.edu

**Table 1. Modeling variables used in various clinical models**

Characteristic	Reduced <sup>a</sup> ATN Model		LASSO Model		Stepwise Model		Parsimonious Model	
	Mortality	Recovery	Mortality	Recovery	Mortality	Recovery	Mortality	Recovery
Age			Age	Age	Age	Age	Age	Age
Chronic hypoxemia	Mechanical ventilation	Mechanical ventilation	Mechanical ventilation	Mechanical ventilation	Mechanical ventilation	Mechanical ventilation	Mechanical ventilation	Mechanical ventilation
Cardiovascular disease	Arterial pH (Z-score)	Arterial pH (Z-score)	Arterial pH (Z-score)	Arterial pH (Z-score)	Arterial pH (Z-score)	Arterial pH (Z-score)	Arterial pH (Z-score)	Arterial pH (Z-score)
Malignancy	Total bilirubin	Total bilirubin	Total bilirubin	Total bilirubin	Total bilirubin	Total bilirubin	Total bilirubin	Total bilirubin
Immunosuppressive therapy	Mean arterial pressure	Mean arterial pressure	Mean arterial pressure	Mean arterial pressure	Platelets	Platelets	Platelets	Platelets
Ischemic AKI	Platelets	Platelets	Platelets	Platelets	Platelets	Platelets	Platelets	Platelets
Postsurgery at RRT initiation			Total APACHE II score	Total APACHE II score	Total APACHE II score	Total APACHE II score	Total APACHE II score	Total APACHE II score
Heart rate			Immunosuppressive therapy	Oliguria	Oliguria	Oliguria	Oliguria	Oliguria
Mean arterial pressure								
Urine volume								
Mechanical ventilation								
Fraction of inspired oxygen $\geq 0.6$								
Interaction of mechanical ventilation and fraction of inspired oxygen $\geq 60\%$								
Arterial pH (Z-score)								
Arterial oxygen partial pressure								
Serum creatinine								
Serum bicarbonate								
Serum phosphate								
Serum albumin								
Total bilirubin								
International normalized ratio								
Platelet count								

ATN, Acute renal failure Trial Network; LASSO, least absolute shrinkage and selection operator; APACHE, Acute Physiology and Chronic Health Evaluation.  
<sup>a</sup>ATN variables that were significant at  $P < 0.05$  in a multivariable logistic analysis in the Biologic Markers of Recovery for the Kidney cohort.

## Materials and Methods

### Study Design and Selection of Participants

The BioMaRK study ( $n=817$ ) was a multicenter, prospective, nested, observational cohort study conducted as an ancillary to the VA/NIH ATN clinical trial. The ATN study was a multicenter randomized clinical trial ( $n=1124$ ) comparing intensive and less intensive RRT strategies in critically ill patients that was conducted between November 2003 and July 2007 and is described in detail elsewhere (3). The BioMaRK study included all participants in the ATN study who gave additional written consent to blood collections for sample banking (1). The institutional review boards of the University of Pittsburgh and all other participating sites approved the study.

### Biomarker Sample and Data Collection and Follow-up

Blood samples were collected after study randomization in the ATN trial (day 1) in 817 participants. Details of the biomarker assays, detection threshold, and censoring are provided in detail elsewhere (1) and in the Supplemental Material (Item 1). We prospectively ascertained baseline characteristics, including demographic variables; cause of AKI; and other clinical, physiologic, and laboratory data at the time of entry into the ATN study. We collected individual comorbid illnesses and assessed comorbidity using the Charlson comorbidity score (5). Severity of illness was ascertained at enrollment using the Acute Physiology and Chronic Health Evaluation (APACHE) II (6), and the Cleveland Clinic Intensive Care Unit Acute Renal Failure score (7). All participants were followed daily until hospital discharge, death, or day 28 after randomization, whichever occurred first. Our primary outcomes of interest were renal recovery and mortality at day 60. Outcomes were ascertained daily during hospitalization and at days 28 and 60 using telephone and/or mail follow-up. Renal recovery was defined as being alive and independent from RRT by day 60, irrespective of the participant's discharge location.

### Statistical Analyses

We imputed missing categorical comorbidities using a regression-based algorithm as implemented in the ATN study, including the Sequential Organ Failure Assessment score ( $n=201$ ), APACHE II score ( $n=44$ ), and Cleveland Clinic Intensive Care Unit Renal Failure score ( $n=146$ ) (4). For all analyses, plasma biomarker data were log transformed and analyzed in the natural logarithm scale. Left-censored biomarker data were imputed by the lowest value and right-censored data by the highest value in measurement (1). Arterial pH was standardized using a normally distributed Z-score with zero mean and unit variance (SD,  $\pm 0.1$ ).

We first determined whether the distribution of clinical variables was similar between the BioMaRK and ATN cohorts using chi-square tests to compare categorical variables and  $t$  tests for continuous variables. We then examined whether the ATN model had good discrimination and calibration in the BioMaRK dataset. The ATN clinical model was a 21-variable model used to predict 60-day mortality. We then performed a random 2-fold split of the BioMaRK data to create separate derivation (model building) and validation data sets. The derivation and validation datasets were constructed using a random Bernoulli so that each record chosen had 50% probability of being in either dataset.

We derived four different clinical models from the BioMaRK derivation cohort: (1) reduced ATN model, (2) least absolute shrinkage and selection operator (LASSO) selected model, (3) stepwise selected model, and (4) parsimonious model restricted to variables routinely available clinically. The reduced ATN model included just those variables from the ATN model that retained significance in multiple logistic regression in the BioMaRK derivation dataset. The LASSO method is a penalized regression-based variable selection method that tends to be more stable and less computationally demanding than other selection methods as the number of variables increases (8). In the stepwise model,

**Table 2. Predictive ability of baseline clinical models for renal recovery and mortality**

Outcome and Clinical Model <sup>a</sup>	AUROC (95% CI)	Difference in AUROC's (95% CI)	P Value
<b>Recovery</b>			
ATN	0.77 (0.72 to 0.81)	Reference model	
Reduced ATN <sup>b</sup>	0.74 (0.69 to 0.79)	-0.02 (-0.05 to -0.001)	0.05
LASSO model <sup>c</sup>	0.76 (0.71 to 0.81)	-0.01 (-0.03 to 0.02)	0.60
Stepwise model <sup>c</sup>	0.75 (0.70 to 0.80)	-0.02 (-0.05 to 0.01)	0.16
Parsimonious model <sup>d</sup>	0.73 (0.68 to 0.78)	-0.03 (-0.06 to -0.01)	0.02
<b>Mortality</b>			
ATN	0.80 (0.76 to 0.85)	Reference model	—
Reduced ATN <sup>b</sup>	0.75 (0.71 to 0.80)	-0.03 (-0.05 to -0.01)	<0.01
Lasso model <sup>c</sup>	0.78 (0.74 to 0.83)	-0.02 (-0.05 to 0.001)	0.07
Stepwise model <sup>c</sup>	0.77 (0.73 to 0.82)	-0.03 (-0.05 to -0.01)	<0.01
Parsimonious model <sup>d</sup>	0.74 (0.69 to 0.78)	-0.06 (-0.10 to -0.03)	<0.001

ATN, Acute renal failure Trial Network; LASSO, least absolute shrinkage and selection operator; AUROC, area under the receiver-operating characteristic curve; 95% CI, 95% confidence interval.

<sup>a</sup>All models were built using the validation cohort consisting of 423 participants and had a Hosmer-Lemeshow  $P$  value  $>0.05$  indicating model calibration. Variables included in various clinical models are shown in Supplemental Table 1.

<sup>b</sup>Variables that remained significant at  $P<0.05$  in a multivariable model in the Biologic Markers of Recovery for the Kidney cohort.

<sup>c</sup>Variables deemed to be significant in a stepwise or penalized least absolute shrinkage and selection operator-based selection model.

<sup>d</sup>Four-variable parsimonious clinical model.

variables were entered using an entry criteria of  $P < 0.1$  and were retained if  $P < 0.05$ . Finally, we derived, using a chi-square scoring variable selection method, a parsimonious model in which variables were restricted to readily available clinical data at the bedside.

The area under the receiver-operating characteristic curve (AUROC) for each of the models in the validation set were compared using the DeLong method (9), their respective net reclassification index (NRI), integrated discrimination improvement (IDI), category-free net reclassification index (CFNRI), and calibration performance using the Hosmer-Lemeshow (HL) goodness-of-fit test (10).

We examined univariate predictive ability of selected biomarkers to determine whether any single biomarker outperformed the clinical model and determined the optimal cutoff for each biomarker using the Youden index (11) (sensitivity+specificity–1) as an optimization criterion using the validation cohort. The cutoff value was found by finding the fitted probability corresponding to the optimal Youden index; this probability, along with the fitted model intercept and slope, was then used to back-solve for the optimal cutoff (Equation 1 in Supplemental Material).

We added biomarkers to the clinical model to determine the added predictive ability. Before adding biomarkers we determined multicollinearity by computing variance inflation factors (12). We considered variance inflation factors  $< 5$  to indicate no serious multicollinearity. The relative performance of the biomarkers, over the clinical set, was determined *via* AUROC, NRI, IDI, and CFNRI, and calibration was applied to the validation set. Statistical analyses were performed using SAS software, version 9.3 (SAS Institute, Cary, NC), with significance set at  $P < 0.05$ .

**Results**

**Baseline Characteristics of BioMaRK and ATN Study Cohorts**

Of the 817 participants who formed the BioMaRK cohort, 298 (36.5%) were alive and free of RRT and 415 (50.8%) died by day 60. The baseline characteristics of BioMaRK and ATN cohorts were similar (Supplemental Table 1). The derivation (model building) and validation cohorts had 394 and 423 participants, respectively.

**Prediction of Renal Recovery and Mortality by Clinical Variables**

Table 1 shows the clinical variables included in different models. In the validation cohort, the ATN clinical model, reduced ATN model, LASSO model, stepwise-selected model, and parsimonious model had similar predictive capacity for renal recovery (AUROC range, 0.74–0.77) (Table 2) and had good calibration (HL  $P$  value range, 0.08–0.45). However, the eight-variable LASSO model, compared with the stepwise model, had a modest incremental discriminative value (IDI, 2% [95% confidence interval (95% CI), 1% to 4%]; CFNRI, 27% [95% CI, 8% to 47%]). A more parsimonious clinical model consisting of age, mean arterial pressure, mechanical ventilation, and bilirubin predicted recovery (AUROC, 0.73 [95% CI, 0.68 to 0.78]) and was equivalent to all four clinical models except the 21-variable ATN clinical model (Table 2).

**Table 3. Univariable association, prediction, and optimal cutoff values of individual biomarkers for clinical outcomes**

Biomarker	Recovery <sup>a</sup>			Mortality <sup>a</sup>		
	Odds Ratio (95% CI)	AUROC (95% CI)	Optimal Cutoff (Median) <sup>c</sup>	Odds Ratio (95% CI)	AUROC (95% CI)	Optimal Cutoff (Median) <sup>c</sup>
<b>Inflammation</b>						
IL-6	0.80 (0.73 to 0.88) <sup>b</sup>	0.61 (0.57 to 0.65)	96.54 (177.68)	1.33 (1.22 to 1.45) <sup>b</sup>	0.63 (0.59 to 0.67)	232.76 (179.47)
IL-8	0.73 (0.65 to 0.81) <sup>b</sup>	0.63 (0.59 to 0.67)	99.48 (103.54)	1.49 (1.34 to 1.66) <sup>b</sup>	0.66 (0.62 to 0.70)	108.48 (103.54)
IL-10	0.83 (0.75 to 0.93) <sup>b</sup>	0.57 (0.53 to 0.61)	13.60 (17.29)	1.30 (1.17 to 1.45) <sup>b</sup>	0.60 (0.56 to 0.63)	30.88 (17.29)
IL-18	0.79 (0.70 to 0.90) <sup>b</sup>	0.58 (0.54 to 0.62)	103.54 (102.51)	1.36 (1.21 to 1.54) <sup>b</sup>	0.60 (0.57 to 0.64)	82.27 (102.51)
MIF	0.85 (0.78 to 0.93) <sup>b</sup>	0.57 (0.53 to 0.61)	92.76 (254.68)	1.33 (1.22 to 1.46) <sup>b</sup>	0.62 (0.58 to 0.66)	113.30 (254.68)
<b>Apoptosis</b>						
TNFR I	0.54 (0.41 to 0.73) <sup>b</sup>	0.61 (0.57 to 0.65)	8955.29 (13095.19)	1.77 (1.33 to 2.37) <sup>b</sup>	0.60 (0.56 to 0.64)	9136.20 (13095.19)
TNFR II	0.73 (0.55 to 0.95) <sup>b</sup>	0.55 (0.51 to 0.59)	5653.33 (5653.33)	1.41 (1.08 to 1.84) <sup>b</sup>	0.55 (0.51 to 0.59)	5710.15 (5653.33)
DR-5	0.82 (0.69 to 0.97) <sup>b</sup>	0.55 (0.51 to 0.60)	202.35 (214.86)	1.41 (1.19 to 1.67) <sup>b</sup>	0.58 (0.54 to 0.62)	206.44 (208.51)

MIF, macrophage migration inhibitory factor; TNFR, TNF receptor; DR, death receptor; 95% CI, 95% confidence interval; AUROC, area under the receiver-operating characteristic curve. <sup>a</sup>For recovery, an odds ratio  $> 1$  indicates that higher biomarker concentration is associated with higher renal recovery, and an odds ratio  $< 1$  indicates nonrecovery. For the mortality model, odds ratio  $> 1$  indicates that higher biomarker concentration is associated with higher mortality and odds ratio  $< 1$  indicates lower mortality. The entire cohort was used to estimate the odds ratios for each biomarker. <sup>b</sup> $P < 0.05$ . <sup>c</sup>Raw biomarker concentration in pg/ml. Optimal cutoff values were derived using the validation cohort of 423 participants.

**Table 4. Added predictive value of biomarkers over the clinical models**

Outcome and Clinical Model <sup>a</sup>	AUROC (95% CI)	Integrated Discrimination Improvement (95% CI)	Category Free Net Reclassification Index (95% CI)	P Value <sup>c</sup>
<b>Recovery</b>				
LASSO	0.76 (0.71 to 0.81)	Reference model		0.45
Add IL-8	0.77 (0.72 to 0.82)	0.013 (0.002 to 0.03) <sup>b</sup>	0.31 (0.11 to 0.51) <sup>b</sup>	0.30
Add IL-10	0.77 (0.72 to 0.82)	0.014 (0.001 to 0.03) <sup>b</sup>	0.18 (−0.01 to 0.40)	0.07
Add TNFR I	0.76 (0.72 to 0.82)	0.003 (−0.003 to 0.009)	0.14 (−0.06 to 0.33)	0.33
LASSO+IL-8	0.77 (0.72 to 0.82)	Reference model		0.30
Add IL-10	0.77 (0.73 to 0.82)	0.01 (−0.003 to 0.01)	0.17 (−0.03 to 0.36)	0.36
Parsimonious model	0.73 (0.68 to 0.78)	Reference model		0.69
Parsimonious model+IL-8	0.76 (0.71 to 0.81)	0.02 (0.01 to 0.04) <sup>b</sup>	0.27 (0.08 to 0.47) <sup>b</sup>	0.50
<b>Mortality</b>				
LASSO	0.78 (0.74 to 0.83)	Reference model		0.08
Add IL-6	0.79 (0.75 to 0.83)	0.01 (−0.002 to 0.01)	0.10 (−0.09 to 0.29)	0.01
Add IL-8	0.80 (0.76 to 0.84)	0.02 (0.01 to 0.04) <sup>b</sup>	0.33 (0.14 to 0.52) <sup>b</sup>	0.09
Add IL-10	0.80 (0.76 to 0.84)	0.02 (0.01 to 0.04) <sup>b</sup>	0.27 (0.07 to 0.45) <sup>b</sup>	0.07
LASSO+IL-8	0.80 (0.76 to 0.84)	Reference model		0.09
Add IL-10		0.01 (−0.001 to 0.02)	0.14 (−0.05 to 0.33)	0.02
Parsimonious model	0.74 (0.70 to 0.79)	Reference model		0.43
Parsimonious model+IL-8	0.78 (0.73 to 0.82)	0.05 (0.03 to 0.07) <sup>b</sup>	0.41 (0.22 to 0.59) <sup>b</sup>	0.17

LASSO, least absolute shrinkage and selection operator; TNFR, TNF receptor; AUROC, area under the receiver-operating characteristic curve; 95% CI, 95% confidence interval.

<sup>a</sup>All models were built using the validation cohort consisting of 423 participants.

<sup>b</sup> $P < 0.05$  implies a statistical improvement for that particular metric by adding biomarkers.

<sup>c</sup> $P < 0.05$  implies lack of fit when adding biomarkers over and above the reference model.

Compared with renal recovery, all clinical models had slightly better predictive value for mortality (AUROC range, 0.75–0.80) (Table 2). Compared with the ATN clinical model, the reduced ATN model (AUROC, ATN versus reduced ATN, 0.80 versus 0.75;  $P < 0.01$ ), the stepwise model (AUROC, ATN versus stepwise, 0.80 versus 0.77;  $P < 0.01$ ), and the parsimonious model (AUROC, ATN versus parsimonious, 0.80 versus 0.74;  $P < 0.001$ ) had lower discrimination. Although the LASSO model and the stepwise model had good calibration (HL  $P$  value=0.08 and 0.09), compared with stepwise model the LASSO model exhibited only a modest improvement in reclassification (IDI, 3% [95% CI, 1% to 4%],  $P < 0.05$ ; CFNRI, 32% [95% CI, 13% to 51%],  $P < 0.05$ ). The parsimonious four-variable clinical model had the lowest predictive value for mortality (AUROC, 0.74 [95% CI, 0.69 to 0.78]) compared with all four models (Table 2).

#### Biomarker Prediction of Renal Recovery and Mortality

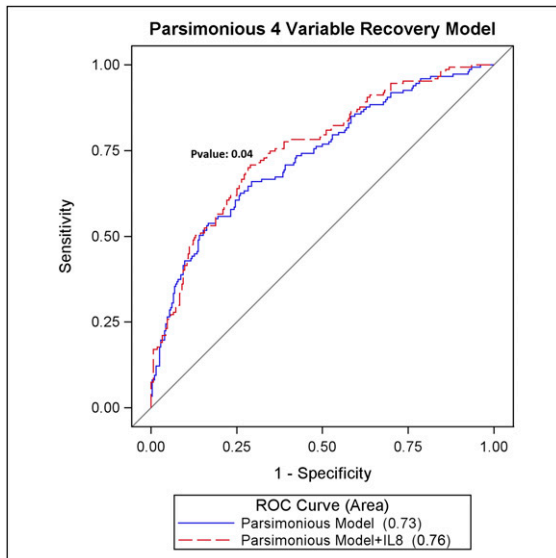
Table 3 shows the univariable association and predictive ability of individual biomarkers in the entire cohort. Optimal cutoff values for each biomarker was derived from the validation cohort. For all biomarkers, per natural log higher concentration in biomarker concentration was associated with lower odds for renal recovery (odds ratio range, 0.54–0.85) and higher odds of mortality (odds ratio range, 1.33–1.77) in the entire cohort. The predictive ability of individual biomarkers was poor for renal recovery (AUROC range, 0.55–0.63) and mortality (AUROC range, 0.55–0.66) (Supplemental Figures 1 and 2). The AUROC for all biomarkers for recovery and mortality were 0.66 and 0.71.

#### Added Predictive Ability of Biomarkers over Clinical Model on Outcomes

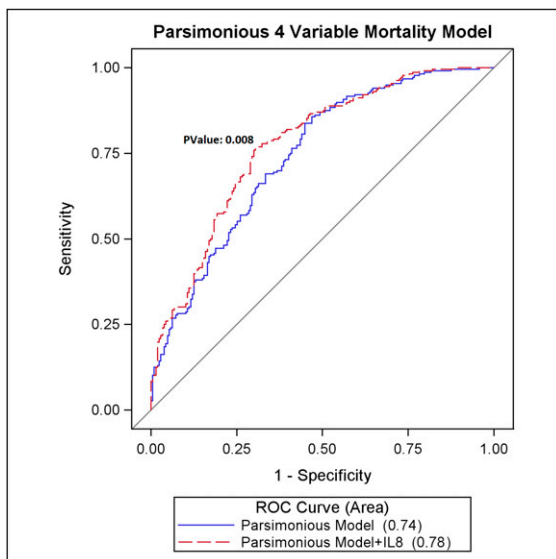
Through use of the validation cohort, Supplemental Figures 3 and 4 show the added value of all biomarkers (IL-6, -8, -10, and -18; TNFR receptor [TNFR] I and II, macrophage migration inhibitory factor, and death receptor-5) or just those biomarkers that remained, from either a separate stepwise or LASSO variable selection algorithm to the clinical model. The biomarkers obtained from these selection algorithms were IL-8, IL-10, and TNFR I for recovery and IL-6, IL-8, and IL-10 for mortality, respectively. Predictive ability for recovery for the LASSO clinical model (AUROC, 0.76 [95% CI, 0.71 to 0.81]), the LASSO clinical model plus all biomarkers (AUROC, 0.78 [95% CI, 0.73 to 0.82]), the LASSO clinical model plus IL-8, IL-10, and TNFR I (AUROC, 0.78 [95% CI, 0.73 to 0.82]), and the LASSO model with IL-8 alone (AUROC, 0.77 [95% CI, 0.73 to 0.82]) were similar, with only a slight improvement in IDI and CFNRI (Table 4).

Adding all of the biomarkers to the clinical model results in highest AUROC for mortality (AUROC, 0.82 [95% CI, 0.78 to 0.86]) (Supplemental Figure 4) and an improvement over the clinical model alone (AUROC, 0.78 [95% CI, 0.74 to 0.83]). The LASSO model with IL-8 alone was superior (AUROC, 0.80 [95% CI, 0.74 to 0.84]) to a LASSO model alone and no different from a LASSO model incorporating IL-6, IL-8, and IL-10 (AUROC, 0.80 [95% CI, 0.76 to 0.85]).

Adding IL-8 to a more parsimonious four-variable clinical model increased the prediction of renal recovery (AUROC, 0.76 [95% CI, 0.71 to 0.81];  $P = 0.04$ ) (Figure 1) and mortality (AUROC, 0.78 [95% CI, 0.73 to 0.82];  $P = 0.008$ ) (Figure 2) compared with clinical model alone. Indeed, these models were



**Figure 1.** | Prediction of renal recovery using a parsimonious four-variable clinical model and IL-8. ROC, receiver-operating characteristic curve.



**Figure 2.** | Prediction of mortality using a parsimonious four-variable clinical model and IL-8. ROC, receiver-operating characteristic curve.

not significantly different from the more complex models. In Table 4 we can see that the addition of IL-8 or IL-10 results in improvement in IDI, CFNRI, and calibration over the reference mortality model ( $P < 0.05$ ). We then reset the reference mortality model to include IL-8 to see whether IL-8 was the major predictive driver, and we found that adding IL-10 to a mortality model containing IL-8 does not aid in the predictive ability of the model. We also performed sensitivity analyses where we reclassified recovery as any time before or including 60 days and found results similar to those already reported: The addition of all biomarkers, or combinations thereof,

do not add to the predictive ability of the baseline clinical models (data not shown).

## Discussion

In this study, we found that the baseline clinical models and individual plasma biomarkers were similar with respect to overall predictive ability for renal recovery and mortality. Adding plasma IL-8 to a more complex eight-variable LASSO clinical model, including the APACHE II score, had similar predictive value for renal recovery and only slightly superior predictive value for mortality compared with a four-variable parsimonious clinical model with IL-8. The simple four-variable model that included age, mean arterial pressure, mechanical ventilation, and bilirubin, along with plasma IL-8, had modest predictive value (AUROC, 0.76 for renal recovery and 0.78 for mortality). To our knowledge, this study is the first large study to examine risk prediction for outcomes after severe AKI using a panel of biomarkers in a large cohort of critically ill patients receiving RRT.

Our findings are important for several reasons. First, the four-variable clinical model along with plasma IL-8 could be used to estimate patient prognosis and clinical decision-making by nephrologists and intensivists. For instance, IL-8 marker levels measured on the day when RRT initiation is being considered by a clinician could be useful to better inform patients and families about prognosis. Because patients who are older, are mechanically ventilated, and have lower mean arterial pressure, high bilirubin, and IL-8 levels, are likely to have worse outcomes, they may not wish to undergo RRT. Although IL-8 is not yet used in clinical practice to treat patients with AKI, the assay is easy to perform, with rapidly available results, and could easily be developed into a clinical test.

Second, identification of a homogeneous group of patients using biomarker-guided risk assessment allows for examination of new interventions or interventions that have previously failed in clinical trials that included a heterogeneous population of patients with severe AKI. Third, while studies have examined various clinical models and biomarkers in severe AKI, none have evaluated risk prediction of renal recovery in severely ill patients using simple clinical models. Finally, early prediction of renal recovery is likely to be helpful with regard to postdischarge monitoring of renal function after critical illness and subsequent progression to CKD and ESRD in patients who are unlikely to have complete renal recovery (13).

In prior work using the ATN dataset, Demirjian *et al.* (4) found that the ATN model consisting of 21 clinical variables was superior (AUROC, 0.85) to existing severity of illness scoring systems for mortality prediction. While such complex models can be used to compare health system performances, they are very difficult for busy clinicians to use at the bedside. In contrast, the combination of the four-variable clinical model and plasma IL-8, despite its modest predictive value, is similar to a more complex eight-variable LASSO model and can potentially be used at the bedside to make patient care decisions. We also found that an eight-variable LASSO model consisting of age, mechanical ventilation, arterial pH, bilirubin, mean arterial pressure, platelets, APACHE II score, and oliguria,

along with IL-8, predicted mortality. However, variables such as arterial pH require an arterial blood sample, and APACHE II requires a complex calculation using many subvariables, which makes it difficult for routine clinical use.

The predictive value of individual markers was only modest and lower in comparison with the clinical models for renal recovery and mortality. Biomarkers remaining after two different selection methods (stepwise and LASSO) were similar (IL-6, IL-8, and IL-10 for mortality and IL-8, IL-10, and TNFR I for recovery) but not identical to those most strongly associated with poor outcomes in our prior work (1). Of note, lack of augmented predictive ability of biomarkers does not imply that these markers are not associated with poor outcomes. IL-8 is a chemokine implicated in the pathogenesis of AKI and higher plasma IL-8 concentrations may reflect a persistent proinflammatory state and multisystem organ failure predicting higher mortality.

Our study has several important limitations. First, we selected candidate biomarkers on the basis of biologic plausibility within inflammation and apoptosis. We did not include more recent novel urinary biomarkers, such as neutrophil gelatinase-associated lipocalin (14), kidney injury molecule-1 (15), or cell-cycle arrest markers (16,17) because the ATN trial did not collect urine samples (3). In a small subset of the BioMaRK study cohort, we previously examined a panel of urinary markers and found that day 1 results were of similar predictive value (AUROC, 0.51–0.66) to the markers examined in this study (15). However, by day 14, results were significantly more predictive, especially when added to the clinical model.

Second, we did not examine local (*i.e.*, renal tissue or urinary) concentrations of these markers, which could have improved risk prediction compared with plasma concentrations. Third, because we used an imputation algorithm for missing data, it is possible that some of the associations and their respective magnitudes may depend on the imputation model used (Supplemental Table 2). Fourth, risk prediction models always demonstrate best calibration in the population in which they are generated; thus, further validation of the model in other cohorts of patients with AKI will be required before it can be widely applied to clinical practice.

Our study also has several strengths. Compared with the ATN mortality prediction model, we used separate model building and validation data folds as opposed to building and validating the model on the same dataset. We chose the most parsimonious model to reduce the number of variables that can be useful to bedside clinicians. We added novel prediction metrics (IDI, NRI, and calibration) to aid in determining the most parsimonious clinical model and the usefulness of adding available plasma biomarkers to the baseline model.

In summary, our results show that in critically ill patients receiving RRT, a simple four-variable clinical model consisting of age, mean arterial pressure, mechanical ventilation, and bilirubin, together with IL-8, augments risk prediction for renal recovery and mortality at day 60 and could potentially be useful at the bedside for clinicians. Although we did a cross-validation within our study cohort, our findings require external validation before they can be applied to other patient populations.

## Acknowledgments

We are indebted to the nurses, respiratory therapists, phlebotomists, physicians, and other health care professionals who participated in the ATN and BioMaRK studies, as well as the patients and their families.

The BioMaRK study was conducted by the BioMaRK investigators and supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). The primary data and samples from the parent trial (A Comparison of Intensive vs. Conventional Renal Support in Acute Renal Failure; the ATN study), from which the analyses reported here were performed, were supplied by the Cooperative Studies Program of the Department of Veterans Affairs Office of Research and Development (CSP #530) and by the NIDDK Central Repositories. This manuscript does not necessarily reflect the opinions or views of the ATN study, NIDDK Central Repositories, or the NIDDK. The VA/NIH ATN study was supported by the Cooperative Studies Program of the Department of Veterans Affairs Office of Research and Development (CSP #530) and by NIDDK by interagency agreement Y1-DK-3508.

## Disclosures

None.

## References

- Murugan R, Wen X, Shah N, Lee M, Kong L, Pike F, Keener C, Unruh M, Finkel K, Vijayan A, Palevsky PM, Paganini E, Carter M, Elder M, Kellum JA; Biological Markers for Recovery of Kidney (BioMaRK) Study Investigators: Plasma inflammatory and apoptosis markers are associated with dialysis dependence and death among critically ill patients receiving renal replacement therapy. *Nephrol Dial Transplant* 29: 1854–1864, 2014
- Kellum JA, Devarajan P: What can we expect from biomarkers for acute kidney injury? *Biomarkers Med* 8: 1239–1245, 2014
- Palevsky PM, Zhang JH, O'Connor TZ, Chertow GM, Crowley ST, Choudhury D, Finkel K, Kellum JA, Paganini E, Schein RM, Smith MW, Swanson KM, Thompson BT, Vijayan A, Watnick S, Star RA, Peduzzi P; VA/NIH Acute Renal Failure Trial Network: Intensity of renal support in critically ill patients with acute kidney injury. *N Engl J Med* 359: 7–20, 2008
- Demirjian S, Chertow GM, Zhang JH, O'Connor TZ, Vitale J, Paganini EP, Palevsky PM; VA/NIH Acute Renal Failure Trial Network: Model to predict mortality in critically ill adults with acute kidney injury. *Clin J Am Soc Nephrol* 6: 2114–2120, 2011
- Charlson ME, Pompei P, Ales KL, MacKenzie CR: A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *J Chronic Dis* 40: 373–383, 1987
- Knaus WA, Draper EA, Wagner DP, Zimmerman JE: APACHE II: A severity of disease classification system. *Crit Care Med* 13: 818–829, 1985
- Paganini EP, Halstenberg WK, Goormastic M: Risk modeling in acute renal failure requiring dialysis: The introduction of a new model. *Clin Nephrol* 46: 206–211, 1996
- Tibshirani R: Regression shrinkage and selection via the lasso. *J R Stat Soc, B* 58: 267–288, 1996
- DeLong ER, DeLong DM, Clarke-Pearson DL: Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* 44: 837–845, 1988
- Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS: Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 27: 157–172, discussion 207–212, 2008
- Youden WJ: Index for rating diagnostic tests. *Cancer* 3: 32–35, 1950
- Kutner MH, Nachtsheim CJN: Applied Linear Regression Models, 4th Ed., New York, McGraw-Hill Irwin, 2004
- Chawla LS, Eggers PW, Star RA, Kimmel PL: Acute kidney injury and chronic kidney disease as interconnected syndromes. *N Engl J Med* 371: 58–66, 2014

14. Srisawat N, Wen X, Lee M, Kong L, Elder M, Carter M, Unruh M, Finkel K, Vijayan A, Ramkumar M, Paganini E, Singbartl K, Palevsky PM, Kellum JA: Urinary biomarkers and renal recovery in critically ill patients with renal support. *Clin J Am Soc Nephrol* 6: 1815–1823, 2011
15. Endre ZH, Pickering JW, Walker RJ, Devarajan P, Edelstein CL, Bonventre JV, Frampton CM, Bennett MR, Ma Q, Sabbisetti VS, Vaidya VS, Walcher AM, Shaw GM, Henderson SJ, Nejat M, Schollum JBW, George PM: Improved performance of urinary biomarkers of acute kidney injury in the critically ill by stratification for injury duration and baseline renal function. *Kidney Int* 79: 1119–1130, 2011
16. Kashani K, Al-Khafaji A, Ardiles T, Artigas A, Bagshaw SM, Bell M, Bihorac A, Birkhahn R, Cely CM, Chawla LS, Davison DL, Feldkamp T, Forni LG, Gong MN, Gunnerson KJ, Haase M, Hackett J, Honore PM, Hoste EA, Joannes-Boyau O, Joannidis M, Kim P, Koyner JL, Laskowitz DT, Lissauer ME, Marx G, McCullough PA, Mullaney S, Ostermann M, Rimmelé T, Shapiro NI, Shaw AD, Shi J, Sprague AM, Vincent J-L, Vinsonneau C, Wagner L, Walker MG, Wilkerson RG, Zacharowski K, Kellum JA: Discovery and validation of cell cycle arrest biomarkers in human acute kidney injury. *Crit Care* 17: R25, 2013
17. Bihorac A, Chawla LS, Shaw AD, Al-Khafaji A, Davison DL, Demuth GE, Fitzgerald R, Gong MN, Graham DD, Gunnerson K, Heung M, Jortani S, Kloorup E, Koyner JL, Krell K, Letourneau J, Lissauer M, Miner J, Nguyen HB, Ortega LM, Self WH, Sellman R, Shi J, Straseski J, Szalados JE, Wilber ST, Walker MG, Wilson J, Wunderink R, Zimmerman J, Kellum JA: Validation of cell-cycle arrest biomarkers for acute kidney injury using clinical adjudication. *Am J Respir Crit Care Med* 189: 932–939, 2014

**Received:** October 6, 2014 **Accepted:** April 27, 2015

Published online ahead of print. Publication date available at [www.cjasn.org](http://www.cjasn.org).

This article contains supplemental material online at <http://cjasn.asnjournals.org/lookup/suppl/doi:10.2215/CJN.09911014/-/DCSupplemental>.