Journal of Veterinary Internal Medicine

Journal of Veterinary Internal Medicine Open Access
ACVIM American College of Veterinary Internal Medicine

STANDARD ARTICLE OPEN ACCESS
Small Animal Internal Medicine Gastroenterology

Multi-Institutional Retrospective Analysis of Prognostic Scoring Systems for Dogs With Acute Pancreatitis (504 Dogs)

Harry Cridge[1] | Daniela Cabrera[1] | Timothy Bolton[2] | Cameron Vaughn[2] | Nolie Parnell[2] | Betty Chow[3] | Alexander Schreiber[3,4] | Andrew Mackin[5] | Paige Albers[1] | Sichao Wang[6]

[1]Department of Small Animal Clinical Sciences, College of Veterinary Medicine, Michigan State University, East Lansing, Michigan, USA | [2]Department Veterinary Clinical Sciences, College of Veterinary Medicine, Purdue University, West Lafayette, Indiana, USA | [3]VCA Animal Specialty and Emergency Center, Los Angeles, California, USA | [4]Department of Clinical Sciences, College of Veterinary Medicine and Biomedical Sciences, Colorado State University, Fort Collins, Colorado, USA | [5]Department of Clinical Sciences, College of Veterinary Medicine, Mississippi State University, Mississippi State, Mississippi, USA | [6]Center for Statistical Training and Consulting (CSTAT), Michigan State University, East Lansing, Michigan, USA

Correspondence: Harry Cridge (cridgeh1@msu.edu)

Received: 19 December 2024 | **Revised:** 16 April 2025 | **Accepted:** 17 April 2025

Funding: The authors received no specific funding for this work.

Keywords: canine | death | gastroenterology | pancreas | prognosis | prolonged hospitalization

ABSTRACT

Background: Acute pancreatitis (AP) in dogs has a broad clinical presentation and variable progression, making prognostication challenging.

Hypothesis/Objectives: (i) To compare predicted prognosis for death and for prolonged (≥ 5 days) hospitalization across scoring schemes for AP in dogs and (ii) to predict concordance of each scoring scheme with death and for prolonged hospitalization.

Animals: Five hundred four client-owned dogs.

Methods: Multi-institutional retrospective study. Data extracted from medical records included: signalment, history, physical examination findings, diagnostic results, length of hospitalization, and death. Five prognostic schemes (OS, CSI, $APPLE_{full}$, CAPS, MCAI) were calculated for each dog.

Results: Overall concordance was low. Only $APPLE_{full}$ ($p = 0.004$) and MCAI ($p = 0.01$) scores were significantly different between survivors and non-survivors. Overall, $APPLE_{full}$ had the greatest concordance (0.632, 95% CI: 0.592–0.672) with length of hospitalization. Of the other more pancreatitis-specific schemes, MCAI had the greatest concordance (0.576, 95% CI: 0.567–0.635) with length of hospitalization, while CSI had the lowest concordance with length of hospitalization (0.525, 95% CI: 0.494–0.556).

Conclusions and Clinical Importance: On a population level, $APPLE_{full}$ and MCAI had the greatest predictive discrimination between dogs of normal and prolonged hospitalization. If an individual dog has any of the 5 prognostic score schemes above the proposed cut-off for death, it should be interpreted with caution because of the low case fatality rate.

Abbreviations: AP, acute pancreatitis; APPLE, acute patient physiologic and laboratory evaluation; AUC, area under the curve; CAPS, canine acute pancreatitis severity; CI, 95% confidence interval; CSI, clinical severity index; iCa, ionized calcium; IQR, interquartile range; MCAI, modified clinical activity index; NPV, negative predictive value; OS, organ score; PPV, positive predictive value; ROC, receiver operating characteristic.

1 | Introduction

Acute pancreatitis (AP) is the most common disorder of the exocrine pancreas in dogs. While abdominal pain and vomiting are considered the predominant clinical signs of the disease, it is well documented that AP has a broad clinical presentation [1, 2]. This broad clinical presentation combined with a variable disease progression can make it challenging to predict the duration of hospitalization and outcome in individual dogs. Efforts have been made to derive prognostic score systems for dogs with AP, to help stratify disease severity and potentially provide prognostic information.

The 'Organ Score' (OS) scheme was published in 1998 and comprised assessment of a dog's leukogram, blood-gas, glucose, β-OH butyrate, hepatic enzymology, and renal values, all as an assessment of multi-organ dysfunction secondary to AP [3]. An OS ≥ 4 was proposed to be associated with a high risk of death, although results might have been impacted by minimal treatment attempts in several of the enrolled dogs [3]. The OS scheme was the first to introduce the concept of concentrating on organ compromise secondary to AP as a prognostic indicator, rather than focusing on pathologic processes occurring within the pancreas. The study cohort came from general practice submissions to a pathology service (Brisbane, Queensland, Australia; $n = 68$) [3]. Diagnosis of pancreatitis was based on elevated serum amylase or lipase activity, which is no longer considered optimally sensitive or specific for a diagnosis of pancreatitis.

The 'Clinical Severity Index' (CSI) scheme was subsequently developed in 2008 and was derived from cases presenting to an emergency center and a veterinary teaching hospital (Perth, Western Australia, Australia; $n = 61$) [4]. Similar to the OS scheme the CSI was a cumulative score based on evaluation of organ compromise. While CSI was correlated with outcome, further analysis of the CSI in additional veterinary hospitals could allow for improvements in outcome prediction and development of a more robust prognostic system [4]. Indeed, in a recent study ($n = 13$) the CSI was not shown to correlate with outcome, suggesting challenges with broad application of the this prognostic scheme [5].

The Acute Patient Physiologic and Laboratory Evaluation ($APPLE_{full}$) score system was developed in 2010, in a cohort of dogs presenting to an academic teaching hospital (Guelph, Ontario, Canada; $n = 810$) [6]. While this score system was not applied to just dogs with pancreatitis, 13.2% (107/810) dogs had clinical signs related to the gastrointestinal tract or pancreas [6]. An $APPLE_{full}$ score > 30 was proposed to have an 81.2% sensitivity and 89.4% specificity for prediction of death in patients presenting to an intensive care unit (ICU) [6]. Given the results of this study, some clinicians have utilized $APPLE_{full}$ when prognosticating in dogs with AP.

The Canine Acute Pancreatitis Severity (CAPS) scheme was developed in 2019 from a cohort of dogs with AP presenting to a university teaching hospital (Alfort, France; $n = 138$) and was subsequently validated in a cohort of dogs from another university teaching hospital (Toulouse, France; $n = 31$) [7]. While this study had the advantage of utilizing a second separate cohort to validate the scoring system and had promising initial results (a CAPS ≥ 11, had a sensitivity of 86% and a specificity of 92%, with a 95% CI area under the curve (AUC) of 0.77–1.00, for prediction of short term death), several studies with smaller numbers of dogs have not yielded similar performance data [7–9].

The Modified Canine Activity Index is a newer scoring system, first published in 2021 ($n = 13$), that was adapted from the Canine Inflammatory Bowel Disease Activity Index (CIBDAI) for use in dogs with AP [5]. This system has also been referred to as the Modified Clinical Activity Index (MCAI) to emphasize that it is a clinical signs-based system, in contrast to the OS, CSI, and CAPS, which are composite evaluations of systemic organ dysfunction [10]. Modified Clinical Activity Index was correlated with outcome in dogs with AP in the 2021 study, which contrasts with the published CSI [5]. To the authors knowledge, MCAI has not been directly evaluated in a larger cohort of dogs in the peer-reviewed literature. Modified Clinical Activity Index was subsequently utilized in a study of 61 dogs evaluating the efficacy of a novel drug for the treatment of AP (fuzapladib sodium) [11]. The MCAI was utilized as the primary outcome variable in the placebo-controlled clinical trial from which fuzapladib sodium achieved conditional approval by the US Food and Drug Administration for the management of the clinical signs associated with the acute onset of pancreatitis [11, 12]. Given this MCAI's performance warrants further investigation [13].

The above scoring systems are limited by the fact they were derived and evaluated from small cohorts, or they have not been externally validated, or both. This, in conjunction with the poor predictive performances in some studies, suggests that these prognostic schemes need to be re-evaluated. Importantly, many of these scoring systems are also used as co-variates or outcome variables in clinical research and study designs. Thus, the influence of these prognostic schemes is large. The aim of our study was therefore to (i) compare predicted prognosis for death and for prolonged (≥ 5 days) hospitalization across scoring schemes in a large cohort of dogs and (ii) to predict concordance of each scoring scheme with death and for prolonged hospitalization. We hypothesized that MCAI would have the greatest concordance with outcome variables when compared with previously described pancreatitis-specific score systems.

2 | Materials and Methods

2.1 | Animals and Case Definitions

Cases were identified by searching the medical record systems at 3 academic teaching hospitals and 1 large multi-specialty private referral hospital for cases of "pancreatitis" in dogs between January 1st, 2018, and December 31st, 2023. All centers practice a similar standard of medicine. Geographic populations represented included the Midwest, Southeast, and West Coast of the United States. For inclusion in the study, review of the medical record confirmed that each dog met criteria for clinically suspected or clinically probable AP, per published guidelines [14].

2.2 | Data Collection and Disease Severity Indices

Medical records were analyzed, and the following data was extracted from each dog: age, sex, weight, history, physical examination findings, clinicopathologic test results, diagnostic imaging results (radiographs, ultrasonographic examination, or computed tomography), length of hospitalization, length of follow-up, and outcome (alive versus dead/euthanized). Cause of death was recorded where known.

For each dog, the following prognostic scoring systems were retrospectively calculated from the first available data in the medical record: OS, CSI, $APPLE_{full}$, CAPS, and MCAI [3–7]. Scores were calculated by DC for 3 institutions, while the 4th institution submitted self-calculated scores for analysis. Ionized calcium (iCa) was assumed to be normal when total calcium was normal and iCa was unavailable. Otherwise, missing data was treated as such, and the corresponding scoring system from which the data was missing was not utilized in the statistical analyses. Duration of hospitalization and death were used as surrogate markers of disease severity.

2.3 | Statistical Analysis

Study cohort characteristics were summarized using descriptive statistics. Wilcoxon rank sum tests were used to compare median scores for each of the scoring systems between dogs that died or were euthanized and those that survived. Similar

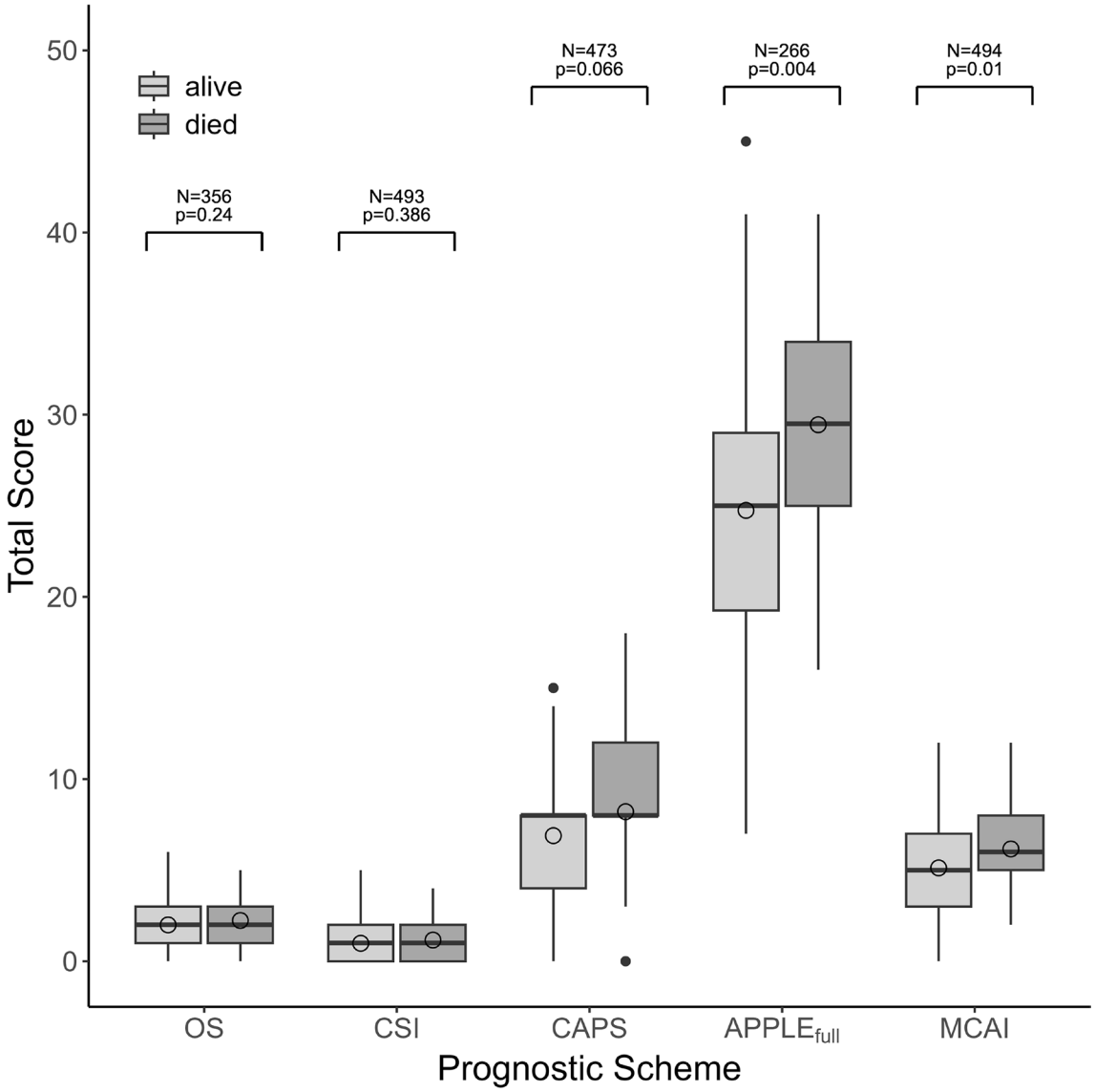

FIGURE 1 | Comparison of prognostic schemes between pancreatitis dogs that survived and dogs that died/were euthanized. $APPLE_{full}$ and MCAI were significantly different between pancreatitis dogs that died or were euthanized and those that survived. The horizontal line inside the box represents the median value. The open circle represents the mean value. The upper edge of the box represents the 3rd quartile and the lower edge of the box represents 1st quartile. The upper whisker represents the 3rd quartile +1.5 times the interquartile range or the maximum value. The lower whisker represents the 1st quartile—1.5 times the interquartile range or the minimum value. Data points outside 1.5 times the interquartile range above the 3rd quartile and below the 1st quartile are marked as an individual point.

analyses were also performed for dogs that were hospitalized for < 5 days compared with those that were hospitalized for ≥ 5 days. Each scoring system was also used to generate receiver operating characteristic (ROC) curves. Optimal cutoff values were selected based on Youden index. The ROCit package was used to derive ROC curves [15]. Concordance was calculated to evaluate the overall predictive capacity of each scoring system and is not restricted to specific cut-off values, thus allowing for a greater overall comparison between prognostic schemes. Concordance represents the area under the ROC curve. All analyses were conducted using R Statistical Software (v. 4.4.1; R Core Team; https://www.r-project.org/) [16]. For all analyses, a $p < 0.05$ was considered statistically significant.

3 | Results

3.1 | Animals

Five hundred four ($n = 504$) dogs were included in the study. The median age of enrolled dogs was 10 years (interquartile range [IQR], 7–12 years). Two hundred and fifty-six dogs were male (237 castrated and 19 intact males) and 248 were females (239 spayed and 9 intact females). The median weight of affected dogs was 9.0 kg (IQR, 5.7–20.8 kg). The median duration of hospitalization was 4 days (IQR: 2–6 days). One hundred and ninety-seven dogs (39.0%) were hospitalized for 5 or more days. The mean duration of follow up after discharge was 111 ± 305 days, with a predominant right skew. The median duration of follow up after discharge

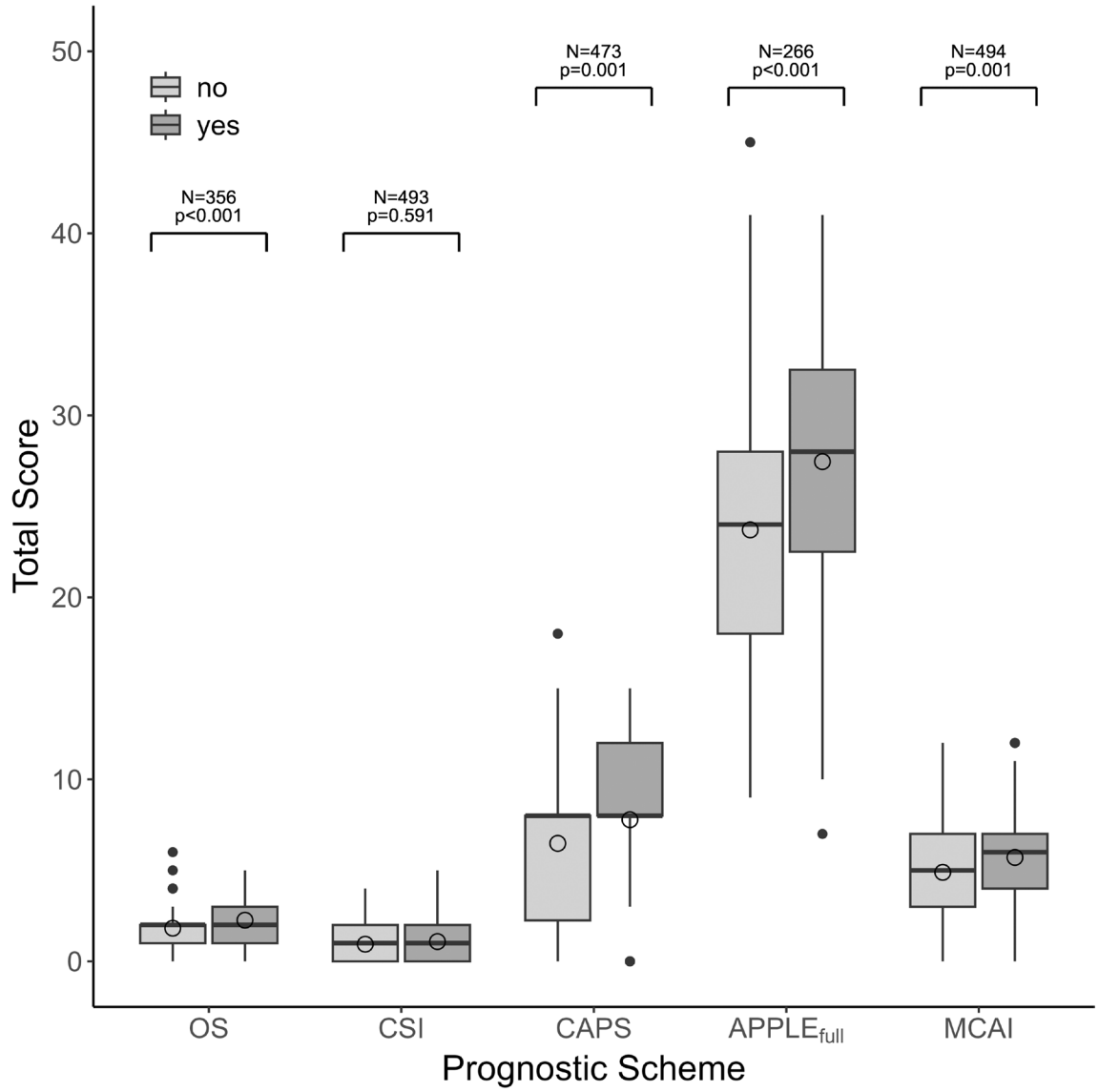

FIGURE 2 | Comparison of prognostic schemes between pancreatitis dogs that were Hospitalized for < 5 days vs. ≥ 5 days. OS, CAPS, $APPLE_{full}$, and MCAI scores were significantly different between pancreatitis dogs that were hospitalized for < 5 days vs. ≥ 5 days. CSI scores were not significantly different. The horizontal line inside the box represents the median value. The open circle represents the mean value. The upper edge of the box represents the 3rd quartile and the lower edge of the box represents 1st quartile. The upper whisker represents the 3rd quartile +1.5 times the interquartile range or the maximum value. The lower whisker represents the 1st quartile—1.5 times the interquartile range or the minimum value. Data points outside 1.5 times the interquartile range above the 3rd quartile and below the 1st quartile are marked as an individual point.

was 0 days (IQR: 0–21 days). Forty-four dogs (8.8%) died or were euthanized, and the remaining dogs survived their episode of pancreatitis. Missing data resulted in the following number of dogs utilized in each scoring system analysis: OS ($n=356$), CSI ($n=493$), $APPLE_{full}$ ($n=266$), CAPS ($n=473$), and MCAI ($n=494$).

3.2 | Wilcoxon Rank Sum Testing

3.2.1 | Prediction of Death

The OS ($p=0.24$), CSI ($p=0.39$), and CAPS ($p=0.07$) schemes were not significantly different between dogs that died and those that survived. The $APPLE_{full}$ ($p=0.004$) and MCAI ($p=0.01$) schemes were significantly different between dogs that died/were euthanized and those that survived. See Figure 1.

3.2.2 | Prediction of Prolonged Hospitalization (≥ 5 Days)

Four of five score systems were significantly different between dogs that had prolonged (≥ 5 days) hospitalization and those that had hospitalization of < 5 days duration (OS, $p<0.001$; CAPS, $p=0.001$; $APPLE_{full}$, $p<0.001$; MCAI, $p=0.001$). The CSI scheme was not significantly different ($p=0.59$) between dogs that had prolonged (≥ 5 days) hospitalization and those that had hospitalization of < 5 days duration. See Figure 2.

3.3 | Receiver Operator Curves

3.3.1 | Prediction of Death

With the OS scheme, ROC analysis revealed an optimal cut-off of 4 to predict death in dogs in the present study (Figure 3). This yielded an area under the curve (AUC) of 0.556 (95% CI: 0.457–0.655), a positive predictive value (PPV) of 23.3% (95% CI: 9.9%–42.3%), and a negative predictive value (NPV) of 90.5% (95% CI: 86.8%–93.4%). Figure 4 demonstrates PPV and NPV data across a range of death prevalences for each prognostic scheme. For the cut-off value of 4, sensitivity was 18.4% (95% CI: 7.7%–34.3%), and specificity was 92.8% (95% CI: 89.3%–95.4%).

With the CSI scheme, ROC analysis revealed an optimal cut-off of 2 to predict death. This yielded an AUC of 0.537 (95% CI: 0.446–0.629), a PPV of 11.1% (95% CI: 6.5%–17.4%), and a NPV of 92.3% (95% CI: 88.9%–94.8%). For the cut-off value of 2, sensitivity was 37.2% (95% CI: 23.0%–53.3%), and specificity was 71.6% (95% CI: 67.1%–75.7%).

With the CAPS scheme, ROC analysis revealed an optimal cut-off of 12 to predict death. This yielded an AUC of 0.579 (95% CI: 0.486–0.672), a PPV of 15.4% (95% CI: 8.7%–24.5%), and a NPV of 92.4% (95% CI: 89.3%–94.9%). For the cut-off value of 12, sensitivity was 32.6% (95% CI: 19.1%–48.5), and specificity was 82.1% (95% CI: 78.1%–85.6%).

With the $APPLE_{full}$ scheme, ROC analysis revealed an optimal cut-off of 25 to predict death. This yielded an AUC of 0.691 (95% CI: 0.558–0.824), a PPV of 11.7% (95% CI: 7.0%–18.1%), and a NPV of 97.5% (95% CI: 92.9%–99.5%). For the cut-off value of 25, sensitivity was 85.0% (95% CI: 62.1%–96.8%), and specificity was 48.0% (95% CI: 41.6%–54.4%).

With the MCAI scheme, ROC analysis revealed an optimal cut-off of 5 to predict death. This yielded an AUC of 0.618 (95% CI: 0.525–0.710), a PPV of 11.4% (95% CI: 8.0%–15.5%), and a NPV of 95.4% (95% CI: 91.4%–97.9%). For the cut-off value of 5, sensitivity was 79.1% (95% CI: 64.0%–90.0%), and specificity was 41.2% (95% CI: 36.7%–45.9%).

3.3.2 | Prediction of Prolonged Hospitalization (≥ 5 Days)

With the OS scheme, ROC analysis revealed an optimal cut-off of 3 to predict prolonged hospitalization (≥ 5 days; Figure 5). This yielded an AUC of 0.621 (95% CI: 0.562–0.679), a PPV of 59.6% (95% CI: 49.8%–68.9%), and a NPV of 61.1% (95% CI: 54.7%–67.2%). Figure 6 demonstrates PPV and NPV data across a range of prolonged hospitalization prevalences for each prognostic scheme. For the cut-off value of 3, sensitivity was 40.4% (95% CI: 32.7%–48.4%), and specificity was 77.4% (95% CI: 70.9%–83.1%).

With the CSI scheme, ROC analysis revealed an optimal cut-off of 2 to predict prolonged hospitalization (≥ 5 days). This yielded an AUC of 0.513 (95% CI: 0.461–0.566), a PPV of 44.4% (95% CI: 36.2%–52.9%), and a NPV of 61.9% (95% CI: 56.6%–67.0%). For the cut-off value of 2, sensitivity was 32.5% (95% CI: 26.0%–39.5%), and specificity was 73.0% (95% CI: 67.5%–77.9%).

With the CAPS scheme, ROC analysis revealed an optimal cut-off of 12 to predict prolonged hospitalization (≥ 5 days). This yielded an AUC of 0.582 (95% CI: 0.529–0.635), a PPV of 51.8%

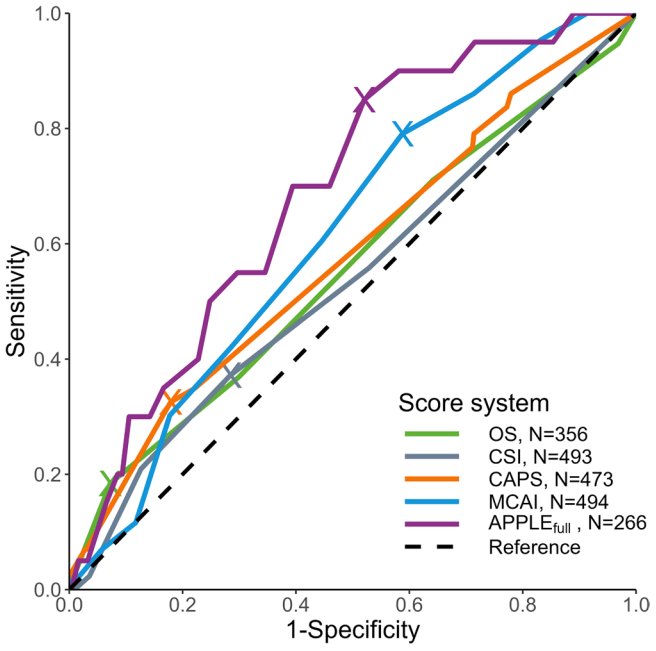

FIGURE 3 | Receiver operator characteristic curves for prediction of death in dogs with pancreatitis. A 'X' mark indicates the optimal cut-off determined by the Youden index. Score system and '*n*' are provided within the figure.

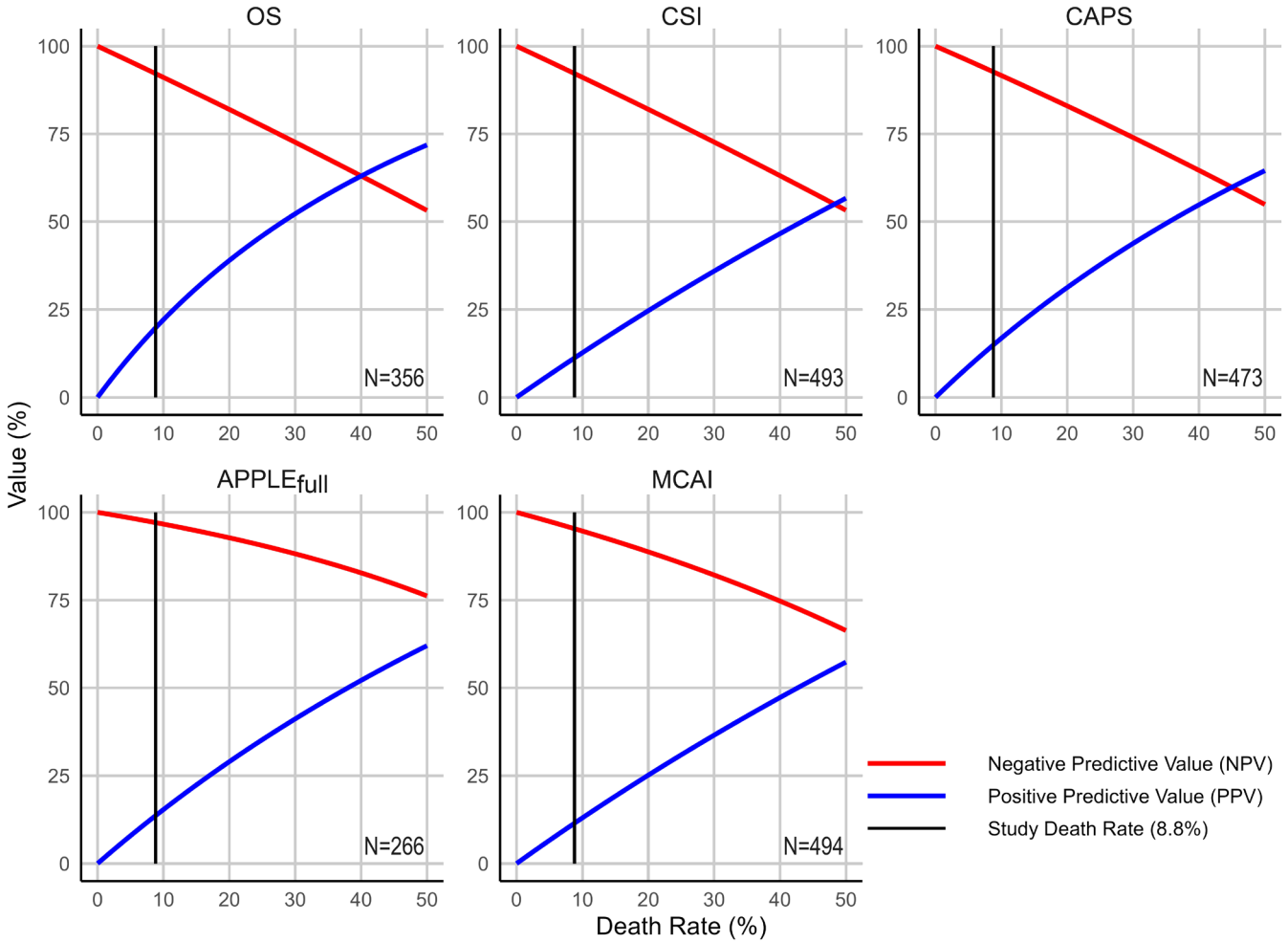

FIGURE 4 | Positive and negative predictive values across a range of prevalences for death in dogs with pancreatitis. The red line indicates the negative predictive value for death of each scoring system across a range of prevalences. The blue line indicates the positive predictive value for death of each scoring system across a range of prevalences. The black line represents the death prevalence (8.8%) in our study cohort.

(95% CI: 42.1%–61.4%), and a NPV of 62.5% (95% CI: 57.3%–67.5%). For the cut-off value of 12, sensitivity was 32.6% (95% CI: 19.1%–48.5%), and specificity was 82.1% (95% CI: 78.1%–85.6%).

With the $APPLE_{full}$ scheme, ROC analysis revealed an optimal cut-off of 29 to predict prolonged hospitalization (≥ 5 days). This yielded an AUC of 0.656 (95% CI: 0.586–0.725), a PPV of 57.1% (95% CI: 45.9%–67.9%), and a NPV of 72.0% (95% CI: 64.9%–78.4%). For the cut-off value of 29, sensitivity was 48.5% (95% CI: 38.3%–58.7%), and specificity was 78.4% (95% CI: 71.4%–84.4%).

With the MCAI scheme, ROC analysis revealed an optimal cut-off of 4 to predict prolonged hospitalization (≥ 5 days). This yielded an AUC of 0.584 (95% CI: 0.532–0.636), a PPV of 44.0% (95% CI: 38.8%–49.3%), and a NPV of 71.9% (95% CI: 63.5%–79.2%). For the cut-off value of 4, sensitivity was 80.6% (95% CI: 74.4%–85.9%), and specificity was 32.6% (95% CI: 27.3%–38.2%).

3.4 | Overall Concordance With Length of Hospitalization

Concordance can be considered as the overall predictive capacity of each scoring system, and is not restricted to specific cut-off values, allowing for greater overall comparison between prognostic schemes. It represents the area under the ROC curve. The $APPLE_{full}$ had the greatest concordance with length of hospitalization (0.632, SE: 0.020, 95% CI: 0.592–0.672), followed by MCAI (0.601, SE: 0.017, 95% CI: 0.567–0.635), OS (0.595, SE: 0.019, 95% CI: 0.558–0.633), CAPS (0.576, SE: 0.017, 95% CI: 0.543–0.608), and CSI (0.525, SE: 0.016, 95% CI: 0.494–0.556). Concordance was not calculated for prediction of death due to low numbers in each group.

4 | Discussion

Our study aimed to compare commonly utilized prognostic scoring systems in a large cohort of dogs ($n = 504$) with suspected or clinically probable pancreatitis. This large study included cases from the Midwest, Southeast, and West Coast of the United States. The complexity of each prognostic scheme varied, with some prognostic schemes utilizing data that is not routinely available in all practice settings, including the use of acid–base status (OS), electrical conduction/ECG data (CSI), cavity fluid assessment ($APPLE_{full}$), and coagulation testing (CAPS). In contrast, the MCAI utilized case data readily available to all clinicians and is likely easier to apply at the dogs side.

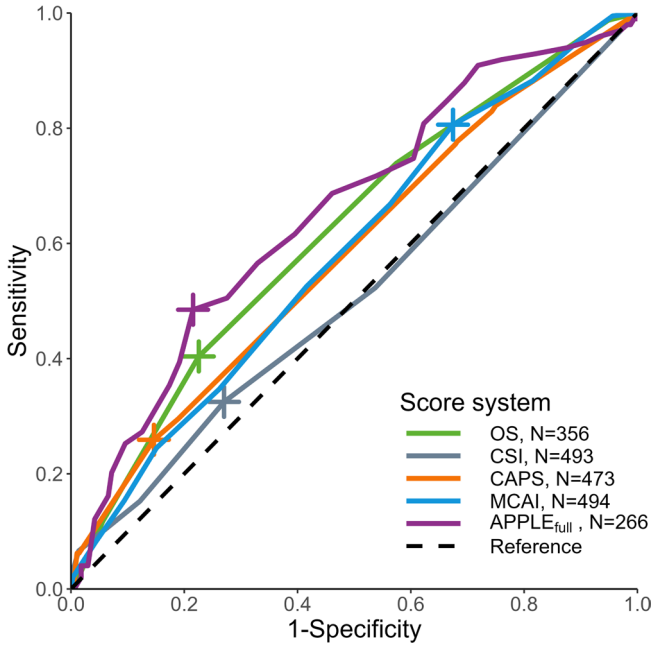

FIGURE 5 | Receiver operator characteristic curves for prediction of hospitalization (≥5 days) in dogs with pancreatitis. A 'X' mark indicates the optimal cut-off determined by the Youden index. Score system and '*n*' are provided within the figure.

Interestingly, the OS, CSI, and CAPS schemes were not significantly different between dogs that died and dogs that survived, suggesting potential limited utility of these scoring systems for use in individual animals or as part of disease severity stratification in statistical models. $APPLE_{full}$ and MCAI were significantly different between dogs that died and dogs that survived, with MCAI being the only pancreatitis-specific scheme that was significantly different between survivors and non-survivors. Our findings appear to contradict commentary made in a recent letter to the editor regarding a fuzapladib sodium study, in which CAPS was recommended to be performed in preference to MCAI [13].

When utilizing length of hospitalization as a surrogate marker of disease severity, 4 of 5 prognostic schemes were different between dogs that were hospitalized for <5 days and those that were hospitalized for ≥5 days. Five days was selected to represent prolonged hospitalization based on clinical perception, and because the mean duration of hospitalization in this study was 4.4 days. While significant differences were noted between these two groups in 80% of the scoring schemes (by duration of hospitalization), it is also important to determine the degree of predictive discrimination, in this case using concordance. Concordance data revealed that $APPLE_{full}$ had the best overall predictive discrimination between those dogs experiencing prolonged hospitalization and those that did not; however, the concordance index (0.632) was only moderately above that of 0.5, a value that represents no predictive discrimination ability above random chance. The $APPLE_{full}$ scheme had the greatest concordance, and this might be due to the fact that this score system was generated in the large cohort of dogs, which could lead to a broad clinical application. Out of the prognostic schemes that were developed to be pancreatitis-specific, MCAI had the highest predictive discrimination for prolonged hospitalization; however, the concordance index was 0.601. The MCAI scheme has been evaluated in a limited number of dogs to date and further optimization of the scoring system in a larger cohort of dogs might allow for improvement in its concordance. Of the schemes that were significantly different between dogs that were hospitalized for <5 days and those hospitalized for ≥5 days, the CAPS scheme had the lowest predictive discrimination. This, in conjunction with a lack of significant difference in CAPS between dogs that died and those that survived, indicates that this scheme performed poorly in our cohort. This mirrors recent studies in which few dogs (0/3 and 1/5) died with a CAPS ≥11, despite CAPS being proposed to represent a 92% specificity for short-term death in the original study [7–9]. Given the above findings, if a prognostic score scheme is to be utilized on a population basis, the authors recommend the use of $APPLE_{full}$ or MCAI; however, the authors also note that these schemes need to be further evaluated in large prospective studies given the overall low concordance and lack of a current optimal scheme. These studies would also allow for further optimization of schemes, or combinations of schemes, to generate higher concordance values, which would be of improved utility. The authors discourage the use of prognostic score data to make life-altering decisions (e.g., euthanasia) on an individual dog basis, particularly in environments in which a similar low death rate would be expected. The PPV and NPV data for death in this study were influenced by the low death rate seen in this cohort of dogs treated at advanced veterinary medical centers across the United States. This likely does not represent PPVs and NPVs across all practice types, and we encourage the use of Figures 4 and 6 to assist in predicting PPV and NPV in scenarios where a different prevalence for death might be seen.

In our study, we also performed receiver operator characteristic curves to further investigate the utility of these prognostic scoring schemes. As part of this study, we focused on discussing PPV and NPV, rather than solely sensitivity and specificity, to more accurately represent the interpretative value of each prognostic scheme in the context of a low prevalence of death (8.8%) and a moderate prevalence of prolonged hospitalization (39.4%). The PPVs for death for the 5 prognostic schemes were low (11.1%–23.3%), indicating only a relatively small likelihood that dogs with a score above the derived cut-off value for each scheme would go on to die as predicted by the scheme. The NPVs for death for the 5 prognostic schemes were relatively high (90.5%–97.5%), suggesting that there is a high likelihood that a dog with a score below the derived cut-off for each score system would indeed not go on to die. The PPVs for prediction of prolonged hospitalization were overall relatively poor (44.0%–59.6%), suggesting limited clinical utility. The NPVs (61.1%–72.0%) for prediction of prolonged hospitalization were mildly improved relative to the PPVs; however, they remained overall relatively poor. It is important to note that these PPV and NPV values are specific to the cut-off value derived by Youden index to maximize the AUC. Manual selection of alternate cut-off values would yield alternate results.

This study was limited by its retrospective nature, non-standardized treatments, and non-standardized case follow-up. Diagnostic criteria for pancreatitis in this study were based on previously published guidelines for clinically suspected and clinically probable pancreatitis and did not include histopathologic

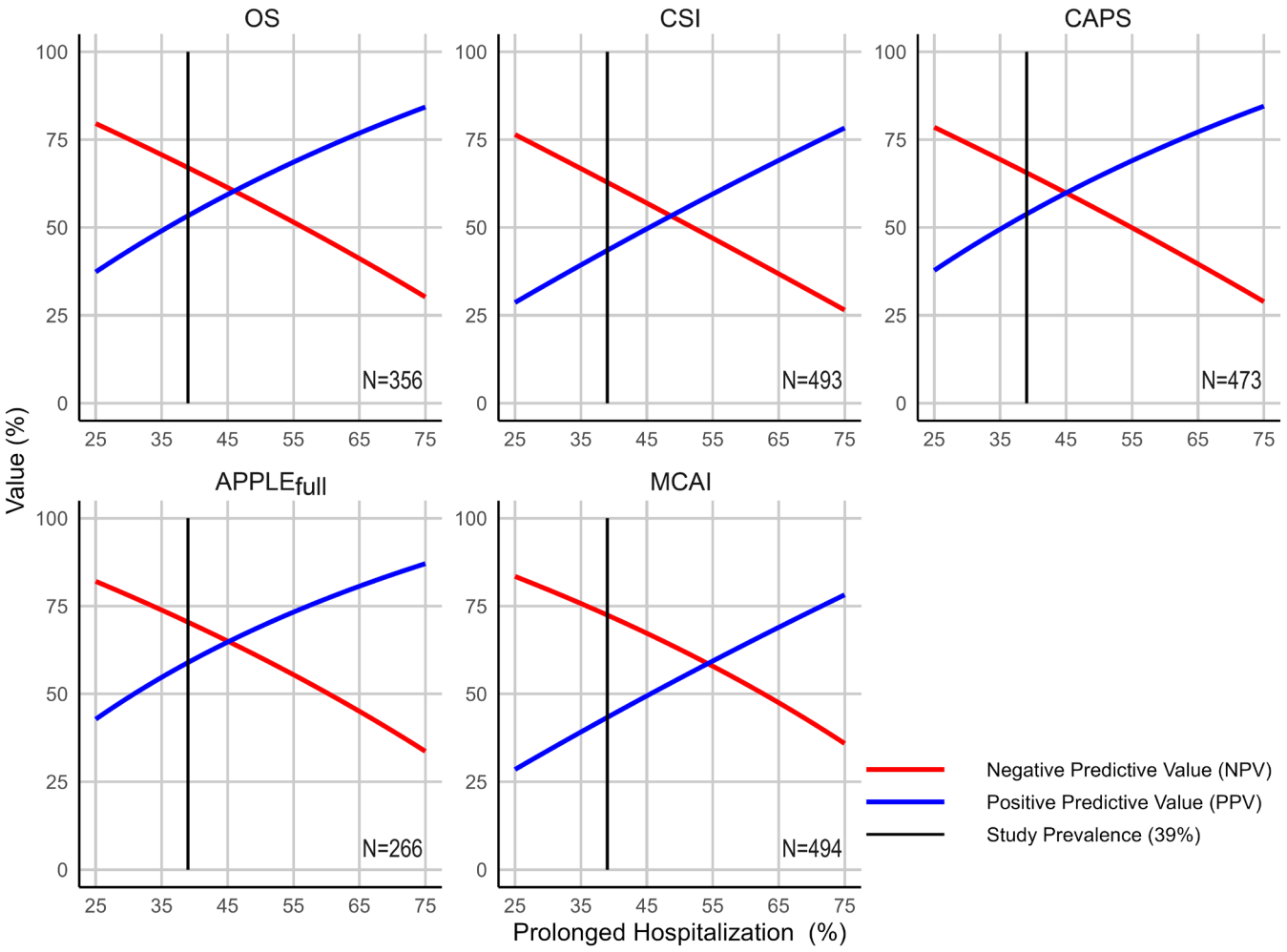

FIGURE 6 | Positive and negative predictive values across a range of prevalences for prolonged hospitalization (≥ 5 days) in dogs with pancreatitis. The red line indicates the negative predictive value for prolonged hospitalization (≥ 5 days) of each scoring system across a range of prevalences. The blue line indicates the positive predictive value for prolonged hospitalization (≥ 5 days) of each scoring system across a range of prevalences. The black line represents the prolonged hospitalization prevalence (39%) in our study cohort.

assessment of pancreatic tissue. Scoring systems were retrospectively assigned based on data available in each dog's medical records, leading to missing data that prevented the assessment of all prognostic schemes in some dogs. Future prospective data collection would allow for comparison of all scoring schemes within every dog in a study, thus allowing for a true head-to-head comparison between schemes. This might also allow for generation of new schemes utilizing artificial intelligence. Factors outside of disease severity, such as financial constraints or co-existing disease, could have influenced duration of hospitalization and death; however, we are hopeful that the large number of dogs in this study helps to mitigate the overall impact of these factors on the study results. Minimal post-discharge data was available and, as such, death data should be considered to reflect death during initial hospitalization. Pre-presentation treatment and pre-presentation duration of illness were unable to be accurately assessed in the retrospective data sets and were not included as confounding variables in this study. Future prospective studies should consider these variables in their analyses. Level of hospitalization care, medication support, and institution standardized cost of care would also be important factors to consider in future prospective analyses.

In conclusion, OS, CSI, and CAPS were not significantly different between dogs that were euthanized/died and those that survived, suggesting that these scores might not be optimal to predict death in dogs with AP. On a population level, the $APPLE_{full}$ had the highest concordance/predictive discrimination for prolonged hospitalization. A prognostic score below the cut-off values for death in this study results in a high likelihood of survival. However, a value above the proposed cut-off for death yields only a limited likelihood for non-survival, suggesting that these scores should not be used to predict poor outcomes on an individual dog basis, in environments where a similar prevalence of death can be expected.

Disclosure

Authors declare no off-label use of antimicrobials.

Ethics Statement

Authors declare no institutional animal care and use committee or other approval was needed. Authors declare human ethics approval was not needed.

Conflicts of Interest

Harry Cridge is a paid speaker for CEVA Animal Health, Royal Canin, and IDEXX. Harry Cridge has received research funding from multiple sources including industry funding from CEVA Animal Health, and Royal Canin (via a Comparative Gastroenterology Society Research Grant). These arrangements did not result in bias. The other authors declare no conflicts of interest.

References

1. H. Cridge, N. Scott, and J. M. Steiner, "Risk Factors and Clinical Presentation in Dogs With Increased Serum Pancreatic Lipase Concentrations-A Descriptive Analysis," *Animals* 12, no. 12 (2022): 1581.

2. P. G. Xenoulis, "Diagnosis of Pancreatitis in Dogs and Cats," *Journal of Small Animal Practice* 56 (2015): 13–26.

3. C. G. Ruaux and R. B. Atwell, "A Severity Score for Spontaneous Canine Acute Pancreatitis," *Australian Veterinary Journal* 76, no. 12 (1998): 804–808, https://doi.org/10.1111/j.1751-0813.1998.tb12331.x.

4. C. S. Mansfield, F. E. James, and I. D. Robertson, "Development of a Clinical Severity Index for Dogs With Acute Pancreatitis," *Journal of the American Veterinary Medical Association* 233 (2008): 936–944.

5. K. M. Keany, G. T. Fosgate, S. M. Perry, S. T. Stroup, and J. M. Steiner, "Serum Concentrations of Canine Pancreatic Lipase Immunoreactivity and C-Reactive Protein for Monitoring Disease Progression in Dogs With Acute Pancreatitis," *Journal of Veterinary Internal Medicine* 35 (2021): 2187–2195.

6. G. Hayes, K. Mathews, G. Doig, et al., "The Acute Patient Physiologic and Laboratory Evaluation (APPLE) Score: A Severity of Illness Stratification System for Hospitalized Dogs," *Journal of Veterinary Internal Medicine* 24 (2010): 1034–1047.

7. V. Fabres, O. Dossin, C. Reif, et al., "Development and Validation of a Novel Clinical Scoring System for Short-Term Prediction of Death in Dogs With Acute Pancreatitis," *Journal of Veterinary Internal Medicine* 33 (2019): 499–507.

8. N. V. Tusa, A. Abuelo, N. A. Levy, J. C. Gandy, D. K. Langlois, and H. Cridge, "Peripheral Biomarkers of Oxidative Stress in Dogs With Acute Pancreatitis," *Journal of Veterinary Internal Medicine* 36 (2022): 1958–1965.

9. H. Cridge, D. K. Langlois, J. M. Steiner, and R. A. Sanders, "Cardiovascular Abnormalities in Dogs With Acute Pancreatitis," *Journal of Veterinary Internal Medicine* 37 (2023): 28–36.

10. L. Mitchell, S. Wang, J. Lawver, and H. Cridge, "Serial Monitoring of Pancreatic Lipase Immunoreactivity, C-Reactive Protein, Abdominal Ultrasonography, and Clinical Severity in Dogs With Suspected Pancreatitis," *Journal of Veterinary Internal Medicine* 38 (2024): 987–994.

11. J. M. Steiner, C. Lainesse, Y. Noshiro, et al., "Fuzapladib in a Randomized Controlled Multicenter Masked Study in Dogs With Presumptive Acute Onset Pancreatitis," *Journal of Veterinary Internal Medicine* 37, no. 6 (2023): 2084–2092.

12. FDA, "Freedom of Information Summary—Application for Conditional Approval—Panoquell CA-1 (Application Number 141–567).", https://animaldrugsatfda.fda.gov/adafda/app/search/public/document/downloadFoi/13134.

13. K. Allenspach, A. Handel, S. Marks, et al., "Letter Regarding "Fuzapladib in a Randomized Controlled Multicenter Masked Study in Dogs With Presumptive Acute Onset Pancreatitis"," *Journal of Veterinary Internal Medicine* 38 (2024): 1282–1284.

14. S. Y. Lim, J. M. Steiner, and H. Cridge, "Understanding Lipase Assays in the Diagnosis of Pancreatitis in Veterinary Medicine," *Journal of the American Veterinary Medical Association* 260 (2022): 1249–1258.

15. M. Khan and T. Brandenburger, "ROCit: Performance Assessment of Binary Classifier With Visualization," 2024 R Package Version 2.1.2, https://CRAN.R-project.org/package=ROCit.

16. R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2024).