Validation of a Clinical Scoring System for Outcome Prediction in Dogs with Acute Kidney Injury Managed by Hemodialysis

G. Segev, C. Langston, K. Takada, P.H. Kass, and L.D. Cowgill

Background: A scoring system for outcome prediction in dogs with acute kidney injury (AKI) recently has been developed but has not been validated.

Hypothesis: The scoring system previously developed for outcome prediction will accurately predict outcome in a validation cohort of dogs with AKI managed with hemodialysis.

Animals: One hundred fifteen client-owned dogs with AKI.

Methods: Medical records of dogs with AKI treated by hemodialysis between 2011 and 2015 were reviewed. Dogs were included only if all variables required to calculate the final predictive score were available, and the 30-day outcome was known. A predictive score for 3 models was calculated for each dog. Logistic regression was used to evaluate the association of the final predictive score with each model's outcome. Receiver operating curve (ROC) analyses were performed to determine sensitivity and specificity for each model based on previously established cut-off values.

Results: Higher scores for each model were associated with decreased survival probability ($P < .001$). Based on previously established cut-off values, 3 models (models A, B, C) were associated with sensitivities/specificities of 73/75%, 71/80%, and 75/86%, respectively, and correctly classified 74–80% of the dogs.

Conclusions and Clinical Relevance: All models were simple to apply and allowed outcome prediction that closely corresponded with actual outcome in an independent cohort. As expected, accuracies were slightly lower compared with those from the previously reported cohort used initially to develop the models.

Key words: Acute renal failure; Hemodialysis; Outcome; Prediction; Survival; Urinary tract.

Acute kidney injury (AKI) leading to severe uremia is associated with high morbidity and mortality.[1,2] Most dogs with AKI are managed medically, but when the disease is severe or the kidney dysfunction is prolonged, the clinical and pathologic consequences of AKI may no longer be managed effectively with conventional medical therapy. Hemodialysis is an advanced extracorporeal renal replacement therapy for uremic and overhydrated patients, with the capability to remove uremic toxins, correct fluid and electrolyte imbalance, restore acid-base balance, and sustain the life of the patient until the kidney injury resolves.[3] As a consequence, hemodialysis extends life expectancies of patients with severe uremia and expands the window of opportunity for recovery and the potential for a favorable outcome. Nonetheless, the mortality of patients with AKI managed with hemodialysis remains as high as 50%.[3]

The availability of renal replacement therapies in veterinary medicine is growing, and it has become the advanced standard of care for animal patients with AKI. However, it remains a costly therapy with limited regional availability and may not be recommended in the light of the unknown prognosis for a favorable outcome. Therefore, it would be valuable to have a prognostic tool that could assist with outcome prediction of animals with AKI when contemplating management with hemodialysis.

Scoring systems are used to assess objectively the severity of diseases and aid clinicians and owners when making decisions regarding the likelihood of survival before electing a complicated and expensive therapeutic option. The information required to assess disease severity and potential for favorable outcome must be available within the first hours of hospitalization, when therapeutic decisions must be made. Recently, we developed a statistically derived scoring system for dogs with AKI managed with hemodialysis.[3] This scoring system was based on clinically available variables obtained at admission from 182 dogs with AKI managed by hemodialysis at 1 tertiary referral center.[3] The scoring

From the Koret School of Veterinary Medicine, Hebrew University of Jerusalem, Israel (Segev); the The Ohio State University Veterinary Medical Center, (Langston); the Department of Medicine and Epidemiology, School of Veterinary Medicine, University of California, Davis, CA (Takada, Cowgill); and the Department of Population Heath and Reproduction, University of California, Davis, CA (Kass).

The animal patients were treated at the William R. Pritchard Veterinary Medical Teaching Hospital, University of California, Davis, USA; the Animal Medical center, New York, NY, USA; The Koret School of Veterinary Medicine, the Hebrew University of Jerusalem, Israel.

Corresponding author: G. Segev, Koret School of Veterinary Medicine, Hebrew University of Jerusalem, Israel, P.O. Box 12, Rehovot 76100, Israel; e-mail: gilad.segev@mail.huji.ac.il.

Abbreviations:

AKI	acute kidney injury
ANOVA	analysis of variance
AUC	area under the curve
CBC	complete blood count
OR	odds ratio
ROC	receiver operating curve

algorithms evaluated clinical signs and clinicopathologic abnormalities available on the first day of hospitalization. From these inputs, a final predictive score was used to assess the probability of 30-day survival. Variables associated with survival were evaluated further by assigning weighting factors to each of the variables based on the odds ratio (OR) relating survival to deviations of each of the variables from their respective reference ranges or reference categories. For each dog, all weighting factors were summed to calculate an individual final predictive score. Four prototype models were developed, of which 3 performed well (models A–C). In model A, the OR for survival for each variable was rounded to an integer value; in model B, the exact OR was used; and in model C, the etiology (if known) also was incorporated into the model. Models A–C correctly classified outcomes in 81–87% of the dogs, with corresponding sensitivities/specificities of 77/85%, 81/85%, and 83/90%, respectively.

Because this AKI scoring system was developed and validated using the same patient population enrolled from a single referral center, it might not perform as well if applied to a more diverse patient population. The aim of this study was to evaluate further and validate the prototype scoring system in an independent cohort of patients.

Material and Methods

Patients and Data Collection

Medical records of dogs with AKI presented to the teaching hospitals of the University of California, Davis, CA; the Animal Medical Center, New York City, NY; and the Koret School of Veterinary Medicine, Hebrew University, Jerusalem, between 2011 and 2015 and managed with hemodialysis but not used in the initial development and validation of the prototype AKI scoring system were reviewed.

Acute kidney injury was defined using the same criteria used in the previous study[3]: (1) acute onset of consistent clinical signs and history (eg, anuria, oliguria, vomiting, diarrhea, inappetence); (2) renal azotemia (serum creatinine concentration >3 mg/dL and urine-specific gravity <1.025); and (3) normal or enlarged kidney size (relative to the dog's body weight) as detected by ultrasound examination. Dogs with historical azotemia or imaging findings (eg, small irregular kidneys, decreased corticomedullary distinction) at presentation consistent with chronic kidney disease were excluded, as were dogs with postrenal azotemia.

Data extracted from the medical records included only variables required to calculate the final predictive score of the previously developed scoring system. Surviving dogs were defined as those remaining nondialysis dependent for at least 30 days after discharge from the hospital. Dogs were excluded from the analysis if euthanized within 2 weeks after initiation of hemodialysis because of financial considerations but dogs euthanized ≥2 weeks after initiation of dialysis because of lack of improvement were considered nonsurvivors.

Laboratory Findings

Blood and urine specimens were collected for CBC, serum biochemistry profile, and urinalysis at admission, and analyses were performed using established methods at the diagnostic laboratories of the respective institutions. Urine production was measured during the first 24 hours of hospitalization.

Organ System Involvement and Etiology

Etiology was classified when known as leptospirosis, ethylene glycol intoxication, or other as previously described.[3] Extra-renal organ involvement was classified based on clinical signs, laboratory abnormalities, radiological or ultrasonographic findings, or some combination of these as previously described.[3]

Calculation of Final Predictive Score

Formulas for the prototype predictive models A, B, and C have been described previously.[3] The clinical value for each required variable was compared with previously established ranges for assignment of the weighting factors for each dog's calculation. All weighting factors were added to produce a final predictive score for each dog. Higher scores corresponded with a poorer prognosis for survival.

Statistical Analysis

Descriptive statistics are presented as mean ± standard deviation or median and range based on evaluation of the data distribution under the normality assumption using the Shapiro-Wilk test. One-way analysis of variance (ANOVA) was used to compare final predicative scores among institutions. Proportions were compared using the chi-square test. Logistic regression was used to assess the relationship between the final predictive score and the probability of 30-day survival. Receiver operating curve analysis was performed to determine the area under the curve (AUC) and to calculate sensitivity and specificity for survival at different cut-off values. The optimal cut-off value was chosen to minimize outcome misclassification. Statistical analyses were performed using statistical software.[a] For all tests, $P < .05$ was considered statistically significant.

Results

Results of clinicopathologic variables assessed at presentation and used to calculate final predictive scores in the validation cohort are shown in Table 1. Median urine production was 0.9 mL/kg/h (range 0–9.9 mL/kg/h). Thirty-four dogs (30%) were anuric during the initial 24 hours after presentation.

Table 1. Mean and standard deviation (SD) of variable used to calculate the final predictive score of 115 dogs with acute kidney injury managed by hemodialysis.

Variable	Mean	SD
Red blood cells ($\times 10^6$ cells/μL)	5.0	1.4
Lymphocyte count ($\times 10^3$ cells/μL)	1.58	1.20
Creatinine (mg/dL)	10.0	4.5
Phosphorus (mg/dL)	13.4	5.7
Ionized calcium (mmol/L)	1.1	0.2
Anion gap (mmol/L)	25.9	9.3
Albumin (g/dL)	2.4	0.5
Alanine aminotransferase (U/L)	271	588

Fifty-two dogs (45%) had respiratory involvement at presentation, 11 dogs (9%) had neurologic involvement, and 26 dogs (23%) fulfilled the criteria for a diagnosis of disseminated intravascular coagulation (DIC). The overall survival proportion was 51.3% (59/115 dogs).

Outcome scores ranged from 12 to 33 (mean, 19.7 ± 4.1) in model A, from 12.2 to 36.3 (mean, 21.4 ± 4.90) in model B, and from 5.3 to 37.3 (mean, 20.4 ± 6.8) in model C. There were no statistically significant differences in the mean final predictive score for any of the models among the different institutions. Higher scores were significantly associated with decreased probability of survival in all models ($P < .001$).

The AUC of the ROC analyses for models A, B, and C were 0.80, 0.81, and 0.85, respectively (Fig 1, Table 2). The optimal cut-off point for the current cohort for model A was 20, with predicted sensitivity and specificity of 75% and 73%, respectively, and cor-
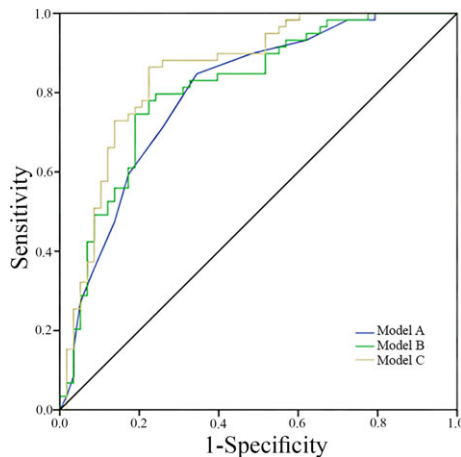

Fig 1. Receiver operating curve (ROC) for Models A, B, and C. Models A, B, and C had associated AUC values equal to 0.80, 0.81, and 0.85, respectively.

Table 2. Area under the receiver operating curve for 3 models for prediction of survival of dogs with acute kidney injury managed with hemodialysis.

Model	AUC	SE	Interval of AUC	P value
Model A	0.80	0.042	0.72–0.88	<.001
Model B	0.81	0.041	0.73–0.88	<.001
Model C	0.85	0.037	0.78–0.92	<.001

AUC, area under the curve; SE, standard error.

rectly classified survival outcome in 85 dogs (74%). The optimal cut-off point for model B was 21.5, with predicted sensitivity and specificity of 77% and 80%, respectively, and correctly classified survival outcome in 90 dogs (78%). The optimal cut-off point for Model C was 21.8 with predicted sensitivity and specificity of 77% and 87%, respectively, and correctly classified survival outcome in 94 dogs (82%).

The previously established (prototype) cut-off value for model A was 20. This cut-off value was identical to the cut-off value for the validation cohort. Therefore, the sensitivity and specificity predictions did not change for this model. The prototype optimal cut-off value for model B established in the previous study[3] was 20.5. This cut-off value generated slightly different sensitivity and specificity predictions of 80% and 68%, respectively, when applied to the validation cohort and correctly classified survival outcome in 74% of the dogs. The prototype optimal cut-off value established for model C was 19.9. This cut-off value resulted in sensitivity and specificity predictions of 86% and 75%, respectively, in the validation cohort and correctly classified survival outcome in 80% of dogs.

Additional sensitivity and specificity values generated from different cut-off values used to increase or to decrease sensitivity or specificity predictions are presented in Table 3.

Discussion

A validated clinical scoring system with the potential to predict outcomes effectively for patients requiring renal replacement therapy can have considerable utility to categorize the severity, prognosis, and categorization of AKI. We have described a prototype scoring system based on clinical variables statistically correlated with disease outcomes.[3] The accuracy reported for this scoring system potentially was flawed by testing its performance on the same patient population used for its generation. This study was designed to validate independently the performance and accuracy of the scoring system in a novel and diverse population of dogs undergoing hemodialysis for AKI.

The results of our study indicate that the overall performance of the prototype scoring system was relatively good at predicting survival outcome, when tested on an independent validation cohort, despite being marginally lower than previously reported. The overall mortality rate for dogs with AKI in this study was comparable to reports in human patients with AKI managed with hemodialysis (32–52.5%)[4] and in previous reports of

Table 3. Selected cut-off values and associated sensitivity and specificity predictions for scoring models for survival in of dogs with acute kidney injury managed by hemodialysis.

Model A			Model B			Model C		
Cut-off	Sensitivity	Specificity	Cut-off	Sensitivity	Specificity	Cut-off	Sensitivity	Specificity
16.0	32	99	18.8	51	91	13.7	34	95
20.0	75	73	20.5	71	80	19.9	71	80
25.0	98	21	26.0	98	30	24.6	97	43

dogs with AKI managed medically (56–62%)[1,2] or with hemodialysis (56%).[3] The comparable mortalities in dogs managed with hemodialysis likely reflect the greater severity of kidney injury in these dogs compared with those treated medically. This is reflected by the higher mean serum creatinine concentration for dogs managed by hemodialysis.[3]

The prognosis of dogs with AKI managed by hemodialysis is affected by several variables including disease severity, number of organs involved, and reversibility of kidney injury, which is influenced directly by its severity and etiology.[5] Etiology is a major determinant of survival, presumably because of the potential of the kidney to recover. The 2 prognostically most important etiologies in the tested cohorts were ethylene glycol intoxication, which is a relatively irreversible injury,[6,7] and leptospirosis, which is readily reversible.[1,8,9] As expected, and model C (that incorporated etiology as a variable in the scoring system) was the most accurate model in both the previous and current evaluations. However, etiology often is unknown at presentation when the decision to initiate hemodialysis must be made and often remains unknown throughout the course of the disease.

The relatively strong influence of etiology on outcome in these studies suggests that the statistical relevance and development of weighting factors for other etiologies relevant to other clinical settings should be evaluated for incorporation into the scoring algorithm. This could influence the importance and accuracy of AKI scoring in other geographic regions where different etiologies predominate or as the causes of AKI in dogs change over time with emerging diseases, new therapeutic agents, and changes in practice patterns. Only 6% of the dogs in the validation cohort were diagnosed with ethylene glycol intoxication compared with 27% in the previous study, and only 23% were diagnosed with leptospirosis compared with 31% in the previous study. Increased awareness by owners of ethylene glycol intoxication and the introduction of multivalent leptospirosis vaccination, as well as regional differences in the validation cohort, likely decreased the prevalence of these 2 etiologies and may have contributed to the slight differences in performance and accuracy compared with the previous evaluation. Assuming these trends persist and the prevalence of leptospirosis and ethylene glycol intoxication decrease further over time, the value of these etiologies in prognostic modeling will decrease and be replaced by more relevant etiologic variable or reliance on scoring models without etiology variables. Models A and B, which do not incorporate etiology as part of the model, were associated with 73% and 76% correct classification of survival outcome, respectively, demonstrating objective assessment of survival probability still is achievable exclusive of etiology. The clinical signs and clinicopathologic abnormalities included in the scoring system may have served as proxies for the severity of the disease as well as its etiology and potential for repair.

Using area under the ROC curves as an indicator of model performance, the scoring system applied to the validation cohort of dogs compared favorably with the previous evaluation in which performance was tested on the same dogs used for its formulation. It also performed similarly to other scoring systems developed for human patients and animals.[10–13] The validation cohort of dogs was derived from 3 geographically distinct referral centers, of which 2 differed from the center used in the previous study. Nevertheless, the score ranges for all models and the optimal cut-off values were almost identical to those reported previously. Their comparability suggests that the severity of kidney injury is similar among dogs with AKI managed by hemodialysis in both studies and suggests that the scoring system will have utility in other independent populations of dogs presenting similarly for evaluation and management.

A concern with this or any scoring system is the relatively high misclassification arising when the final predictive score is close to the cut-off value. The scoring system is most accurate when the final predictive score is either low or high and is more likely to misclassify outcome when the score is close to the cut-off point. When the score is close to the cut-off point, a small change in 1 of the clinical values has potential to move the final predictive score above or below the cut-off and alter the predicted outcome. This also is an inherent limitation of diagnostic tests using defined cut-off values to rule in or rule out a disease. A way to overcome this limitation is to define a "gray zone" of scores for which predictions and recommendation are avoided, especially when outcome predictions could discourage owners from pursuing medical options. For the current scoring system, the sensitivity and specificity of the models increased when predictions were not made when the final predictive score was within a defined "gray zone" (Table 3). Clinicians also may elect to use higher or lower cut-off values to further emphasize sensitivity or specificity predictions. By increasing the cut-off value, sensitivity will increase (higher correct prediction that an animal will survive), but this will be at the expense of decreased specificity (higher false prediction that an animal will not survive).

The outcomes of dogs managed by hemodialysis are affected by treatment duration, which in turn often is limited by financial constraints. Repair and regeneration of the injured kidney and compensatory hypertrophy may persist over several months, and owners may not be able financially to maintain their dogs on hemodialysis for the prolonged time required to facilitate recovery and achieve successful outcome. However, when financial constraints do not limit treatment duration, the prognosis for survival may be better than forecasted by the scoring system predicting outcome only for the first 30 days after presentation.

Scoring systems always should be applied judiciously and with caution in individual patients, where misclassification can have detrimental consequences. Despite the relatively high performance of this scoring system, 20–26% of the dogs were misclassified using models in which etiology was unknown, and 20% were misclassified even when etiology was known. Therefore, scoring systems alone cannot supplant proper clinical assessment or serve as a sole prognostic tool, but rather should be incorporated as part of the overall assessment.

The proposed scoring system has been tailored for dogs requiring hemodialysis and should not be used in dogs with AKI that are managed medically until first tested and validated in that patient population. Prognostic factors in AKI dogs managed medically likely differ from those of dogs managed by hemodialysis. For example, hyperkalemia was not included in the current scoring systems because, once hemodialysis is initiated, dogs are unlikely to die from consequences of hyperkalemia. Conversely, hyperkalemia is more likely to influence the prognosis of dogs managed medically.

A scoring system like that used in our study also can be used to objectively compare or classify disease severity among different patient populations. The final predictive scores were comparable among all 3 institutions, suggesting hemodialysis was initiated in patients with equivalent degrees of AKI severity. This comparability is likely due to the uniformity of guidelines used for dialytic intervention employed by these 3 centers.

There are several limitations to this study. First, the number of dogs included was relatively low compared with the numbers used in the development and the evaluation of scoring systems used in humans, which can reach thousands.[10–12] Second, dogs euthanized ≥2 weeks after initiation of hemodialysis because of lack of improvement were considered nonsurvivors. Some of these dogs might have recovered if hemodialysis had been continued for a longer period of time. Nonetheless, in veterinary medicine, financial constraints are more likely to influence advanced therapeutic decisions and treatment duration than are favorable outcome predictions, which may be linked to costly or extended therapies. Scoring systems developed for veterinary patients may be more relevant and practical if directed at predicting outcomes within weeks as opposed to months.

In summary, scoring systems can be used as an additional tool to aid clinicians in the overall assessment of dogs with AKI, managed by hemodialysis.

Footnote

Acknowledgments

References

1. Behrend EN, Grauer GF, Mani I, et al. Hospital-acquired acute renal failure in dogs: 29 cases (1983-1992). J Am Vet Med Assoc 1996;208:537–541.

2. Vaden SL, Levine J, Breitschwerdt EB. A retrospective case-control of acute renal failure in 99 dogs. J Vet Intern Med 1997;11:58–64.

3. Segev G, Kass PH, Francey T, et al. A novel clinical scoring system for outcome prediction in dogs with acute kidney injury managed by hemodialysis. J Vet Intern Med 2008;22:301–308.

4. McCarthy JT. Prognosis of patients with acute renal failure in the intensive-care unit: A tale of two eras. Mayo Clin Proc 1996;71:117–126.

5. Segev G. Prognosis of acute renal failure. In: Bartges J, Polzin D, eds. Nephrology and Urology of Small Animals. UK: Blackwell Publishing; 2012:524–527.

6. Connally HE, Thrall MA, Forney SD, et al. Safety and efficacy of 4-methylpyrazole for treatment of suspected or confirmed ethylene glycol intoxication in dogs: 107 cases (1983-1995). J Am Vet Med Assoc 1996;209:1880–1883.

7. Rollings CE, Francey T, Cowgill LD. Use of hemodialysis in uremic and non-uremic dogs with ethylene glycol toxicity. J Vet Intern Med 2004;18:416.

8. Adin CA, Cowgill LD. Treatment and outcome of dogs with leptospirosis: 36 cases (1990-1998). J Am Vet Med Assoc 2000;216:371–375.

9. Birnbaum N, Barr SC, Center SA, et al. Naturally acquired leptospirosis in 36 dogs: Serological and clinicopathological features. J Small Anim Pract 1998;39:231–236.

10. Knaus WA, Draper EA, Wagner DP, et al. APACHE II: A severity of disease classification system. Crit Care Med 1985;13:818–829.

11. 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.

12. Knaus WA, Zimmerman JE, Wagner DP, et al. APACHE-acute physiology and chronic health evaluation: A physiologically based classification system. Crit Care Med 1981;9:591–597.

13. Segev G, Aroch I, Savoray M, et al. A novel severity scoring system for dogs with heatstroke. J Vet Emerg Crit Care (San Antonio) 2015;25:240–247.