Development and assessment of a deep neural network to predict functional outcome and perioperative adverse events after decompressive thoracolumbar hemilaminectomy for canine acute intervertebral disc extrusion: a retrospective cohort study of 460 cases

Daniel Low, Scott Rutherford

frank. Pet Surgeons., Leeds

Acute intervertebral disc extrusion (IVDE) is a common condition in chondrodystrophic dogs, which may be surgically treated with decompressive thoracolumbar hemilaminectomy (DTLH). Prognostic factors have been previously investigated with traditional statistical analysis. Understanding prognosis and risk is important for patient selection, client counselling, postoperative management and surveillance, and clinical trial design. In human neurosurgery, machine learning algorithms have been investigated as prognostic tools. In DTLH for IVDE, a binary classification deep neural network may be able to predict a negative event.

Retrospective institutional records from 2018-2023 were searched for “IVDE” and “DTLH”. Patients without follow-up data or acute IVDE were excluded. Follow-up was obtained via neurological exam four weeks postoperatively. Input data included age, breed, sex, bodyweight, duration of clinical signs, admission Frankel score, imaging modality, time to surgery, day of the week, surgeon, and use of fenestration. Functional outcome was defined as recovery of ambulation and perioperative adverse events included gastrointestinal signs, lower urinary tract signs, and complications secondary to recumbency; these were defined binarily as the output. The dataset was randomly split into training and testing sets in an 80:20 ratio for a three-layer feedforward deep neural network developed using Keras on the Google Colaboratory Python 3.0 environment. Logistic regression analysis of the same dataset served as the control. Model performances were compared with McNemar's test.

The search yielded 556 cases, 96 were excluded, leaving 460 in the sample population. The negative event rate was 127/460 (27.6%). The model identified 89/102 cases with a negative event in the training set with a peak accuracy of 94.8%, a sensitivity of 87.3% (95% CI: 80.0-93.5%), a specificity of 97.7% (95% CI: 95.7-99.3%), a positive predictive value of 93.7% (95% CI: 88.5-97.8%), a negative predictive value of 95.2% (95% CI: 92.4-97.7%), and an AUC of 0.925 (95% CI: 0.887-0.958). Logistic regression analysis identified 36/102 dogs with a negative event ($p<0.001$) in the training set. The model identified 5/25 dogs with a negative event in the test set with a peak accuracy of 68.5%, a sensitivity of 20.0% (95% CI: 0.0-30.4%), a specificity of 86.6% (95% CI:75.8-92.4%), a positive predictive value of 35.7% (95% CI: 0.0-50.0%), a negative predictive value of 74.4% (95% CI: 63.3-82.5%), and an AUC of 0.533 (95% CI: 0.420-0.587). Logistic regression analysis identified 4/25 dogs with a negative event ($p=0.02$) in the test set.

This study demonstrates the application of a deep neural network in veterinary neurosurgery. The model identified complex associations between clinical variables and outperformed traditional regression methods. The model was less accurate in the test set but this may be due to the limited sample size, data quality, and model architecture. A deep neural network has the potential to predict negative events after veterinary neurosurgery. Further prospective investigation of machine learning algorithms, and with larger or external datasets are warranted.