Received: 28 April 2022 | Revised: 25 July 2022 | Accepted: 3 September 2022

DOI: 10.1111/vec.13350

ORIGINAL STUDY

Journal of Veterinary Emergency and Critical Care

WILEY

Outcome prediction in dogs admitted through the emergency room: Accuracy of staff prediction and comparison with an illness severity stratification system for hospitalized dogs

Alice Le Gal DVM, MVetMed, DACVECC, DECVECC | **Dominic Martin Barfield BSc, BVSc, MVetMed, DACVECC, DECVECC** | **Roseanne Helen Wignall RVN** | **Simon David Cook BSc, BVSc, MVetMed, DACVECC, DECVECC**

Section of Emergency and Critical Care, Department of Clinical Science and Services, The Royal Veterinary College, Hatfield, UK

Correspondence
Alice Le Gal, Department of Clinical Science and Services, The Royal Veterinary College, Hawkshead Lane, North Mymms, Hatfield, Hertfordshire AL9 7TA, UK.
Email: alegal7@rvc.ac.uk

Presented in part as an abstract at the 2021 European Veterinary Emergency and Critical Care Congress.

Abstract

Objective: To determine whether emergency staff and students can predict patient outcome within 24 hours of admission, comparing the accuracy of clinician prognostication with outcome prediction by Acute Patient Physiologic and Laboratory Evaluation (APPLE)$_{fast}$ scoring and identifying whether experience or mood would be associated with accuracy.

Design: Prospective observational study between April 2020 and March 2021.

Setting: University teaching hospital.

Animals: One hundred and sixty-one dogs admitted through an Emergency Service were assessed. Where data were available, an APPLE$_{fast}$ score was calculated per patient. An APPLE$_{fast}$ score of >25 was deemed a predictor for mortality.

Interventions: None.

Measurements and Main Results: Emergency staff and students were asked to complete surveys about dogs admitted to the emergency room. All clinicopathological data were available for review, and the animals were available for examination. Data collected included opinions on whether the patient would be discharged from hospital, a mood score, position, and experience in Emergency and Critical Care. One-hundred and twenty-five dogs (77.6%) were discharged; 36 dogs (22.4%) died or were euthanized. Two hundred and sixty-six responses were obtained; 202 responses (75.9%) predicted the correct outcome. Students, interns, residents, faculty, and nurses predicted the correct outcome in 81.4%, 58.3%, 83.3%, 82.1%, and 65.5% of cases, respectively. Of 64 incorrect predictions, 43 (67.2%) predicted death in hospital. APPLE$_{fast}$ scores were obtained in 121 cases, predicting the correct outcome in 83 cases (68.6%). Of 38 cases in which APPLE$_{fast}$ was incorrect, 27 (71.1%) were dogs

Abbreviations: APPLE, Acute Patient Physiologic and Laboratory Evaluation; ECC, Emergency and Critical Care; ER, emergency room; MBI, Maslach Burnout Inventory.

surviving to discharge. Mean $APPLE_{fast}$ score was 22.9 (± 6.2). There was no difference in outcome prediction accuracy between staff and $APPLE_{fast}$ scores ($P = 0.13$). Neither experience nor mood score was associated with outcome prediction ability ($P = 0.55$ and $P = 0.74$, respectively).

Conclusions: Outcome prediction accuracy by staff is not significantly different to $APPLE_{fast}$ scoring where a cutoff of >25 is used to predict mortality. When predictions were incorrect, they often predicted nonsurvival.

KEYWORDS
canine, critical illness, nonsurvival, prognostication, survival

1 | INTRODUCTION

In people, physicians make clinical decisions based on their judgement, experience, and care available in the hospital. A study evaluating criteria used by physicians to triage patients into the ICU identified patient-specific criteria (eg, age, underlying disease, and self-sufficiency) and also physician-specific criteria such as their seniority and ability to examine the patient.[1] Another study showed that physicians were more likely to withdraw mechanical ventilation, inotropes or vasopressors, and dialysis if they had predicted a less than 10% probability of ICU survival.[2]

Objective scoring systems and models have been evaluated as tools to predict mortality more accurately. These include the Acute Physiology and Chronic Health Evaluation (APACHE) score, the Simplified Acute Physiology Score (SAPS), and the Mortality Probability Model (MPM0) score. When compared to different scoring systems, physicians' predictions of mortality have been found to be more accurate[1,3–5]; however, both have only a moderate ability to discriminate between survivors and nonsurvivors. ICU nurses have also been shown to be able to accurately predict patient outcome and mortality,[6] and to be more accurate than scoring systems.[5]

Prognostic assessments of veterinary patients influence management decisions and may be relied upon by owners for decision-making on concepts such as euthanasia. This is particularly true in emergent situations or when the animal is critically ill. Staff may not necessarily rely on quantified risks or literature to make prognostications and these can also be a source of debate among veterinary staff.

Scoring systems have been developed in veterinary medicine (or adapted from people) to objectively assess illness severity and have predictive validity for hospitalized patients.[7–9] These include severity of illness scores such as the Sequential Organ Failure Assessment (SOFA) score and the Acute Patient Physiologic and Laboratory Evaluation (APPLE) score. The APPLE score exists as 10-variable model ($APPLE_{full}$ containing creatinine, WBC count, albumin, SpO_2 as detected by pulse oximetry, total bilirubin, mentation score, respiratory rate [bpm], age [years], lactate, and presence of free fluid in a body cavity as detected by ultrasonographic screening) and as a 5-variable model ($APPLE_{fast}$ containing glucose, albumin, mentation score, platelet count, and lactate).[8] It has been shown that an $APPLE_{fast}$ score cutoff of >25 had a specificity of 85% and a sensitivity of 67% in predicting mortality in hospitalized patients.[8]

These objective scores may support clinical judgement, benchmark hospital performance for quality assurance, assist in research, provide quantitative measures of illness to guide clinicians in determining diagnostic and therapeutic recommendations, and advise pet owners regarding resources required, but their accuracy compared to clinicians' prognostication has not been investigated.

To the authors' knowledge, there are no studies evaluating the accuracy of staff predictions in the veterinary profession. The objectives of this study were to evaluate the accuracy of veterinary staff and students' predictions regarding patient outcome and mortality compared to the $APPLE_{fast}$ score. Secondary objectives were to assess whether experience, mood, and number of sequential days worked would affect the accuracy of staff predictions. We hypothesized that veterinary staff would be able to accurately predict patient outcome and would be more accurate than the $APPLE_{fast}$ score, that experience would increase the accuracy of predictions, and that the more time between admission and prediction, the more accurate the prediction would be. Finally, we hypothesized that the more days staff had worked, the worse their mood would be, and therefore that they would be more pessimistic about patient outcome.

2 | MATERIALS AND METHODS

Staff and students working in the emergency room (ER) and ICU of a university veterinary teaching hospital were asked to complete a survey (Appendix S1) evaluating dogs admitted within the previous 24 hours. The dogs were selected as a convenience sample: the survey was open between April 2020 and March 2021, and completion was voluntary and anonymous. Any dog admitted to the hospital and hospitalized in the ER or the ICU could be evaluated. The surveys could be completed at 2 timepoints: in the 6 hours postadmission or between 12 and 24 hours postadmission. Interns, residents, faculty, and nurses working in the Emergency and Critical Care (ECC) department were eligible to fill in the survey as well as undergraduate veterinary students on their clinical rotations within ECC. In the hospital where the study was carried out, interns working in the ECC department were

rotating small animal interns who were qualified veterinarians with at least 2 years of experience in clinical practice. Residents who were asked to participate were enrolled in European or American Board of Veterinary Specialist training programs. Each participant could fill out up to 2 surveys per animal (1 in the first 6 h and 1 in the 12–24 h postadmission period) and could assess as many dogs as they wanted during the study period. All medical records and the opportunity to examine the patient were available to the assessor. Staff and students were asked for their opinion on the patient's discharge outcome (discharged vs death or euthanasia in hospital), but also their qualitative outcome (positive vs. negative; where a positive outcome was defined as discharged home healed, on supportive care, or following a treatment plan, and a negative outcome was defined as discharged for home euthanasia, discharged against medical advice, discharged for short-term palliative care, euthanasia, or death) and their 30-day survival (alive or dead 30 days post discharge). Surveys were excluded if they were incomplete or if they were completed outside of the study timepoints.

Respondents were then asked a series of questions to gauge their level of experience in ECC (number of years worked in a specialists-lead ECC department), their mood at the time (from 0 = *very unwell* to 10 = *very well*), and an objective assessment of depersonalization according to the Maslach Burnout Inventory (MBI).[10] The 5 components of the depersonalization score measure a loss of consideration toward other people (colleagues, owners, or patients) or a loss of empathy. A high score denotes a lack of feeling or an impersonal feeling toward patient care or outcome and more generally a higher degree of experienced burnout. The depersonalization score is considered low if it is less than 5, moderate if between 6 and 9, and high if it is 10 or more.[11] The MBI has been validated to measure burnout in medical students.[12,13] Respondents were asked to select the criteria they had used to make their prediction with an option for free text insertion if they had used an unlisted criterion. Finally, staff and students were asked if they had thought about the dog's outcome prior to making a prediction in the survey and if they had shared their opinion with their colleagues or the owners.

Where data were available, an $APPLE_{fast}$ score was calculated in the first 6 hours. An $APPLE_{fast}$ score of more than 25 was used as a predictor of mortality. This study received ethical approval by the university's Clinical Research Ethical Review Board (SR2020-0199).

2.1 | Statistical methods

Data were assessed for normality using the Shapiro–Wilk test and presented as mean (± standard deviation) for parametric data. Prediction accuracies were calculated descriptively as percentages (where accuracy is equal to the sum of [i] the number of cases in which mortality was correctly predicted and [ii] the number of cases in which survival was correctly predicted, divided by [iii] the total number of cases for which a prediction was made). *T*-tests were used to examine whether experience, number of days worked, depersonalization score, and mood score were significantly associated with the accuracy of the predictions. Chi-squared or Fisher's exact tests were used to assess whether the assessor's position, the time between admission and assessment, and if being the clinician in charge of the case were associated with the accuracy of the predictions and to compare the accuracy of staff predictions with that of the $APPLE_{fast}$ score. Differences were considered significant at a *P*-value of <0.05. Data were analyzed using open-source software.[a]

3 | RESULTS

Two hundred and eighty-six surveys were obtained, of which 20 were excluded (14 were assessed between 6 and 12 h or after 24 h postadmission and 6 were incomplete): the 266 surveys included in the study evaluated 161 dogs. Seventy-four (28%) of responses were anonymous, and the remaining 192 responses were from 54 different respondents. The mean age of the dogs was 6.14 years (± 4.73); there were 25 intact females, 42 neutered females, 36 intact males, and 58 neutered males. Fifty-nine different breeds were represented; Labradors ($n = 13$), cocker spaniels ($n = 12$), French bulldogs ($n = 7$), and German shepherd dogs ($n = 7$) were the most commonly represented.

One hundred and eighty-one surveys (68.0%) were completed within the first 6 hours and 85 (32.0%) were completed between 12 and 24 hours following admission. One hundred and six surveys (39.8%) were completed regarding dogs hospitalized in the ER and 160 (60.2%) in the ICU. One hundred and twenty-one cases (75.2%) had data available to calculate an $APPLE_{fast}$ score: mean $APPLE_{fast}$ score was 22.9 (± 6.2); 41 dogs (33.9%) had an $APPLE_{fast}$ score >25 and 80 dogs (66.1%) had ≤25. Eighty-four percent of patients (68/81) hospitalized in the ER and 71.2% of patients (57/80) hospitalized in the ICU when assessed were discharged from hospital. Patients hospitalized in the ER were not significantly more likely to be discharged than patients hospitalized in the ICU ($P = 0.053$).

Responders were students for 43 of 266 surveys (16.2%), interns for 12 of 266 surveys (4.5%), residents for 96 of 266 surveys (36.1%), faculty members for 28 of 266 surveys (10.5%), and nurses for 87 of 266 surveys (32.7%). For all respondents, the mean experience in ECC was 4.6 years (± 4.0) and the mean number of days worked in a row was 2.51 (± 1.62). Mean mood score was 7.32 (± 1.71) and mean MBI depersonalization score was 5.49 (± 4.21) (between low and moderate).[10] Clinicians in charge of the cases assessed filled 36.8% (98/266) of surveys.

3.1 | Discharge outcome

One hundred and twenty-five dogs (77.6%) were discharged from the hospital and 36 dogs (22.4%) died or were euthanized in hospital. Two hundred and two assessors (75.9%) predicted the correct outcome and 64 (24.1%) the incorrect outcome. When the predictions were incorrect, they were more pessimistic than reality, with 43 of 64 (67.2%) wrong assessments predicting death in hospital for dogs that were discharged. Students, interns, residents, faculty, and nurses predicted the

correct outcome in 35 of 43 (81.4%), 7 of 12 (58.3%), 80 of 96 (83.3%), 23 of 28 (82.1%), and 57 of 87 (65.5%) cases, respectively. Qualified veterinarians (interns, residents, and faculty members) predicted the correct outcome in 110 of 136 (80.9%) cases. There were too few surveys from interns to perform statistical analysis on this group. The role in the team (nurse, student, resident, or faculty) was significantly associated with the accuracy of outcome prediction ($P = 0.023$) with nurses being significantly less likely to predict the correct outcome compared to veterinarians (interns, residents, and faculty members) ($P = 0.0098$). When the nurses' predictions were incorrect, they were more pessimistic than reality, with 19 of 30 (63.3%) wrong assessments predicting death in hospital for dogs that were discharged. Neither experience, number of days worked, mood, nor depersonalization score was associated with ability to predict the discharge outcome ($P = 0.55$, $P = 0.13$, $P = 0.74$, and $P = 0.88$, respectively). The number of days worked, mood, and depersonalization score were not associated with the likelihood of predicting mortality in hospital ($P = 0.12$, $P = 0.79$, and $P = 0.30$, respectively). Clinicians in charge of the cases assessed were not more likely to predict the correct outcome compared to other assessors ($P = 0.25$). Assessors were significantly more accurate at predicting patient's outcome in the first 6 hours following admission compared to the 12–24 hours postadmission ($P = 0.0085$).

The $APPLE_{fast}$ score matched the discharge outcome in 83 of 121 cases (68.6%) but did not match in 38 cases (31.4%). Of 38 cases in which $APPLE_{fast}$ was incorrect, 27 (71.1%) were patients surviving to discharge. The sensitivity and specificity of the $APPLE_{fast}$ score when using a score of greater than 25 as a predictor of mortality were 56.0% and 71.9%, respectively, in this study. Assessors were not significantly more accurate than $APPLE_{fast}$ score at predicting mortality in hospital ($P = 0.13$). Qualified veterinarians were significantly more likely to predict the correct discharge outcome than the $APPLE_{fast}$ score ($P = 0.023$).

3.2 | Qualitative outcome

Ninety-nine dogs (61.5%) had a positive outcome and 62 dogs (38.5%) had a negative outcome. Two hundred and four assessors (76.7%) predicted the correct qualitative outcome and 62 (23.3%) the incorrect outcome. Students, interns, residents, faculty, and nurses predicted the correct qualitative outcome in 33 of 43 (76.7%), 8 of 12 (66.7%), 80 of 96 (83.3%), 23 of 28 (82.1%), and 60 of 87 (69.0%) cases, respectively. There were too few surveys from interns to perform statistical analysis on this group. The role in the team (nurse, student, resident, or faculty) was not associated with the accuracy of qualitative outcome prediction ($P = 0.12$), but nurses were significantly less likely to predict the correct outcome compared to veterinarians (interns, residents, and faculty members) ($P = 0.029$). When the predictions were incorrect, assessors had tended to be more pessimistic, with 35 of 62 (56.4%) surveys predicting a poor outcome when the patient had a good outcome. Neither experience, number of days worked, mood, nor depersonalization score was associated with ability to predict the qualitative outcome ($P = 0.89$, $P = 0.06$, $P = 0.94$, and $P = 0.34$, respectively). The number of days worked, mood, and depersonalization score were not associated with the likelihood of predicting a poor outcome ($P = 0.66$, $P = 0.64$, and $P = 0.38$, respectively). Clinicians in charge of the cases were not more likely to predict the correct qualitative outcome compared to other assessors ($P = 0.25$). Assessors were significantly more accurate at predicting patient's qualitative outcome in the first 6 hours following admission compared to the 12–24 hours postadmission ($P = 0.0043$).

3.3 | Thirty-day survival

Thirty-day survival information was available for 154 dogs (95.6%) included in the study. One hundred and three dogs (66.9%) were alive 30 days postdischarge and 51 dogs (33.1%) were deceased. There were 254 surveys evaluating these 154 dogs. One hundred and ninety-seven assessors (77.6%) predicted the correct 30-day status and 57 (22.4%) the incorrect status. Students predicted the correct 30-day status in 32 of 42 cases (76.2%), interns in 7 of 11 cases (63.6%), residents in 77 of 93 cases (82.8%), faculty in 23 of 28 cases (82.1%), and nurses in 58 of 80 cases (72.5%). There were too few surveys from interns to perform statistical analysis on this group. The role in the team (nurse, student, resident, or faculty) was not associated with the accuracy of the 30-day survival prediction ($P = 0.39$). When the predictions were incorrect, assessors had tended to be more pessimistic with 30 of 57 (52.6%) surveys predicting nonsurvival at 30 days when the dog's status was alive at that time. Neither experience, number of days worked, mood, nor depersonalization score was associated with ability to predict the 30-day postdischarge outcome ($P = 0.95$, $P = 0.30$, $P = 0.51$, and $P = 0.50$, respectively). The number of days worked, mood, and depersonalization score were not associated with the likelihood of predicting nonsurvival at 30 days ($P = 0.43$, $P = 0.45$, and $P = 0.52$, respectively). Clinicians in charge of the cases were not more likely to predict the correct 30-day postdischarge status compared to other assessors ($P = 0.20$). Assessors were significantly more accurate at predicting patient's 30-day survival in the first 6 hours following admission compared to the 12–24 hours postadmission ($P = 0.022$).

3.4 | Sharing opinions

One hundred and fifty-four assessors (57.9%) had thought about the dogs' outcomes prior to completing the surveys and 44 of them (28.6%) had shared their opinion with other people: 29 of 44 (65.9%) had shared their opinion with colleagues and 21 of 44 (47.7%) had shared it with owners (6 assessors had shared their opinion with both colleagues and owners). Among the 79 assessors who thought the dog assessed would die in hospital, 22 (27.8%) had shared their opinion, of whom 10 had shared their opinion with the owner. These 10 assessors were all the clinicians in charge of the case assessed. All 10 dogs had a poor outcome, but 3 were alive at discharge and 1 was alive 30 days postdischarge. Assessors who had shared their opinion were not more accurate at predicting the discharge outcome than those that had not ($P = 0.52$).

3.5 | Criteria used for predictions

Forty assessors (15.0%) reported using only 1 criterion to base their judgement, 86 (32.3%) used 2 criteria, and 139 (52.2%) used 3 or more criteria. Assessors who used 3 or more criteria were not more likely than others to make a correct prediction on the patient's discharge outcome ($P = 0.31$).

The most commonly used criteria to make predictions were the underlying disease process (179/266; 67.3%), physical examination findings (159/266; 59.8%), and the history of the current episode (148/266; 55.6%). Criteria independent from the patient's clinical status were used in 131 of 266 (49.2%) surveys: experience was stated in 80 of 266 (30.1%) responses, impressions from the owners in 46 of 266 (17.3%) responses, and financial situation (estimate given for treatment and insurance status) in 36 of 266 (13.5%) responses. When death or euthanasia was predicted, after disease, physical examination findings, and history, frailty was stated as a criterion in 26 of 79 cases.

4 | DISCUSSION

In this study, we investigated whether the ECC veterinary team were able to predict patient outcome. Accuracy of these predictions is important as they are commonly requested by owners and likely significantly influence their decision-making. Our results show that staff and students can predict the outcome of dogs that present to the ER within 24 hours of admission with moderate accuracy regardless of whether or not they are primarily managing the patient.

Recognizing the limitations of staff opinions may even be reassuring and enhance job satisfaction for those caring for critical inpatients. When discharge, qualitative outcome, and $APPLE_{fast}$ predictions were incorrect, they were more often needlessly negative than overly optimistic. It is an important finding that consistently better outcomes were documented than expected in these cases.

Although objective scoring systems can be helpful to clinicians and nurses, they are imperfect. The specificity of $APPLE_{fast}$ score in predicting death was reported to be 85% in a previous study.[8] In our study, 71.1% of cases with incorrect $APPLE_{fast}$ predictions actually survived, meaning an overcautious prognosis was calculable.

Illness severity models such as the APPLE score were designed to provide an objective risk stratification system that is helpful to analyze patient data in veterinary studies[8]; therefore, they should not be used as outcome predictors on an individual basis. Furthermore, the APPLE score was published in 2010, and advances in veterinary medicine and treatments available mean that it could currently be overestimating mortality in our patients. That said, some clinicians may rely on such models to advise owners on continuation of care or even euthanasia and would have been inaccurate in over a quarter of cases in this study.

Staff and students were not significantly more accurate than the $APPLE_{fast}$ score in spite of being able to take more parameters into account. Veterinarians were more accurate than $APPLE_{fast}$ score and nurses at predicting the discharge outcome of the patients evaluated in this study. It is difficult to postulate a hypothesis for this without further investigation; it could be related to veterinarians using more variables in considering outcome, confirmational or anchoring bias, or a myriad of other factors. The nurses' group was also smaller than the veterinarian group (87 surveys from nurses and 136 surveys from veterinarians), retaining a chance of type 1 error. The majority of assessors reported using at least 3 criteria to make their predictions, including criteria that were not associated with the disease process and would be difficult to integrate to an objective scoring system. Interestingly, being the clinician in charge of the case did not affect the accuracy of the prediction. It is nonetheless reassuring that clinicians in charge, with a similar frequency to other veterinary staff, make incorrect predictions rather than consciously or subconsciously fulfilling their predictions by owner influence. Nurses' predictions were less accurate and more pessimistic than reality, which could reflect a lack of objectivity when working in close contact with the patient as in the hospital where this study was carried out, nurses in the ICU are nursing an average of 5 patients each and thus spend a lot of time with 1 patient, especially with critical cases. The difference in accuracy between nurses and veterinarians could also come from the lack of contact with the owners (specifically during this study period as no owners were allowed in the hospital) or from the lack of information on the whole history and disease process of the patient, which would have been reviewed by the clinician. In a study of critically ill people, nurses were also found to be more pessimistic than doctors when assessing survival of ICU patients, especially the sicker patients, but were more accurate in their predictions as a result.[14] However, they considered treatment withdrawal more often than doctors did in dying patients, including in patients who survived and had a good quality of life 6 months after their ICU stay.[14]

Overall, predictions were more accurate in the first 6 hours following admission compared to the 12–24 hours postadmission. This was an unexpected finding as we hypothesized that case evolution and familiarity would contribute to improving accuracy of prediction. However, due to the shift patterns in the hospital where the study was performed, the assessors would unlikely be the same in the 2 different windows of assessment. The other possible reason for this contrast is the amount of information available to the assessor as staff and students might be more likely to read the previous clinical notes and take the history from the owner, having seen or performed a thorough physical examination of the patient who has recently presented to the ER compared to a patient for whom they might rely more on written documentation about the patient, their problem list, patient handover, and the patient's current physical exam findings.

Unlike reports in people,[1,3,15] experience did not affect the accuracy of the predictions and students as well as junior clinicians were able to make accurate predictions. This finding is also reflected in the criteria quoted by assessors to make their prediction as only 30% of them reported using their experience as a criterion. This might be a reflection of the population who participated in the surveys as the mean experience in ECC was 4.6 years when clinicians who participated in similar human studies had more experience and a difference of at least 10 years of experience between groups.[1]

This study was carried out during the COVID-19 pandemic, a period that increased the workload and lowered the levels of mental well-being for members of the veterinary profession,[16] which was already a profession in which well-being was lower than the general population.[17,18] Psychological distress has been shown to increase with the number of hours worked per week,[17] so we also evaluated the impact of the number of days worked on the accuracy and negativity of the predictions. Even in this more difficult period of time, neither the number of days worked in a row, mood, nor the depersonalization score of the MBI was found to be associated with the accuracy of the predictions. The mean number of days worked in a row was relatively low, at 2.5 days, and the mean global mood was good, at 7.3. The mean depersonalization score was between low and moderate, which is consistent with the previously reported score in veterinary students[19] but differs from the reported high scores in two thirds of the residents in people.[20] The assessors reporting a globally good level of well-being likely explain the absence of relation found between the parameters evaluated and the accuracy or negativity of the predictions. However, staff and students were in general more pessimistic in their predictions and would have likely advised toward withdrawal of care or even euthanasia in cases that had a good outcome. These findings are similar to those reported in people where actual ICU survival rates were higher than physician-predicted survival rates[2] and where nurses were reported to be more pessimistic than doctors proposing treatment withdrawal in some very sick patients who survived.[14] In this latter study, quality of life after ICU stay was not correctly predicted by nurses or physicians, with both wrong optimistic and pessimistic appreciations and only 6% of patients indicating a bad quality of life 6 months after their admission in ICU.

This study has several limitations. Firstly, it was performed in a referral hospital with a large first opinion emergency and referral emergency caseload. This likely explains the relatively high mean $APPLE_{fast}$ score of the patients. The $APPLE_{fast}$ score was constructed and validated to be scored from the most abnormal patient data points collected over the first 24 hours from ICU admission, whereas it was calculated within the first 6 hours postadmission in this study in an ER setting. This may have affected the accuracy of the $APPLE_{fast}$ score. This study was performed in an academic setting with veterinary students, interns, residents, diplomats, and nurses working exclusively in ECC; the findings might not be generalizable to all veterinary hospitals. Finally, the fact that dogs may be euthanized for a variety of reasons makes prognostication very complicated. Euthanasia is an inherent feature of all veterinary clinical practice and research, so inclusion and consideration of those animals is clinically very relevant.

5 | CONCLUSION

Both staff and scoring systems can predict patient outcome with moderate accuracy in the first 24 hours following presentation to the ER. Veterinarians are more accurate with their predictions than $APPLE_{fast}$ scores and nurses. However, when incorrect, staff predictions are pessimistic and may potentially lead to discontinuation of care or euthanasia when better outcomes are achievable.

ORCID

Alice Le Gal DVM, MVetMed, DACVECC, DECVECC https://orcid.org/0000-0002-5370-5974

ENDNOTE

[a] Tanagra Version 1.4.50 (Lyon, France, 2003).

REFERENCES

1. Garrouste-Orgeas M, Montuclard L, Timsit JF, Misset B, Christias M, Carlet J. Triaging patients to the ICU: a pilot study of factors influencing admission decisions and patient outcomes. *Intensive Care Med*. 2003;29(5):774-781.
2. Rocker G, Cook D, Sjokvist P, et al. Clinician predictions of intensive care unit mortality. *Crit Care Med*. 2004;32(5):1149-1154.
3. Scholz N, Bäsler K, Saur P, Burchardi H, Felder S. Outcome prediction in critical care: physicians' prognoses vs. scoring systems. *Eur J Anaesthesiol*. 2004;21(8):606-611.
4. Sinuff T, Adhikari NKJ, Cook DJ, et al. Mortality predictions in the intensive care unit: comparing physicians with scoring systems. *Crit Care Med*. 2006;34(3):878-885.
5. Copeland-Fields L, Griffin T, Jenkins T, Buckley M, Wise LC. Comparison of outcome predictions made by physicians, by nurses, and by using the Mortality Prediction Model. *Am J Crit Care*. 2001;10(5):313-319.
6. Detsky ME, Harhay MO, Bayard DF, et al. Discriminative accuracy of physician and nurse predictions for survival and functional outcomes 6 months after an ICU admission. *JAMA*. 2017;317(21):2187-2195.
7. Ripanti D, Dino G, Piovano G, Farca A. Application of the Sequential Organ Failure Assessment Score to predict outcome in critically ill dogs: preliminary results. *Schweiz Arch Tierheilkd*. 2012;154(8):325-330.
8. Hayes G, Mathews K, Doig G, et al. The Acute Patient Physiologic and Laboratory Evaluation (APPLE) score: a severity of illness stratification system for hospitalized dogs. *J Vet Intern Med*. 2010;24(5):1034-1047.
9. King LG, Wohl JS, Manning AM, Hackner SG, Raffe MR, Maislin G. Evaluation of the survival prediction index as a model of risk stratification for clinical research in dogs admitted to intensive care units at four locations. *Am J Vet Res*. 2001;62(6):948-954.
10. Maslach C, Jackson SE. The measurement of experienced burnout. *J Organ Behav*. 1981;2(2):99-113.
11. Maslach C, Jackson S, Leiter M, The Maslach Burnout Inventory manual. In: Zalaquett CP, Wood RJ, eds. *Evaluating Stress: A Book of Resources*. 3rd ed. The Scarecrow Press; 1997:191-218.
12. Obregon M, Luo J, Shelton J, Blevins T, MacDowell M. Assessment of burnout in medical students using the Maslach Burnout Inventory-Student Survey: a cross-sectional data analysis. *BMC Med Educ*. 2020;20(1):376.
13. Shi Y, Gugiu PC, Crowe RP, Way DP. A Rasch analysis validation of the Maslach Burnout Inventory-Student Survey with preclinical medical students. *Teach Learn Med*. 2019;31(2):154-169.
14. Frick S, Uehlinger DE, Zuercher Zenklusen RM. Medical futility: predicting outcome of intensive care unit patients by nurses and doctors—a prospective comparative study. *Crit Care Med*. 2003;31(2):456-461.
15. Christensen C, Cottrell JJ, Murakami J, Mackesy ME, Fetzer AS, Elstein AS. Forecasting survival in the medical intensive care unit: a comparison of clinical prognoses with formal estimates. *Methods Inf Med*. 1993;32(4):302-308.
16. Mair TS, Mountford DR, Radley R, Lockett E, Parkin TD. Mental wellbeing of equine veterinary surgeons, veterinary nurses and vet-

erinary students during the COVID-19 pandemic. *Equine Vet Educ*. 2021;33(1):15-23.
17. Volk JO, Schimmack U, Strand EB, Lord LK, Siren CW. Executive summary of the Merck Animal Health Veterinary Wellbeing Study. *J Am Vet Med Assoc*. 2018;252(10):1231-1238.
18. Hatch P, Winefield H, Christie B, Lievaart J. Workplace stress, mental health, and burnout of veterinarians in Australia. *Aust Vet J*. 2011;89(11):460-468.
19. Chigerwe M, Boudreaux KA, Ilkiw JE. Assessment of burnout in veterinary medical students using the Maslach Burnout Inventory-Educational Survey: a survey during two semesters. *BMC Med Educ*. 2014;14(1):255.
20. Mendelsohn D, Despot I, Gooderham PA, Singhal A, Redekop GJ, Toyota BD. Impact of work hours and sleep on well-being and burnout for physicians-in-training: the Resident Activity Tracker Evaluation Study. *Med Educ*. 2019;53(3):306-315.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Le Gal A, Barfield DM, Wignall RH, Cook SD. Outcome prediction in dogs admitted through the emergency room: Accuracy of staff prediction and comparison with an illness severity stratification system for hospitalized dogs. *J Vet Emerg Crit Care*. 2023;1-7.
https://doi.org/10.1111/vec.13350