CHAPTER 6

Kidney-Specific Severity Scores

Shigehiko Uchino

OBJECTIVES

This chapter will:
1. Review published kidney-specific severity scores.
2. Evaluate accuracy of severity scores for acute renal failure.
3. Summarize risk factors for hospital death in patients with acute renal failure.

Despite continuing progress in medical treatment, acute renal failure (ARF) in critical illness carries a hospital mortality of more than 60%.[1] Several randomized controlled trials have been unsuccessful in decreasing such mortality.[2,3] One of the difficulties with the conduct of clinical trials in ARF is that there is no reliable scoring system to stratify patient selection and confirm balanced randomization. General severity scores do not reliably predict outcome of patients with ARF,[4] partly because data from only a few patients with ARF were collected to generate these scores. Multiple epidemiologic studies have looked at risk factors for hospital mortality in patients with ARF,[1,5–14] and in some of these studies, kidney-specific severity scores have been published.[15–27] However, these scores also have problems and limitations, mainly because they are based on small populations.

In this chapter, published kidney-specific severity scores and their problems are reviewed. Subsequently, major risk factors for hospital mortality in patients with ARF are discussed to provide information for the future development of more accurate kidney-specific severity scores.

KIDNEY-SPECIFIC SEVERITY SCORES AND THEIR EXTERNAL VALIDATION

Multiple kidney-specific severity scores have been published in the literature; they are listed in Table 6.1. Most of these scores were developed in single centers or, if multicenter, in single countries. Although two studies were conducted in a multicenter setting,[23,27] they were originally randomized controlled studies. One was for atrial natriuretic peptide for ARF,[28] and the sample size was smaller than that of other, later studies.[24–26] The other was for intensity of renal replacement therapy (RRT); thus all patients included for score derivation were treated with RRT.[29]

The following ARF scores are based on larger samples than others, have been often validated externally, or have been published recently.

Bullock's Score

Bullock et al.[16] generated their score with data from a population of 462 patients with ARF who were admitted to a single center in the United States from January 1971 to January 1978. ARF was diagnosed when a serum creatinine (SCr) value of 2.5 mg/dL or greater and/or a blood

TABLE 6.1

Multiple Kidney-Specific Severity Scores

STUDIES REPORTING KIDNEY-SPECIFIC SEVERITY SCORES								
STUDY (YEAR)	COUNTRY(IES)	NUMBER OF CENTERS	NUMBER OF PATIENTS	POPULATION	REQUIREMENT FOR RENAL REPLACEMENT THERAPY (%)	HOSPITAL MORTALITY (%)	NUMBER OF VARIABLES	EXTERNAL VALIDATION STUDIES*
Cioffiet al. (1984)[15]	US	1	65	Surgical	100	81	8	29
Bullock et al. (1985)[16]	US	1	462	Hospital	62	68	6	29
Rasmussen et al. (1985)[17]	Australia	1	148	Hospital	—	53	10	30
Lohr et al. (1988)[18]	US	1	126	Hospital	100	75	5	29, 30
Schaefer et al. (1991)[19]	Germany	1	134	Intensive care unit (ICU)	100	57	6	24, 30
Liaño et al. (1993)[20]	Spain	1	328	Hospital	51	53	9	20, 24-26, 30, 33
Barton et al. (1993)[21]	UK	1	250	ICU	100	51	5	—
Paganini et al. (1996)[22]	US	1	512	ICU	100	67	8	23, 24, 26, 33
Chertow et al. (1998)[23]	US, Canada	48	256	Hospital	42	36	9	24, 33
Mehta et al. (2002)[24]	US	4	605	ICU	50	52	9	33
SHARF-II (2004)[25]	Belgium	8	293	ICU	37	51	8	24
PICARD (2006)[26]	US	5	618	ICU	64	—	5-8	—
Demirjian et al. (2011)[27]	US	27	1122	ICU	100	53% (60-day)	21	

*Superscript numbers are chapter references.

urea nitrogen (BUN) concentration of 100 mg/dL or greater was found during the clinical course. Patients with chronic renal insufficiency were included if their SCr values rose by at least 2.5 mg/dL over baseline levels. Bullock's score was calculated as follows:

$$\begin{aligned}\text{Log odds of death} = &-1.765 - 0.687 \times (\text{CP1} + 0.037) + 0.822 \\ &\times (\text{CP2} + 0.100) + 1.053 \times ([\text{pulmonary complications}] \\ &- 0.087) + 0.050 \times (\text{age} - 61.1) + 0.7 \times ([\text{jaundice}] + 0.143) \\ &+ 0.608 \times ([\text{cardiovascular complications}] - 0.247) \\ &+ 0.365 \times ([\text{hypercatabolism}] + 0.0303)\end{aligned}$$

where *CP1* and *CP2* are categorical variables for nonoliguria and anuria, respectively (e.g., a patient with anuria would have CP1 = 0 and CP2 = 1).

Although this score was published two decades ago, data from more than 400 patients were used to generate it. However, only one external validation has been conducted so far, making it difficult to evaluate the score's accuracy. Halstenberg et al.[30] validated three kidney-specific severity scores (Lohr's,[18] Bullock's, and Cioffi's[15]) and the second Acute Physiology and Chronic Health Evaluation (APACHE-II) score[31] through the use of a registry of 512 patients who received acute dialysis at their institution between 1988 and 1992. Using the same population, this group generated another kidney-specific severity score.[22] When they tested the Bullock's score in their population, the Q statistic was only 20%, compared with 77% in the original population. (Q values greater than 50% are associated with good discriminatory power.)

Liaño's Score

The population that Liaño et al.[20] used to generate their score consisted of 328 patients with ARF who were admitted to a single hospital in Madrid from November 1977 to June 1988. ARF was diagnosed when a sudden rise in SCr to more than 2 mg/dL was found in subjects with prior normal renal function as documented by an SCr value less than 1.5 mg/dL. No patients with previous reduction in renal function were included. The researchers used two models of multiple regression analysis, linear and logistic, comparing the models' ability to predict hospital mortality. They found that the linear regression model's receiver operating characteristic (ROC) curve was better than that for the logistic regression model, so they adopted the equation from linear regression as their final model of the score, as follows:

$$\begin{aligned}\text{Probability of death} = &(0.032 \times \text{age in decades}) \\ &- (0.086 \times [\text{male}]) - (0.109 \times [\text{nephrotoxic}]) \\ &+ (0.109 \times [\text{oliguria}]) + (0.116 \times [\text{hypotension}]) \\ &+ (0.122 \times [\text{jaundice}]) + (0.150 \times [\text{coma}]) \\ &- (0.154 \times [\text{consciousness}]) \\ &+ (0.182 \times [\text{assisted respiration}]) + 0.210\end{aligned}$$

Liaño et al.[20] found that there were no survivors above a discriminant score of 0.9, and at this level, the sensitivity of the score was 28%, the specificity 100%, and the positive predictive value 100%. Finally, they applied the score to 25 patients in another hospital in Spain for external validation; in this population as well, there were no survivors above a discriminant score of 0.9.

Of all the kidney-specific severity scores published, Liaño's score has been externally validated in other populations most frequently. For example, Douma et al.[32] retrospectively examined 238 patients who were treated with renal RRT for ARF in their unit. They evaluated several general severity scores and four kidney-specific severity scores (Rasmussen's,[17] Lohr's,[18] Schaefer's,[19] and Liaño's), reporting that Liaño's score had the highest area under the receiver operating curve (AUROC) (Rasmussen's,[17] 0.63;

Lohr's,[18] 0.65; Schaefer's,[19] 0.69; Liaño's, 0.78).[32] (AUROC values greater than 0.7 are associated with good discriminatory power.) They also found that patients in the highest quintile of Liaño's score had a near-100% mortality (98%). However, Mehta et al.[24] and Chertow et al.[26] reported much lower AUROC values for Liaño's score, 0.630 and 0.53 to 0.56, respectively.

Mehta's Score

Mehta et al.[24] generated their kidney-specific severity score from a population of 605 patients who had undergone nephrology consultation for ARF in intensive care units (ICUs) at four hospitals in Southern California between October 1989 and September 1995. ARF was defined either as a BUN value greater than 40 mg/dL or as an SCr value greater than 2 mg/dL for patients with no prior history of kidney disease. For patients with preexisting renal insufficiency, ARF was defined as a sustained rise in SCr value of more than 1 mg/dL over baseline level. Multiple logistic regression analysis was used to generate the score as follows:

$$\begin{aligned}\text{Log odds of death} = {} & (0.0170 \times \text{age}) + (0.8605 \times [\text{male}]) \\ & + (0.0144 \times \text{BUN}) - (0.3398 \times \text{SCr}) + (1.2242 \\ & \times [\text{hematologic failure}]) + (1.1183 \times [\text{liver failure}]) \\ & + (0.9637 \times [\text{respiratory failure}]) + (0.0119 \times \text{heart rate}) \\ & - (0.4432 \times \textit{log}(\text{urine output})) - 0.7207\end{aligned}$$

Using the same population, these researchers compared the AUROC[33] and Hosmer-Lemeshow goodness-of-fit[34] of their score with those of several general severity scores and kidney-specific severity scores (Schaefer's,[19] Liaño's, Paganini's,[22] PICARD,[26] and SHARF-I/SHARF-II [see later]). Mehta's score had the best discrimination (AUROC 0.832) and calibration ability of all scores tested. However, the Beginning and Ending Supportive Therapy for the Kidney (BEST Kidney) study[35]—a multinational epidemiologic study of ARF conducted at 54 centers in 23 countries and involving more than 1700 patients—found that none of the scores tested, including Mehta's score, had good discrimination or calibration ability (AUROC for Mehta's score, 0.670).

SHARF-II Score

The SHARF-II score is a modified version of a previously published kidney-specific severity score, Stuivenberg Hospital Acute Renal Failure (SHARF-I). Originally, Lins et al.[36] generated a score on the basis of data from 197 patients treated in a single center in Belgium from March 1996 to April 1997. ARF was defined as an SCr value more than 2 mg/dL or an increase in SCr value of more than 50% observed in patients with previous mild-to-moderate chronic renal failure. Data were collected at study inclusion (T_0) and 48 hours later (T_{48}), and two scores were generated for each data collection point (SHARF-I_0 and SHARF-I_{48}). Both of these scores involved the same five variables—age, albumin level, prothrombin time, mechanical ventilation, and heart failure. Both scores showed good discrimination ability (AUROC 0.87 and 0.89, respectively) and good calibration ability (goodness-of-fit C *P* values 0.83 and 0.28, respectively) in the study population.

When Lins et al.[25] reevaluated validation of these scores in eight ICUs using the same inclusion criteria (from September 1997 to March 1998, 293 patients); however, they found that the discrimination ability of the original SHARF scores was not as strong as in the original population (AUROC 0.67 and 0.78, respectively). They also found three additional variables, bilirubin, sepsis, and hypotension, related to hospital mortality. Therefore they generated new scores (SHARF-II) at T_0 and T_{48}. For example, SHARF-II at T_0 (SHARF-T_0) is as follows:

$$\begin{aligned}\text{SHARF-II}_0 = {} & (3.0\,\&\,\text{ZeroWidthSpace}; \times \text{age in decades}) \\ & + (2.6 \times [\text{albumin category}]) \\ & + (1.3 \times [\text{prothrombin category}]) \\ & + (16.8 \times [\text{mechanical ventilation}]) \\ & + (3.9 \times \text{heart failure}]) + (2.8 \times [\text{bilirubin}]) \\ & + (27 \times [\text{sepsis}]) + (21 \times [\text{hypotension}]) - 17\end{aligned}$$

AUROC values for these new scores (AUROC 0.82 and 0.83, respectively) were significantly better than those of the original scores. SHARF II_0 was validated externally by Mehta et al.,[24] who found that it showed fairly good discrimination ability (AUROC 0.733) but bad calibration ability (goodness-of-fit *p* value .03).

PICARD Score

The Program to Improve Care in Acute Renal Disease (PICARD) score was published as a substudy of the PICARD study, which was conducted in five centers in the United States from February 1999 to August 2001 and involved 618 patients.[26] ARF was defined as an increase in SCr value of more than 0.5 mg/dL in patients with baseline values lower than 1.5 mg/dL or an increase in SCr of more than 1.0 mg/dL in patients with baseline SCr values greater than 1.5 mg/dL but less than 5.0 mg/dL. Patients with baseline SCr values greater than 5.0 mg/dL were not included. Using multiple logistic regression analysis, Chertow and associates[26] created three scores: on day of ARF diagnosis, on day of consultation, and on day of first RRT. For example, the score on day of consultation had eight variables, including adult respiratory distress syndrome (ARDS), as follows:

$$\begin{aligned}\text{Log odds of death} = {} & (0.1241 \times \text{age in decades}) \\ & - (0.2063 \times \log \text{urine output}) \\ & + (0.6900 \times [\text{SCr} < 2\ \text{mg/dL}]) \\ & + (0.0828 \times \text{BUN per 10 mg/dL}) \\ & + (0.4811 \times [\text{liver failure}]) \\ & + (0.5800 \times [\text{ARDS}]) + (0.5074 \\ & \times [\text{platelet count } 150 \times 10^6 \times \text{L}]) \\ & + (0.4083 \times [\text{sepsis}]) - 1.2563\end{aligned}$$

Using the study population, these researchers compared the AUROC value of their score with that of several general severity scores and kidney-specific severity scores (Liaño's, Paganini's,[22] and SHARF-I). Surprisingly, the AUROC of their score was only 0.68 at day of consultation, which was lower than that of APACHE-III,[37] the second Simplified Acute Physiology Score (SAPS-II),[38] and the Sepsis-Related Organ Failure Assessment (SOFA) score[39] (AUROC values for all, 0.70). Ohnuma et al., using the Japanese Society for Physicians and Trainees in Intensive Care (JSEPTIC) database, which collected data retrospectively from 343 patients with ARF who required continuous renal replacement therapy (CRRT) in 14 ICUs, assessed several ARF severity scores, including the PICARD score.[40] The AUROC for the PICARD score was only 0.64.

Demirjian's Score

This newest kidney-specific severity score was generated from the largest population in the literature.[27] The score

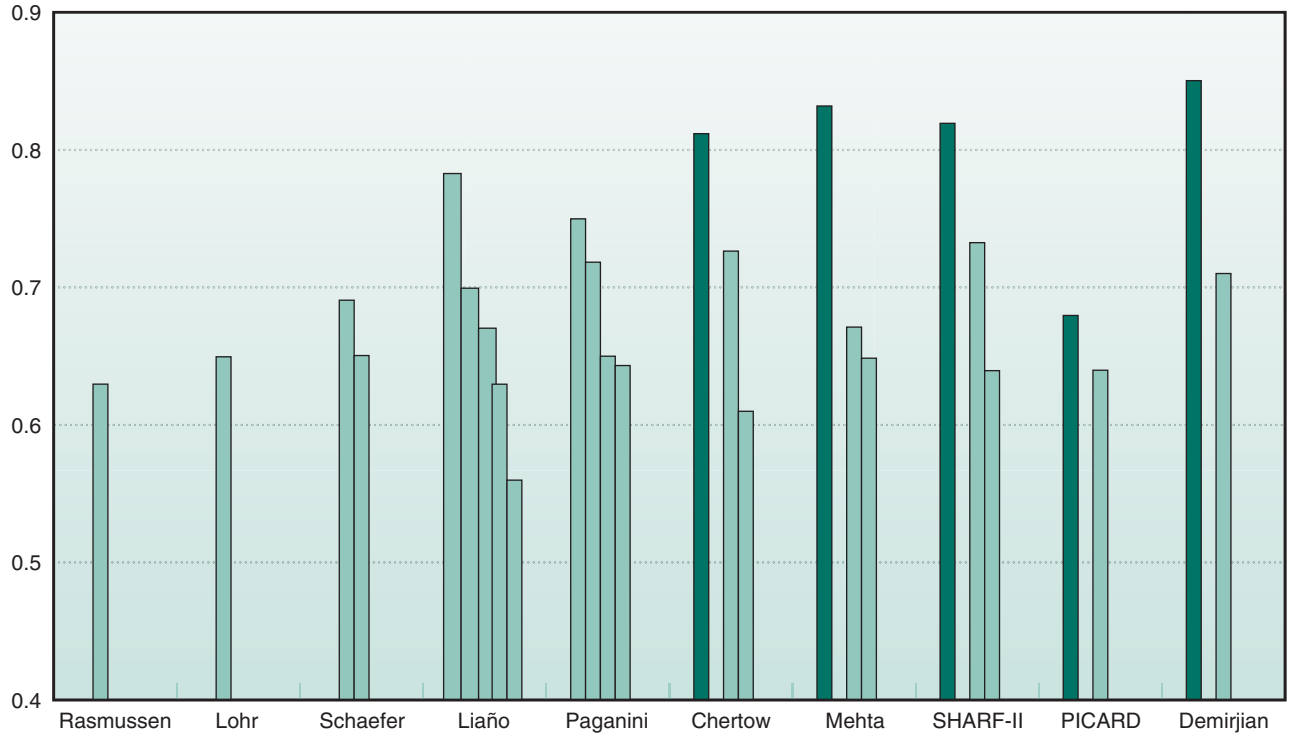

FIGURE 6.1 Area under the receiver operating curve (AUROC) values for mortality reported in the original studies and external validation studies. Dark green columns indicate AUROC values in original articles, and light green columns are from external validation studies. SHARF, Stuivenberg Hospital Acute Renal Failure; PICARD, Program to Improve Care in Acute Renal Disease.

was generated using data from 1122 subjects enrolled in a multicenter randomized trial of intensive versus less intensive renal support in critically ill patients requiring RRT conducted between November 2003 and July 2007 at 27 centers.[29] Patients were included if they had a clinical diagnosis of ARF requiring dialysis because of acute tubular necrosis. Patients with chronic kidney disease (defined as premorbid serum creatinine 2 mg/dL in men and 1.5 mg/dL in women) or prior kidney transplantation were excluded. Twenty-one independent predictors of 60-day mortality were identified and the logistic regression model using these variables had an AUROC of 0.85. An integer risk score also was created, with an AUROC of 0.80. This newest score was validated externally by Ohnuma et al. using the JSEPTIC database, which found an AUROC of 0.71, the highest among the scores they tested.[40]

PROBLEMS OF CURRENTLY AVAILABLE KIDNEY-SPECIFIC SEVERITY SCORES

For general severity scores—SAPS-II,[38] APACHE-II,[31] and APACHE-III[37]—several external validation studies have been conducted.[41,42] These studies have found that general severity scores have good discrimination ability in different settings, with an AUROC value greater than 0.8. However, their calibration abilities were not as good as their discrimination abilities, and recalibration to fit these scores to each center or country has been recommended.[42]

Kidney-specific severity scores have not reached this level of assessment. External validation studies have shown that no kidney-specific severity score has good calibration or discrimination ability.[31,34] Fig. 6.1 shows reported AUROC values for ARF scores in the literature. Among five scores that reported AUROC values in the original studies, four of them reported AUROC values exceeding 0.8, indicating good discrimination. However, AUROC values reported in external validation studies for these scores were lower. AUROC values for other scores in external validation studies were also low. No score had an AUROC value greater than 0.8, and the values were often lower than 0.7 in the external validation studies, suggesting poor discrimination ability (see Fig. 6.1).

One of the major reasons for such a difference between general and kidney-specific severity scores is size of population. General severity scores were based on multicenter, multinational databases involving more than 5000 patients. Most of the kidney-specific scores were generated from populations in one center, and none of them involved more than 1200 patients. Therefore a large database collected from multiple centers in many nations would be required to generate more accurate severity scores for ARF.

Risk Factors for Hospital Death in Patients With Acute Renal Failure

Because of differences in case mix, collected variables, sample size, and statistical power in the studies, reported risk factors for hospital death in patients with ARF are variable. To determine common and relevant risk factors, Table 6.2 lists such factors that have been reported in more than two epidemiologic studies that used multivariate regression analysis. The most frequently reported risk factor is age,

TABLE 6.2

Risk Factors Reported in More Than Two Epidemiologic Studies Using Multivariate Regression Analysis

RISK FACTORS	CHAPTER REFERENCES
Demographics	
Age	1,7,9,11,12,14–16,20,21,24–27
General severity scores	1,7,9,10,12–14
Male gender	15,22–24
Development of acute renal failure after admission	1,7,12
Comorbidity	
Previous health status scores	7,9,13
Preexisting heart disease	17,18,27
Diagnosis	
Sepsis/septic shock	1,5,7,8,13,18,25,26
(Hematologic) malignancy	1,6,17
Surgical patients	12,15,17,27
Organ Failure	
Oliguria	7–9,12,16,17,20,21,23,24,26,27
Hypotension	18–20,25,27
Respiratory failure	16,17,24,26,27
Heart failure	1,16,25
Laboratory Data	
High bilirubin	16,20–23,25,27
Low creatinine	21,22,24,26,27
High urea	22,24,26
Low albumin	8,23,25,27
(Metabolic) acidosis	10,21,23,27
Low platelet count	22,26,27
Treatment	
Mechanical ventilation	1,6,10–12,18–23,25,27
Vasoactive medication	1,10,11,21
Renal replacement therapy requirement	9,11,13

followed by mechanical ventilation, oliguria, sepsis/septic shock, and high serum bilirubin value. When hypotension and vasoactive medication are combined as a single risk factor, this factor also becomes common. Most of these risk factors are included in general severity scores as well, except for sepsis/septic shock.

Sepsis has been reported to be a leading precipitant of ARF, with 50% to 70% of cases of ARF being related to sepsis.[1,13] Hoste et al.,[11] observing 185 septic patients in a surgical ICU, found that 16.2% of the patients had ARF and 70% of those patients required RRT. Bagshaw et al.[43] reported that, using the BEST kidney database, sepsis was considered the cause in 47.5% of study patients, and that septic ARF was associated with greater aberrations in hemodynamics and laboratory parameters, greater severity of illness, and higher need for mechanical ventilation and vasoactive therapy. Septic ARF also had a higher hospital mortality compared with the nonseptic control (70.2 vs. 51.8%; $p < .001$). Therefore sepsis is an important condition for development of ARF and hospital mortality. Any researchers planning to develop new kidney-specific severity scores should probably include sepsis as a variable.

Several risk factors are related to renal function: oliguria, low SCr value, high BUN value, and RRT requirement. Among them, low SCr value seems to be unique because general severity scores usually classify higher SCr value as a risk factor.[31,37,38] Low SCr value could be related to diminished muscle mass or hemodilution resulting from volume overload. Therefore low SCr value with high BUN value (marker of renal dysfunction) can be found as an independent variable when both values are entered in multiple regression analysis. This is the case in the three out of four studies that identified low SCr value as a risk factor.[22,24,26] In the fourth study, reported by Barton et al.,[21] the association between low SCr value and poor prognosis was attributed to the clinical decision by ICU care providers to request early RRT for sicker patients. For this reason, Barton et al.[21] eliminated serum SCr value from their final model of the severity score. When two variables with strong colinearity are entered in multiple regression analysis, they weaken each other's explanatory powers, and estimating their separate effects can be difficult. Therefore it does not seem to be correct to enter both low SCr value and high BUN value in multiple regression analysis, because these two variables have obviously strong colinearity. To evaluate SCr or BUN as a possible variable for a severity score, SCr value adjusted for body weight and/or gender may be useful.

Several studies reported that development of ARF after ICU or hospital admission was related to mortality. For example, the BEST Kidney researchers found that the odds ratio of duration between hospital admission and study inclusion, in 1-day increments, was as high as 1.02 ($p < .001$).[1] A possible explanation for the relationship between development of ARF and poor prognosis is that patients in whom ARF developed after hospital or ICU admission had a worsened clinical condition despite supportive therapy.

It may be intuitive that RRT requirement is a marker of severe renal dysfunction, which should relate to higher hospital mortality.[9,11,13] However, of the six studies generating ARF severity scores that included patients who were not treated with RRT, none included RRT requirement in their scores. Possible explanations are that some patients were too sick to receive RRT or that RRT was not given because of expected poor prognosis. For example, in the BEST Kidney study, hospital mortality rates for patients with and without RRT were 62.1% and 55.7%, respectively ($p = .021$).[1] This slight difference was eliminated by multiple regression analysis, thus removing RRT requirement as a risk factor.

Key Points

1. Currently available kidney-specific severity scores have good discrimination and calibration ability in their study populations.
2. This accuracy is not demonstrated when the scores are validated externally.
3. The discrepancy may be due to the small sample size.
4. A large multinational database will be required to generate a more precise kidney-specific severity score.

Key References

20. Liaño F, Gallego A, Pascual J, et al. Prognosis of acute tubular necrosis: An extended prospectively contrasted study. *Nephron.* 1993;63:21-31.
24. Mehta RL, Pascual MT, Gruta CG, et al. Refining predictive models in critically ill patients with acute renal failure. *J Am Soc Nephrol.* 2002;13:1350-1357.

25. Lins RL, Elseviers MM, Daelemans R, et al. Re-evaluation and modification of the Stuivenberg Hospital Acute Renal Failure (SHARF) scoring system for the prognosis of acute renal failure: An independent multicentre, prospective study. *Nephrol Dial Transplant.* 2004;19:2282-2288.
26. Chertow GM, Soroko SH, Paganini EP, et al. Mortality after acute renal failure: Models for prognostic stratification and risk adjustment. *Kidney Int.* 2006;70:1120-1126.
27. Demirjian S, Chertow GM, Zhang JH, et al. Model to predict mortality in critically ill adults with acute kidney injury. *Clin J Am Soc Nephrol.* 2011;6:2114-2120.

A complete reference list can be found online at ExpertConsult.com.

References

1. Uchino S, Kellum JA, Bellomo R, et al. Beginning and Ending Supportive Therapy for the Kidney (BEST Kidney) Investigators. Acute renal failure in critically ill patients: A multinational, multicenter study. *JAMA*. 2005;294:813-818.
2. Hirschberg R, Kopple J, Lipsett P, et al. Multicenter clinical trial of recombinant human insulin-like growth factor I in patients with acute renal failure. *Kidney Int*. 1999;55:2423-2432.
3. Lewis J, Salem MM, Chertow GM, et al. Atrial natriuretic factor in oliguric acute renal failure. Anaritide Acute Renal Failure Study Group. *Am J Kidney Dis*. 2000;36:767-774.
4. Fiaccadori E, Maggiore U, Lombardi M, et al. Predicting patient outcome from acute renal failure comparing three general severity of illness scoring systems. *Kidney Int*. 2000;58: 283-292.
5. Corwin HL, Teplick RS, Schreiber MJ, et al. Prediction of outcome in acute renal failure. *Am J Nephrol*. 1987;7:8-12.
6. Chertow GM, Christiansen CL, Cleary PD, et al. Prognostic stratification in critically ill patients with acute renal failure requiring dialysis. *Arch Intern Med*. 1995;155:1505-1511.
7. Brivet FG, Kleinknecht DJ, Loirat P, et al. Acute renal failure in intensive care units—causes, outcome, and prognostic factors of hospital mortality: A prospective, multicenter study. French Study Group on Acute Renal Failure. *Crit Care Med*. 1996;24:192-198.
8. Obialo CI, Okonofua EC, Nzerue MC, et al. Role of hypoalbuminemia and hypocholesterolemia as copredictors of mortality in acute renal failure. *Kidney Int*. 1999;56:1058-1063.
9. Guerin C, Girard R, Selli JM, et al. Initial versus delayed acute renal failure in the intensive care unit: A multicenter prospective epidemiological study. Rhone-Alpes Area Study Group on Acute Renal Failure. *Am J Respir Crit Care Med*. 2000;161:872-879.
10. Metnitz PG, Krenn CG, Steltzer H, et al. Effect of acute renal failure requiring renal replacement therapy on outcome in critically ill patients. *Crit Care Med*. 2002;30:2051-2058.
11. Hoste EA, Lameire NH, Vanholder RC, et al. Acute renal failure in patients with sepsis in a surgical ICU: Predictive factors, incidence, comorbidity, and outcome. *J Am Soc Nephrol*. 2003;14:1022-1030.
12. Ostermann ME, Chang RW. Riyadh ICU Program Users Group. Prognosis of acute renal failure: An evaluation of proposed consensus criteria. *Intensive Care Med*. 2005;31:250-256.
13. Bagshaw SM, Laupland KB, Doig CJ, et al. Prognosis for long-term survival and renal recovery in critically ill patients with severe acute renal failure: A population-based study. *Crit Care*. 2005;9:700-709.
14. du Cheyron D, Parienti JJ, Fekih-Hassen M, et al. Impact of anemia on outcome in critically ill patients with severe acute renal failure. *Intensive Care Med*. 2005;31:1529-1536.
15. Cioffi WG, Ashikaga T, Gamelli RL. Probability of surviving postoperative acute renal failure: Development of a prognostic index. *Ann Surg*. 1984;200:205-211.
16. Bullock ML, Umen AJ, Finkelstein M, et al. The assessment of risk factors in 462 patients with acute renal failure. *Am J Kidney Dis*. 1985;5:97-103.
17. Rasmussen HH, Pitt EA, Ibels LS, et al. Prediction of outcome in acute renal failure by discriminant analysis of clinical variables. *Arch Intern Med*. 1985;145:2015-2018.
18. Lohr JW, McFarlane MJ, Grantham JJ. A clinical index to predict survival in acute renal failure patients requiring dialysis. *Am J Kidney Dis*. 1988;11:254-259.
19. Schaefer JH, Jochimsen F, Keller F, et al. Outcome prediction of acute renal failure in medical intensive care. *Intensive Care Med*. 1991;17:19-24.
20. Liaño F, Gallego A, Pascual J, et al. Prognosis of acute tubular necrosis: An extended prospectively contrasted study. *Nephron*. 1993;63:21-31.
21. Barton IK, Hilton PJ, Taub NA, et al. Acute renal failure treated by haemofiltration: Factors affecting outcome. *Q J Med*. 1993;86:81-90.
22. Paganini EP, Halstenberg WK, Goormastic M. Risk modeling in acute renal failure requiring dialysis: The introduction of a new model. *Clin Nephrol*. 1996;46:206-211.
23. Chertow GM, Lazarus JM, Paganini EP, et al. Predictors of mortality and the provision of dialysis in patients with acute tubular necrosis. The Auriculin Anaritide Acute Renal Failure Study Group. *J Am Soc Nephrol*. 1998;9:692-698.
24. Mehta RL, Pascual MT, Gruta CG, et al. Refining predictive models in critically ill patients with acute renal failure. *J Am Soc Nephrol*. 2002;13:1350-1357.
25. Lins RL, Elseviers MM, Daelemans R, et al. Re-evaluation and modification of the Stuivenberg Hospital Acute Renal Failure (SHARF) scoring system for the prognosis of acute renal failure: An independent multicentre, prospective study. *Nephrol Dial Transplant*. 2004;19:2282-2288.
26. Chertow GM, Soroko SH, Paganini EP, et al. Mortality after acute renal failure: Models for prognostic stratification and risk adjustment. *Kidney Int*. 2006;70:1120-1126.
27. Demirjian S, Chertow GM, Zhang JH, et al. Model to predict mortality in critically ill adults with acute kidney injury. *Clin J Am Soc Nephrol*. 2011;6:2114-2120.
28. Allgren RL, Marbury TC, Rahman SN, et al. Anaritide in acute tubular necrosis. Auriculin Anaritide Acute Renal Failure Study Group. *N Engl J Med*. 1997;336:828-834.
29. VA/NIH Acute Renal Failure Trial Network, Palevsky PM, Zhang JH, et al. Intensity of renal support in critically ill patients with acute kidney injury. *N Engl J Med*. 2008;359:7-20.
30. Halstenberg WK, Goormastic M, Paganini EP. Validity of four models for predicting outcome in critically ill acute renal failure patients. *Clin Nephrol*. 1997;47:81-86.
31. Knaus WA, Draper EA, Wagner DP, et al. APACHE II: A severity of disease classification system. *Crit Care Med*. 1985;13:818-829.
32. Douma CE, Redekop WK, van der Meulen JH, et al. Predicting mortality in intensive care patients with acute renal failure treated with dialysis. *J Am Soc Nephrol*. 1997;8:111-117.
33. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143:29-36.
34. Lemeshow S, Hosmer DW Jr. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol*. 1982;115:92-106.
35. Uchino S, Bellomo R, Morimatsu H, et al. Beginning and Ending Supportive Therapy for the Kidney (BEST Kidney) investigators. External validation of severity scoring systems for acute renal failure using a multinational database. *Crit Care Med*. 2005;33:1961-1967.
36. Lins RL, Elseviers M, Daelemans R, et al. Prognostic value of a new scoring system for hospital mortality in acute renal failure. *Clin Nephrol*. 2000;53:10-17.
37. Knaus WA, Wagner DP, Draper EA, et al. The APACHE III prognostic system: Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest*. 1991;100:1619-1636.
38. Le Gall J-R, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on an European/North American multicentre study. *JAMA*. 1993;270:2957-2963.
39. Vincent JL, de Mendonça A, Cantraine F, et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: Results of a multicenter, prospective study. Working group on "sepsis-related problems" of the European Society of Intensive Care Medicine. *Crit Care Med*. 1998;26:1793-1800.
40. Ohnuma T, Uchino S, Toki N, et al, JSEPTIC (Japanese Society of Education for Physicians and Trainees in Intensive Care) Clinical Trial Group. External validation for acute kidney injury severity scores: a multicenter retrospective study in 14 Japanese ICUs. *Am J Nephrol*. 2015;42:57-64.
41. Metnitz PG, Lang T, Vesely H, et al. Ratios of observed to expected mortality are affected by differences in case mix and quality of care. *Intensive Care Med*. 2000;26:1466-1472.
42. Harrison DA, Brady AR, Parry GJ, et al. Recalibration of risk prediction models in a large multicenter cohort of admissions to adult, general critical care units in the United Kingdom. *Crit Care Med*. 2006;34:1378-1388.
43. Bagshaw SM, Uchino S, Bellomo R, et al. Beginning and ending supportive therapy for the kidney (BEST kidney) investigators. Septic acute kidney injury in critically ill patients: clinical characteristics and outcomes. *Clin J Am Soc Nephrol*. 2007;2:431-439.