UC Davis
UC Davis Previously Published Works

Title

Evaluation of a machine learning tool to screen for hypoadrenocorticism in dogs presenting to a teaching hospital

Permalink

Journal

ISSN

Authors

Reagan, Krystle L
Pires, Jully
Quach, Nina
et al.

Publication Date

DOI

Peer reviewed

STANDARD ARTICLE

Evaluation of a machine learning tool to screen for hypoadrenocorticism in dogs presenting to a teaching hospital

Krystle L. Reagan[1] | Jully Pires[2] | Nina Quach[2] | Chen Gilor[3]

[1]Department of Medicine and Epidemiology, School of Veterinary Medicine, University of California-Davis, Davis, California, USA

[2]Veterinary Medical Teaching Hospital, School of Veterinary Medicine, University of California-Davis, Davis, California, USA

[3]Department of Small Animal Clinical sciences, College of Veterinary Medicine, University of Florida, Gainesville, Florida, USA

Correspondence
Krystle L. Reagan, Department of Medicine and Epidemiology, School of Veterinary Medicine, University of California-Davis, 1 Garrod Drive, Davis, CA 95616, USA.
Email: kreagan@ucdavis.edu

Abstract

Background: Dogs with hypoadrenocorticism (HA) have clinical signs and clinicopathologic abnormalities that can be mistaken as other diseases. In dogs with a differential diagnosis of HA, a machine learning model (MLM) has been validated to discriminate between HA and other diseases. This MLM has not been evaluated as a screening tool for a broader group of dogs.

Hypothesis: An MLM can accurately screen dogs for HA.

Animals: Dogs (n = 1025) examined at a veterinary hospital.

Methods: Dogs that presented to a tertiary referral hospital that had a CBC and serum chemistry panel were enrolled. A trained MLM was applied to clinicopathologic data and in dogs that were MLM positive for HA, diagnosis was confirmed by measurement of serum cortisol.

Results: Twelve dogs were MLM positive for HA and had further cortisol testing. Five had HA confirmed (true positive), 4 of which were treated for mineralocorticoid and glucocorticoid deficiency, and 1 was treated for glucocorticoid deficiency alone. Three MLM positive dogs had baseline cortisol ≤2 μg/dL but were euthanized or administered glucocorticoid treatment without confirming the diagnosis with an ACTH-stimulation test (classified as "undetermined"), and in 4, HA was ruled out (false positives). The positive likelihood ratio of the MLM was 145 to 254. All dogs diagnosed with HA by attending clinicians tested positive by the MLM.

Conclusions and Clinical Importance: This MLM can robustly predict HA status when indiscriminately screening all dogs with blood work. In this group of dogs with a low prevalence of HA, the false positive rates were clinically acceptable.

KEYWORDS
ACTH, Addison's disease, artificial intelligence, cortisol

1 | INTRODUCTION

Hypoadrenocorticism (HA) in dogs is caused most commonly by immune mediated destruction of the adrenal cortex with subsequent

deficiencies in hormone production. Most cases of CHA lack both glucocorticoids and mineralocorticoids (GMDH); in 30% to 40% of dogs, glucocorticoid deficiency (GDH) is diagnosed without overt mineralocorticoid deficiency.[1,2] Clinical presentations of HA can vary widely, and signs can be vague and episodic.[3] The manifestation of disease can mimic primary gastrointestinal disease, renal disease, hepatic insufficiency, and many other diseases processes, and dogs with mild signs, especially those with GDH, might have HA overlooked as a potential differential diagnosis.[3-5]

Characteristic clinicopathologic findings that are associated with HA include hyponatremia, hyperkalemia, and azotemia that are associated with aldosterone deficiency, and hypoalbuminemia, hypocholesterolemia, hypoglycemia, and lack of a stress leukogram that are associated with cortisol deficiency.[2,6,7] Observing all, or some, of these abnormalities in a dog with compatible clinical signs should prompt the clinician to test for HA. However, in many cases, these variables can be within the normal reference range, and these variables alone are neither sensitive nor specific for the screening of HA.[3,6,7] This is especially true for cases of GDH that lack the characteristic electrolyte changes associated with aldosterone deficiency.[8]

Clinical decision making has been augmented with statistical and mathematical tools that can screen for diseases or make diagnostic predictions. A machine learning algorithm was utilized to train a model that predicts a diagnosis of HA, regardless of GMDH or GDH subcategory.[1] The trained model has a sensitivity of 96.3% (95% CI, 81.7%-99.8%) and specificity of 97.2% (95% CI, 93.7%-98.8%) and outperforms other screening tools.[1] The data utilized to train and validate this model was extracted in a retrospective study from the medical records of dogs that had a baseline cortisol or ACTH stimulation test performed; therefore, the attending clinician seemingly had an index of suspicion for HA and the pretest probability was higher than it would be in a general population presenting to a veterinary hospital. In this study, our objective was to assess the performance of the machine learning model (MLM) as a screening tool, when applied sequentially and indiscriminately to all dogs that had a CBC and serum chemistry ordered by the attending clinician, regardless of clinical signs or suspicion for HA.

2 | METHODS

All dogs presented to the William R. Prichard Veterinary Medical Teaching Hospital (VMTH) at the University of California, Davis School of Veterinary Medicine between March 26, 2019 and June 17, 2019 were considered for enrollment in this study. Dogs were included if they had a CBC (Advia 120; Siemens, Erlangen, Germany) and serum chemistry panel (Cobas c501/6000 series; Roche Diagnostics, Indianapolis, Indiana) performed on the same day in the VMTH Clinical Diagnostic Laboratory, regardless of presenting complaint or history. Dogs were presented for primary care, secondary, or tertiary referrals through the community practice service, emergency service, or specialty services. If a dog had multiple paired CBC and chemistry panels during the study period, only the first pair was included in the study. An MLM classification of HA status was made for each paired

CBC and chemistry panel using a trained and validated MLM as described.[1] In brief, data were collected retrospectively from 908 control dogs (suspected to have HA, but disease ruled out) and 133 dogs with confirmed HA. A boosted tree algorithm (AdaBoost) was trained with 80% of the collected data, and the remaining 20% of the data was then utilized to assess MLM performance.

The electronic medical record was evaluated for each enrolled dog. The age, breed, and sex were recorded for each dog. The dog's primary disease process was categorized into one of 17 categories based on the primary clinical diagnosis assigned by the attending veterinarian and upon review by an author (KR). These categories include cardiac, dental, dermatologic, endocrine disease (other than HA), gastrointestinal/hepatic/pancreatic (nonneoplastic), apparently healthy, HA (newly diagnosed), HA (medically managed), infectious/immune mediated disease, neoplasia, neurologic, ocular, orthopedic, respiratory, trauma/toxin exposure, unknown diagnosis, or urogenital disease.

Serum cortisol concentrations were measured if the attending veterinarian had a clinical suspicion of HA, without the knowledge of the MLM results at the point of care. Dogs with a baseline cortisol (Immulite 2000; Siemens, Erlangen, Germany) concentration <2 μg/dL, had an ACTH stimulation test performed by administration of ≥5 mcg/kg of cosyntropin (Cortrosyn, Amphastar Pharmaceuticals, Inc) IV. A 1-hour post stimulation cortisol concentration <2 μg/dL was considered confirmatory of HA in dogs with no history of being administered medications that altered the pituitary-adrenal gland axis. Dogs with a baseline or post stimulation serum cortisol concentration of >2 μg/dL were categorized as disease negative for HA. If a dog was euthanized, treated with corticosteroids, or lost to follow up before either ruling out HA with a baseline cortisol or confirming HA with an ACTH stimulation test, their HA status was categorized as undetermined.

Demographic data are reported as median (range). Prevalence of HA was calculated. The positive likelihood ratio of the MLM as compared to gold-standard cortisol measurement was calculated (Prism 7; GraphPad, La Jolla, California). Chi-squared analysis was performed to compare proportions of sex in each classification group and a Kruskal-Wallis test was performed to compare weight and age of dogs in each classification group. The positive and negative percent agreement of the MLM were calculated as compared to the clinical diagnosis assigned by the attending clinician as HA or non-HA related diagnosis. The positive percent agreement was calculated as true positives/(true positives + false negatives). The negative percent agreement was calculated as true negatives/(true negatives + false positives).

3 | RESULTS

3.1 | Dogs

During the study period, 8116 dogs were evaluated, and 1151 paired CBC and serum chemistry panels were obtained from 1025 dogs. Five dogs were newly diagnosed with HA, resulting in a prevalence of 0.06%. Dogs with newly diagnosed HA had a median age of 7 years (1-9 years) and a median body weight of 8.5 kg (5.9-14 kg) (Table 1).

TABLE 1 Summary of the dogs based on MLM and HA categorization

MLM prediction and clinical categorization	Number	Age (years)	Sex	Body weight (kg)
MLM positive				
Confirmed HA (newly diagnosed with stimulation test)	5	7 (8)	FS—2 MC—3	8.5 (8.4)
Suspected HA (baseline cortisol ≤2 μg/dL but euthanized or administered steroids before ACTH stimulation)	3	9 (6)	FS—2 MC—1	31 (3.2)
HA ruled out with cortisol >2 μg/dL	4	6 (7)	FS—2 MC—2	24 (41)
No cortisol testing (euthanized before testing)	2	5 (8)	FS—1 MC—1	23 (30)
MLM negative				
HA clinically negative	1007	9 (18)	F—28 FS—446 M—92 MC—440 U—1	19 (107)
Baseline cortisol <2 μg/dL without follow up	4	5 (7)	F—1 FS—2 MC—1	26 (34)

Note: Age and body weight are shown as median (range). Sex is abbreviated as male castrated (MC), male (M), female spayed (FS), female (F), or unknown (U). No significant difference in age ($P = .1$), body weight ($P = .4$), or sex distribution ($P = .9$) across groups.

TABLE 2 Primary clinical diagnosis of all dogs included in study

| Clinical diagnosis	No. (%)
Cardiac	24 (2.3%)
Dental	52 (5.1%)
Dermatologic	36 (3.5%)
Endocrine	29 (2.8%)
GI/liver/pancreas	175 (17%)
Healthy	51 (5.0%)
Hypoadrenocorticism—newly diagnosed	5 (0.5%)
Hypoadrenocorticism—medically managed	3 (0.3%)
Infectious/immune	68 (19%)
Neoplasia	192 (13%)
Neurologic	136 (13%)
Ocular	40 (3.9%)
Orthopedic	49 (4.8%)
Respiratory	54 (5.3%)
Trauma/toxin	13 (1.3%)
Unknown	22 (2.1%)
Urogenital	76 (7.4%)

Of these 5 dogs, 2 were spayed female and 3 were male castrated dogs, and breeds included one each of an Australian cattle dog, a Boston terrier mix, Jack Russel terrier, miniature poodle, and a pug. For the remaining dogs, the age was 11 years (0.2-18 years), body weight was 19 kg (1.3-108 kg), with 453 spayed females, 445 castrated males, 92 intact males, 29 intact females, and 1 intersex dog. Breeds with more than 25 dogs represented included 74 Labrador retrievers, 43 German shepherd dogs, 33 golden retrievers, and 32 Chihuahuas. The primary clinical diagnosis category for all study dogs is summarized in Table 2.

3.2 | Machine learning model output

Of the 1025 dogs enrolled in the study, 14 were MLM positive for HA and the remaining 1011 were MLM negative for HA (Figure 1). Of the 14 dogs that were MLM positive, 2 dogs were euthanized before cortisol was measured and necropsy was not performed. The MLM correctly classified the 5 dogs that were newly diagnosed with HA based on serum cortisol concentration <2 μg/dL 1 hour after ACTH stimulation. This included 4 dogs that were treated for GMDH, two of which had aldosterone measured 1 hour after ACTH stimulation, confirming aldosterone deficiency. One dog was treated for GDH, with GMDH ruled out based on aldosterone measurement after ACTH stimulation test. Three of the 14 dogs that were MLM positive had a baseline cortisol ≤2 μg/dL, consistent with HA, but the dogs were euthanized or treated with glucocorticoids before HA could be confirmed with an ACTH stimulation test. Excluding the 2 dogs that did not have serum cortisol measured, and considering the three dogs that had a low baseline cortisol without an ACTH stimulation test as "HA suspects," the positive likelihood ratio ranged from 145, considering HA suspects as negative, and 254, considering HA suspects as positive. The positive percent agreement between the MLM and diagnosis made by clinicians was 100% (95% CI, 56.5%-100%). Four of the

FIGURE 1   Consort diagram.
Machine learning model
prediction results for all dogs
enrolled in the study and their HA
associated clinical diagnosis.
Figure made with Biorender.com
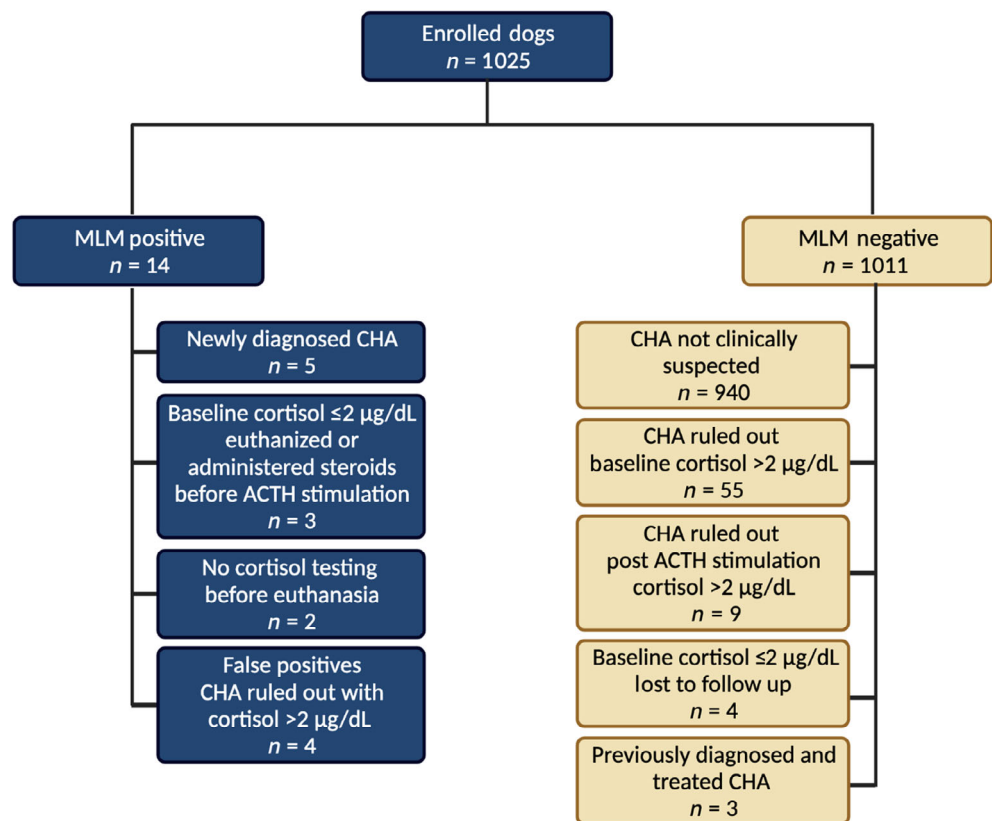
14 dogs that were MLM positive had cortisol levels >2 μg/dL and were categorized as false positive MLM results and included 2 dogs with gastrointestinal disease, 1 with ascites secondary to heart failure, and 1 dog with neurologic disease and profound polydipsia.

Of the 1011 dogs that were MLM negative for HA, 55 dogs had HA definitively ruled out with a baseline cortisol and 9 had the disease ruled out with an ACTH stimulation test. Three dogs had been diagnosed and treated for HA before the study period, and all had MLM negative results consistent with resolution of the clinicopathologic changes consistent with untreated HA. None of the dogs with a negative MLM result were newly diagnosed with HA by the attending clinician. Four dogs with negative MLM results had baseline cortisol ≤2 μg/dL without follow up cortisol testing at our institution and were assigned a clinical diagnosis other than HA by the attending clinician. Follow up medical information was available for 2 dogs. One was euthanized for prostatic carcinoma 2 months later, and the other had no diagnosis of or further clinical suspicion of HA 4 years later. The negative percent agreement was 99.3% (95% CI, 98.6%-99.7%). The sex distribution ($P = .9$), age ($P = .1$), and body weight ($P = .4$) of dogs did not differ between MLM classification groups (Table 1).

4 | DISCUSSION

In this study, we used a previously described MLM, that was trained using clinicopathologic data from dogs with a high pretest probability of HA, to screen for HA in an entire hospital population with low pretest probability of HA. We demonstrate that using CBC and serum biochemistry data, this MLM was as effective in screening for HA as clinicians are despite the latter having access to additional clinical data. Of note, cases included in this study were evaluated by clinicians at various levels of training and experience, in a hospital that receives primary, secondary, and tertiary referral cases.

Machine learning methods have been applied to a wide range of clinical syndromes in people and veterinary medicine[9-13] and can detect subtle patterns that allow for the early detection of disease. The gold standard diagnostic test for HA, serum cortisol measurement, has several shortcomings including the need for a clinician to recognize the need for the diagnostic test, prolonged turnaround time in some settings, increased cost to the client, and the need for specialized laboratory equipment to measure. Leveraging routinely collected clinicopathologic data and applying machine learning methods identified patterns that can streamline the diagnostic process and flag patients that are considered high risk for HA for confirmatory testing. Because recognition of HA as a differential diagnosis is challenging, automated methods to identify at risk dogs in a general hospitalization population are needed.

In this study, the prevalence of HA was 0.06%, which is similar to previous reports.[5,8,14] With a low prevalence, even diagnostic tests with high accuracy will incur false positives when screening an entire population indiscriminately, rather than only those with a high pretest probability of disease. Here, a low false positive rate was incurred, suggesting that positive MLM results should be followed by confirmatory tests. From a clinical perspective, the consequences of a false positive result are acceptable, as the confirmatory testing is noninvasive and cost effective. In contrast, recent studies have demonstrated that as a screening test, even when used in a population with higher pretest probability, the baseline cortisol test is associated with a higher false positive rate.[1,15] The MLM might therefore be used not

only to automatically screen patients for HA, but also refine test selection and save money in HA suspects.

The false positive MLM results in this group of dogs encompassed cases with cavitary effusions and gastrointestinal disease. This is similar to previous findings that indicate this model is predisposed to misclassifying these types of disease process, which likely represents a high similarity of CBC and serum chemistry panel findings between HA and these disease processes.[1]

A limitation of this study is the lack of gold standard cortisol testing in all dogs with a negative MLM result, and instead comparison to the diagnosis made at a veterinary medical teaching hospital where access to specialist veterinarians and advanced diagnostics are readily available. Without gold standard cortisol results in every dog, sensitivity and specificity could not be calculated. Instead, we report the positive and negative percent agreement relying upon the judgment of the attending clinician and their clinical suspicion and available cortisol measurements, rather than cortisol testing alone. All dogs that were diagnosed with HA in this study were MLM positive, resulting in a positive percent agreement of 100%. It is possible that a diagnosis of HA was missed by the attending clinician and the MLM, which would decrease the sensitivity of the MLM. Further, the negative percent agreement was >99%, indicating that the MLM has similar accuracy to a cohort of clinicians with various levels of training and experience at ruling out HA. Further assessment with large scale cortisol testing would be required to fully assess the sensitivity and specificity when used as a screening tool in a general population of dogs having blood work performed. An additional limitation is that some of the dogs with positive MLM results were euthanized or treated with corticosteroids before complete cortisol testing, which limited the ability to fully evaluate the MLM positive dogs.

This MLM tool can augment clinical decision making by incorporating objective clinicopathologic data to individualize the risk assessment of dogs for HA with high agreement compared to clinicians with various levels of training and experience, regardless of the subclassification of HA type (GMDH or GDH). False positive MLM results occur at a clinically acceptable rate when this model is applied to a general hospital population with a wide variety of clinical presentations and a low HA prevalence.

ACKNOWLEDGMENT

CONFLICT OF INTEREST DECLARATION

Authors Krystel L. Reagan and Chen Gilor are inventors on a pending patent that incorporates this trained model. No other authors have a conflict of interest.

OFF-LABEL ANTIMICROBIAL DECLARATION

Authors declare no off-label antimicrobial administration.

INSTITUTIONAL ANIMAL CARE AND USE COMMITTE (IACUC) OR OTHER APROVAL DECLARATION

Authors delcare no IACUC or other approval was neeed.

HUMAN ETHICS APPROVAL DECLARATION

Authors declare human ethics approval was not needed for this study.

ORCID

Krystle L. Reagan https://orcid.org/0000-0003-3426-6352
Chen Gilor https://orcid.org/0000-0003-0393-4135

REFERENCES

1. Reagan KL, Reagan BA, Gilor C. Machine learning algorithm as a diagnostic tool for hypoadrenocorticism in dogs. *Domest Anim Endocrinol*. 2020;72:106396.
2. Borin-Crivellenti S, Garabed RB, Moreno-Torres KI, Wellman ML, Gilor C. Use of a combination of routine hematologic and biochemical test results in a logistic regression model as a diagnostic aid for the diagnosis of hypoadrenocorticism in dogs. *Am J Vet Res*. 2017;78:1171-1181.
3. Peterson ME, Kintzer PP, Kass PH. Pretreatment clinical and laboratory findings in dogs with hypoadrenocorticism: 225 cases (1979-1993). *J Am Vet Med Assoc*. 1996;208:85-91.
4. Wakayama J, Furrow E, Merkel L, et al. A retrospective study of dogs with atypical hypoadrenocorticism: a diagnostic cut-off or continuum? *J Small Anim Pract*. 2017;58:365-371.
5. Hanson J, Tengvall K, Bonnett B, et al. Naturally occurring adrenocortical insufficiency—an epidemiological study based on a Swedish-insured dog population of 525,028 dogs. *J Vet Intern Med*. 2016;30:76-84.
6. Seth M, Drobatz K, Church D, et al. White blood cell count and the sodium to potassium ratio to screen for hypoadrenocorticism in dogs. *J Vet Intern Med*. 2011;25:1351-1356.
7. Adler JA, Drobatz KJ, Hess RS. Abnormalities of serum electrolyte concentrations in dogs with hypoadrenocorticism. *J Vet Intern Med*. 2007;21:1168-1173.
8. Hauck C, Schmitz SS, Burgener IA, et al. Prevalence and characterization of hypoadrenocorticism in dogs with signs of chronic gastrointestinal disease: a multicenter study. *J Vet Intern Med*. 2020;34:1399-1405.
9. Bradley R, Tagkopoulos I, Kim M, et al. Predicting early risk of chronic kidney disease in cats using routine clinical laboratory tests and machine learning. *J Vet Intern Med*. 2019;33:2644-2656.
10. Li WT, Ma J, Shende N, et al. Using machine learning of clinical data to diagnose COVID-19: a systematic review and meta-analysis. *BMC Med Inform Decis Mak*. 2020;20:1-13.
11. De Bruyne S, Speeckaert MM, Van Biesen W, et al. Recent evolutions of machine learning applications in clinical laboratory medicine. *Crit Rev Clin Lab Sci*. 2021;58:131-152.
12. Schofield I, Brodbelt DC, Niessen SJ, et al. Development and internal validation of a prediction tool to aid the diagnosis of Cushing's syndrome in dogs attending primary-care practice. *J Vet Intern Med*. 2020;34:2306-2318.
13. Biourge V, Delmotte S, Feugier A, Bradley R, McAllister M, Elliott J. An artificial neural network-based model to predict chronic kidney disease in aged cats. *J Vet Intern Med*. 2020;34:1920-1931.
14. Kelch WJ. *Canine Hypoadrenocorticism (Canine Addison's Disease): History, Contemporary Diagnosis by Practicing Veterinarians, and Epidemiology*. Knoxville, Tennessee: The University of Tennessee; 1996.
15. Gallego AF, Gow AG, Boag AM. Evaluation of resting cortisol concentration testing in dogs with chronic gastrointestinal signs. *J Vet Intern Med*. 2022;36:525-531.