Received: 14 August 2023 | Accepted: 25 February 2024

DOI: 10.1111/evj.14087

ORIGINAL ARTICLE

Equine Veterinary Journal WILEY

Comparison of veterinarians and a deep learning tool in the diagnosis of equine ophthalmic diseases

Annabel Scharre[1] | Dominik Scholler[1] | Stefan Gesell-May[2] | Tobias Müller[3] | Yury Zablotski[4] | Wolfgang Ertel[5] | Anna May[1]

[1]Equine Clinic, Ludwig Maximilians University, Oberschleissheim, Germany

[2]Center for Equine Ophthalmology, Munich, Germany

[3]anirec, Munich, Germany

[4]Clinic for Ruminants, Ludwig Maximilians University, Oberschleissheim, Germany

[5]Institute for Artificial Intelligence, Ravensburg-Weingarten University, Weingarten, Germany

Correspondence
Anna May, Equine Clinic, Ludwig Maximilians University, Oberschleissheim, Germany.
Email: anna.may@lmu.de

Abstract

Background/Objectives: The aim was to compare ophthalmic diagnoses made by veterinarians to a deep learning (artificial intelligence) software tool which was developed to aid in the diagnosis of equine ophthalmic diseases. As equine ophthalmology is a very specialised field in equine medicine, the tool may be able to help in diagnosing equine ophthalmic emergencies such as uveitis.

Study design: In silico tool development and assessment of diagnostic performance.

Methods: A deep learning tool which was developed and trained for classification of equine ophthalmic diseases was tested with 40 photographs displaying various equine ophthalmic diseases. The same data set was shown to different groups of veterinarians (equine, small animal, mixed practice, other) using an opinion poll to compare the results and evaluate the performance of the programme. Convolutional Neural Networks (CNN) were trained on 2346 photographs of equine eyes, which were augmented to 9384 images. Two hundred and sixty-one separate unmodified images were used to evaluate the trained network. The trained deep learning tool was used on 40 photographs of equine eyes (10 healthy, 12 uveitis, 18 other diseases). An opinion poll was used to evaluate the diagnostic performance of 148 veterinarians in comparison to the software tool.

Results: The probability for the correct answer was 93% for the AI programme. Equine veterinarians answered correctly in 76%, whereas other veterinarians reached 67% probability for the correct diagnosis.

Main limitations: Diagnosis was solely based on images of equine eyes without the possibility to evaluate the inner eye.

Conclusions: The deep learning tool proved to be at least equivalent to veterinarians in assessing ophthalmic diseases in photographs. We therefore conclude that the software tool may be useful in detecting potential emergency cases. In this context, blindness in horses may be prevented as the horse can receive accurate treatment or can be sent to an equine hospital. Furthermore, the tool gives less experienced veterinarians the opportunity to differentiate between uveitis and other ocular anterior segment disease and to support them in their decision-making regarding treatment.

KEYWORDS
artificial intelligence, blindness, deep learning, equine uveitis, horse, ophthalmology

1 | INTRODUCTION

Equine uveitis is a threatening disease which can lead to blindness or loss of the eye in affected horses.[1–3] Therefore, diagnosing uveitis or other ophthalmic diseases prompting quick decisions on whether the condition is an emergency and if the horse needs specialised treatment, is very important for the equine practitioner. Because early detection can be critical to save affected eyes, image analysis using deep learning (artificial intelligence) can help to diagnose ophthalmic conditions in horses. A software tool, which was developed for this reason, proved to be reliable in diagnosing various eye diseases.[4]

In recent years, the application of artificial intelligence has started to transform daily routines through applications that can be used by everyone, such as photo or image creation, speech recognition, translation of languages, or text formation. Deep learning, as a part of machine learning, has made vast advances as assistant tools in human and veterinary medicine in the past years.[5–8] Deep learning tools are composed of multiple processing layers and are capable of processing large amounts of data. The technique uses artificial neural networks, which are algorithms that resemble the functions of the human brain by structuring unstructured data. Deep learning algorithms can be trained with photographs of diseases and then categorise patterns, which is especially useful in fields with repetitive actions such as diagnostic imaging, histology, or cytology.[5,6] The main difference to other computer programmes is the fact that the filter criteria of the 'layers' are created autonomously by the algorithm itself and not by a software developer. Examples for this technique in human ophthalmology include diagnostic tools for retinal pathologies,[7–9] macular degeneration or glaucomatous optic neuropathy.[10,11] Systematic reviews on deep learning algorithms found that the tools can have equivalent sensitivity and specificity to health-care professionals.[12] Concerns are raised on whether the findings are generalisable and can be applied to the real-world setting.

In this study, a deep learning software tool trained to diagnose equine ophthalmic conditions (healthy, uveitis, other ophthalmic diseases) was compared with veterinarians of various specialisations. The same dataset (40 images of equine eyes) was shown to the deep learning tool and the veterinarians and had to be categorised into healthy, uveitis, and other ophthalmic diseases.

The aim of this study was to determine whether the deep learning tool showed equivalent results to veterinarians when eye conditions were diagnosed on images, to evaluate if the deep learning tool may help to detect ophthalmic emergencies in the future.

2 | MATERIALS AND METHODS

2.1 | Clinical material and development of the AI software tool

For the development of a deep learning tool to diagnose equine ophthalmic diseases Convolutional Neural Networks (CNNs) were used. Photographs were taken from various angles and displayed a wide range of equine ophthalmic diseases. Horses included in the preceding study[4] were examined according to the same protocol: the pupil was dilated with mydriatics (Tropicamide), and the horses were examined via routine direct (WelchAllyn® direct ophthalmoscope) and indirect ophthalmoscopy (HEINE Omega 500 LED indirect binocular ophthalmoscope and HEINE indirect ophthalmoscopy lens 20D), as well as slit lamp biomicroscopy (Keeler PSL Classic LED) and tonometry (Icare® Tonovet). If necessary, sedation or eye lid blocks were used. Ophthalmologic findings and diagnoses were determined by a board-certified internal medicine specialist and a veterinarian with extensive experience in equine ophthalmology. In total 2346 training images (90% of the dataset) were used. The data was expanded to 9384 images using augmentation. To validate the data, 10% (261 images) were used which were presented to the tool for the first time. Cross validation revealed an accuracy of 99.82% in the training data and an accuracy of 96.66% in the validation data (distinction between the three categories healthy, uveitis, other).

2.2 | Selection of photographs for opinion poll

Only images in which significant ophthalmologic findings were visible and in which both examiners agreed on all findings and the diagnosis, were included in the study. The data set used in the opinion poll and for assessment of the deep learning tool comprised 10 photographs of healthy eyes, 12 photographs showing uveitis, and 18 images of other diseases. The photographs used for the poll had not been assessed by the deep learning tool before.

2.3 | Categorisation of equine ophthalmic diseases/definition of uveitis

Inclusion criteria for uveitis therefore were typical findings of inner eye involvement such as fibrin or flare in the anterior chamber, miosis, inflammatory deposits on anterior or posterior lens capsule or in vitreous body (irregularities in the pupil), a turbid greenish fundic reflex in the acute cases, as well as synechiae, cataract and depositions in the vitreous body in the chronic cases. The 'other' diseases group consisted of horses showing various types of keratitis, corneal ulcers, or glaucoma, which were diagnosed by ophthalmic examination and clearly visible on the photographs. Horses with unremarkable ophthalmic examination and photographs with no visible pathologic findings were categorised as healthy.

2.4 | Selection of veterinarians for opinion poll

A commercial survey software (UmfrageOnline; www.umfrageonline.com) was used to create the survey. Approximately 200 emails with invitations to participate in the survey were sent to universities in Germany and other European countries, as well as private practices and animal hospitals, so that a large number of veterinarians was reached. Since it is not possible to trace how often the email was forwarded among colleagues, the exact number of veterinarians contacted is unknown.

The link and the questionnaire were active from 1 April to 23 May 2021. The questionnaire included 45 questions (Supplementary Item 1). The first five questions addressed the veterinary field, professional experience, and possible professional titles such as veterinary specialist or diplomate. In addition, the first question asked for consent to data storage. The next 40 questions assessed the evaluation of equine eye images by the veterinarians. Each question included a photo of a horse's eye. Participants could choose one of three possible answers for each photo. The answers were identical for all 40 photos: healthy eye, uveitis, and other eye diseases. The 40 selected photos for the survey were numbered randomly. To limit the time for completion of the survey to a maximum of 10 min, a selection of 40 photos was used.

In total, 237 veterinarians participated in the survey, but only 148 participants completed the questionnaire, so that their answers could be used for statistical analysis. The participants were divided into equine veterinarians (59%), and non-equine veterinarians, consisting of small animal (18%), mixed practice (20%) and other veterinarians working with poultry or ruminants (3%). In each group, the probability of a correct answer was assessed.

2.5 | Data analysis

Given the exploratory nature of our study, no power analysis was conducted. Due to the presence of repeated measures a mixed effects logistic regression model with both individual person and photo ID as random effects and assessor group (AI software vs. equine and non-equine veterinarian groups) as a fixed effect was chosen for analysis ('glmer' function with binomial family from 'lme4' R package). The variable correct distinguishes between the correct answer (code = 1) and one of the two incorrect answers (code = 0). All contrasts (differences) between particular groups were assessed after model-fitting by the estimated marginal probabilities and odds ratios (R package—emmeans) with Tukey *p*-value correction for multiple comparisons. Results with a *p*-value <0.05 were considered statistically significant. Data analysis was performed using R 4.2.1 (2022-06-23).

3 | RESULTS

3.1 | Assessment of the AI software tool (group 1)

The deep learning tool reached a 93% (95% CI: 72%–99%) probability for the correct answer when the 40 equine eye images were used to test the device. Certainties for each diagnosis are shown in Table S1.

When reviewing the specific photographs that were misdiagnosed by the deep learning tool, it was found that the tool diagnosed one keratitis eye as false healthy (2.5% false negative results), while one healthy eye was diagnosed as 'other' disease (2.5% false positive) (Figures 1 and 2). Two images of 'uveitis' eyes were falsely categorised as 'other' disease.

3.2 | Assessment of equine and non-equine veterinarians and the AI software tool

Equine veterinarians correctly diagnosed 76% (95% CI: 69%–81%) of the diseases shown in the photographs. The other veterinarians (mixed practice veterinarians: 19.59%, small animal veterinarians: 17.57%, and other: 3.38%) (group 3) made a correct diagnosis in 67% (95% CI: 59%–74%) of the diseases shown in the photographs.

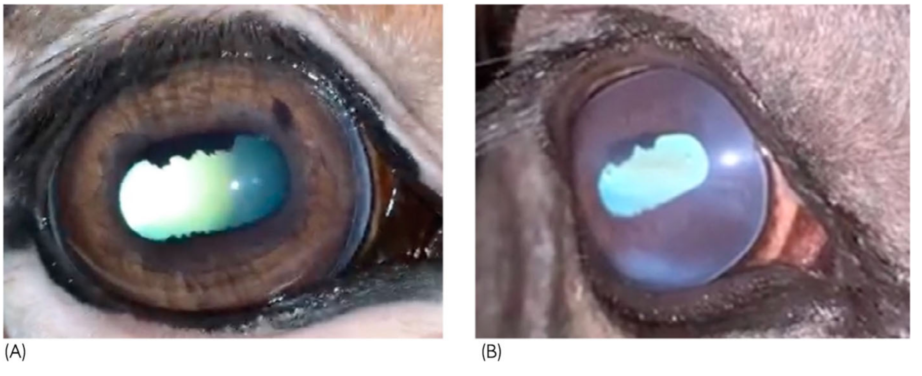

FIGURE 1 Two of the photographs misdiagnosed by the deep learning tool: (A) was falsely diagnosed as uveitis while being healthy, whereas (B) was diagnosed as healthy, though it showed signs of a keratitis (ventral corneal opacity).

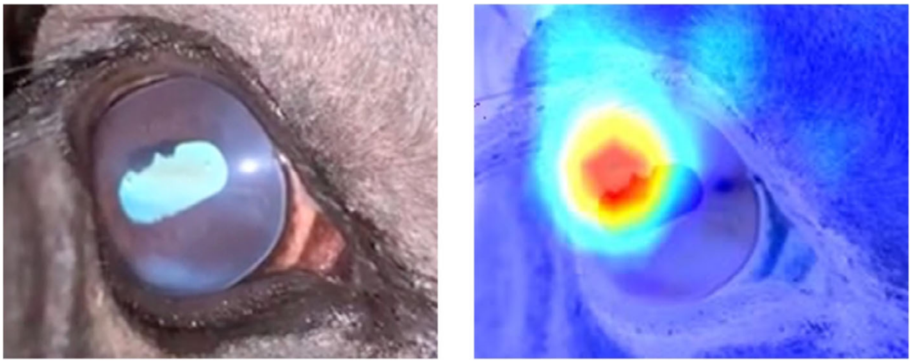

FIGURE 2 The photograph falsely diagnosed as healthy by the deep learning tool with original photograph on the left and relevant areas scrutinised by the tool superimposed with the original photograph on the right. The superimposition shows the area the deep learning tool uses to categorise the photographs. As the area does not match the changes in the cornea the deep learning programme did not diagnose the condition correctly. This may be because the photograph was taken from a slightly different angle than those images which the deep learning tool was trained with.

TABLE 1 OR, 95% CI and *p*-values (corrected with Tukey method for multiple comparisons) comparing the likelihood of achieving the correct diagnoses between the deep learning tool and equine and non-equine veterinarians evaluating 40 photographs of equine ophthalmic diseases.

Comparison	Odds ratio	95% confidence interval	*p*-value
Equine veterinarians versus deep learning tool	0.23	0.03, 1.63	0.2
Non-equine veterinarians versus deep learning tool	0.15	0.02, 1.08	0.06
Non-equine versus equine veterinarians	0.66	0.49, 0.87	0.001

The odd ratios, the 95% confidence intervals and *p*-values for comparisons between the tool and each group of veterinarians and equine versus non-equine veterinarians are shown in Table 1.

4 | DISCUSSION

In a preceding study, the deep learning tool proved to be suitable to differentiate between photographs showing healthy eyes and those showing ophthalmic diseases.[4] To evaluate whether the programme is equivalent to ophthalmic diagnosis of field veterinarians and may therefore serve as an aid in veterinary practice, a comparison of the tool and veterinarians was conducted. Analyses on the diagnostic accuracy of health-care professionals versus deep learning algorithms using medical imaging have been performed in various medical fields, not only in human medicine[13–15] but also in veterinary medicine. For example, one study compared the performance of human experts and a deep learning tool regarding cytologic scoring of equine exercise-induced pulmonary haemorrhage and found out that AI can improve the reproducibility and routine applicability of cytologic scoring.[16] Another study compared AI to a veterinary radiologist's diagnosis of canine cardiogenic pulmonary oedema and confirmed a high accuracy, sensitivity, and specificity of the AI-based software.[17] Overall, most of the studies concluded that deep learning algorithms can reach clinical-level accuracy.[18,19] But there are also important methodological deficiencies which must be considered.[20] Similar to our study, most studies assessed deep learning algorithms as isolated methods, which does not reflect the situation in clinical practice. Most of the studies, including our approach, did not provide health-care professionals with additional clinical information, as they would have in clinical practice. This isolated approach limits the ability to transfer the results to the real-life medical situation unless the deep learning algorithm is used for mass screening.[21] There may be a difference in veterinary medicine regarding this point, as horse owners may not provide the veterinarian with sufficient clinical information, and the history may also be unknown in some of the cases. In this context, the access to sufficient patient information for establishing correct diagnoses has been pointed out.[19]

Another flaw in many studies was the fact, that health-care professionals and the deep learning algorithms were not provided with the same test dataset.[13] In the current study, the deep learning tool and the veterinarians were all provided with the same photographs of equine eyes.

While studies on deep learning algorithms using high-quality images show identical or even greater performance than that of human doctors in diagnosing diseases, deep learning usually performs poorly when low-quality images are used.[22] It is not known whether human doctors also perform poorly in assessing low-quality images. If the results of health-care professionals evaluating low-quality images are better than deep learning, it exposes a weakness of deep learning systems. If the performance of deep learning systems is similar to human doctors, there is an indication that detection of diseases from low-quality images may be difficult. One study compared the performance of deep learning systems with that of cornea specialists in diagnosing corneal diseases from low-quality slit lamp images.[22] The study concluded that cornea specialists performed better than the deep learning system which was trained only on high-quality images. A system which was trained on both low-quality and high-quality images performed better than the previously mentioned system but was still inferior to senior corneal specialists. The researchers suggested training deep learning systems also with low-quality images to reduce this performance gap.[22] The deep learning tool for diagnosis of equine ophthalmic diseases had a large sample size of low-quality images in the training set as it was trained on various images that had been taken under different light conditions and from various angles. Therefore, its performance on low-quality images with sufficient diagnostic certainty was good.[4] In the current study excellent performance was observed when the same dataset of 40 equine eye images was evaluated by veterinarians and the tool. Because the deep learning tool had been trained with various images before, whereas the veterinarians were presumably not used to looking solely at photographs of the eye, the tool's performance was better than that of the veterinary professionals. In the current study, performance of the equine veterinarians was superior to that of the other veterinarians, and comparable to the deep learning tool. Possible explanations are that equine veterinarians are better in diagnosing equine ophthalmic diseases and that they may have more experience in analysing low-quality images of equine eyes, for example, through photographs sent by horse owners.

A main limitation of the study was the fact that images only showed the anterior segments of the eye, whereas the posterior parts (fundus region and retina) were not visible. Nevertheless, the tool can detect changes in the anterior segment and to a lesser extent the lens. In cases of emergency (e.g., corneal ulcer, acute uveitis) the tool still proved to be useful, as it is capable to detect many ophthalmic features associated with these diseases, for example, fibrin in the anterior chamber or an irregular pupil. Emergencies of the posterior segment (e.g., retinal detachment, inflammation of the optic nerve) will not be diagnosed by the tool though. In this study, four photographs were misdiagnosed by the deep learning tool, but the system recognised three out of them correctly as diseased eyes. The photograph of the diseased eye which was falsely categorised as healthy showed clear signs of a keratitis (corneal opacity) but was taken from an angle which was less represented in the training data set of the tool.

Another limitation is the fact that photographs were categorised solely by two clinicians (one board-certified equine internal medicine specialist and a veterinarian with extensive experience in equine ophthalmology). To categorise the eyes correctly, diagnosis was made by a thorough ophthalmic examination and not just by looking at the images. Inclusion criteria for uveitis therefore were typical findings of inner eye involvement such as fibrin or flare in the anterior chamber, miosis, inflammatory deposits on anterior or posterior lens capsule, a turbid greenish fundic reflex in the acute cases, as well as synechiae, cataract and depositions in the vitreous body in the chronic cases.[4] Photographs without clear signs of the classic or insidious form of uveitis were excluded from the selection, which may have caused a bias. But both veterinarians and the deep learning tool may have equally benefited from 'easier' to detect images with clear signs of disease. The other ophthalmic conditions consisted of glaucomas, various types of keratitis, and corneal ulcers. Pictures in the healthy group did not show any ophthalmologic findings. The test dataset of 40 images for evaluation by the deep learning tool and the veterinarians consisted of photographs with complete agreement in findings and diagnosis between the two clinicians.

The deep learning tool tested in this study showed to be a useful solution for detecting different eye conditions of the equine patient. It can help to identify emergency conditions and may therefore be an aid on whether the horse needs specialised treatment or surgery, for example, corneal flaps. Whereas it should not be used as a sole 'opinion' it can serve as an additional tool for a full ophthalmic examination performed by a veterinarian. Further information on previous medication, general horse health, and location of the horse has to be considered when performing an ophthalmic exam. As the deep learning tool proved to be at least equivalent to equine veterinarians in categorising and diagnosing images of equine ophthalmic diseases, it can be especially useful in areas with little emergency veterinary coverage. As it was superior to less specialised veterinarians it may serve as a 'second opinion' diagnostic tool. Nonetheless the deep learning tool should be used carefully and well chosen, it can never substitute for a thorough examination by a veterinarian and should therefore stay in expert hands, rather than be provided publicly, which could lead to misdiagnosis by medical laypersons and insufficient or wrong treatment of equine patients. Making a diagnosis of uveitis based on photographs can be of low sensitivity and miss many early cases of uveitis in the absence of other diagnostics such as tonometry and a skilled examiner. Furthermore Tamori et al. pointed out that there is still a lack of acceptance of using deep learning tools in medicine in public, while there is a higher acceptance among doctors[23] and medical students.[24] To increase public acceptance of using deep learning tools in medicine, further investigation and improvement of these tools may be helpful. In this case, further studies comparing the deep learning tool with veterinarians using photographs of ophthalmic diseases with only subtle findings or implementing a wider range of ophthalmic diseases to improve accuracy of the diagnostic deep learning tool could be helpful.

Artificial intelligence can improve diagnostics in almost all areas of human and veterinary medicine if used carefully.[6,8,12,20,25,26] However, exaggerated claims on whether a system is superior to clinicians may mislead the public and may potentially lead to inappropriate treatment for the patient.[27,28] With the increasing digitalisation of medical data, deep learning systems can be trained better and find suitable patterns for improving diagnostics in equine medicine conditions. Implemented in thorough examinations by health-care and veterinary professionals, specialised deep learning programmes can make human and veterinary medicine more accurate, faster and can therefore improve outcome for the human and veterinary patient.

FUNDING INFORMATION

No funding has been obtained for this study.

ACKNOWLEDGEMENTS

Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

Tobias Müller is employed with anirec, a company whose products are used in the study, and which may enter the market at some stage.

AUTHOR CONTRIBUTIONS

Annabel Scharre: Conceptualization; data curation; investigation; methodology; writing – original draft; writing – review and editing. **Dominik Scholler:** Data curation; investigation; methodology; writing – original draft; writing – review and editing. **Stefan Gesell-May:** Conceptualization; data curation; formal analysis; methodology; project administration; validation. **Tobias Müller:** Conceptualization; data curation; formal analysis; investigation; software. **Yury Zablotski:** Validation; visualization; writing – review and editing. **Wolfgang Ertel:** Conceptualization; data curation; methodology. **Anna May:** Conceptualization; data curation; formal analysis; investigation; methodology; project administration; supervision; validation; writing – original draft; writing – review and editing.

DATA INTEGRITY STATEMENT

Anna May has full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

ETHICAL ANIMAL RESEARCH

According to the Ethics Committee of LMU Munich, no ethics approval was necessary for this study.

INFORMED CONSENT

Survey respondents gave consent for participation in the study.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/evj.14087.

DATA AVAILABILITY STATEMENT

The data that support the findings will be available in Open Data LMU at https://data.ub.uni-muenchen.de/443/ and https://doi.org/10.5282/ubm/data.443 following an embargo from the date of publication.

ORCID

Stefan Gesell-May https://orcid.org/0000-0002-0834-6677
Anna May https://orcid.org/0000-0003-3038-5952

REFERENCES

1. Gerding JC, Gilger BC. Prognosis and impact of equine recurrent uveitis. Equine Vet J. 2016;48:290–8.
2. Gilger B, Hollingsworth S. Diseases of the uvea, uveitis, and recurrent uveitis. In: Gilger B, editor. Equine ophthalmology. 3rd ed.; Ames, Iowa: Wiley-Blackwell; 2016. p. 369–415.
3. Gilger BC. Equine recurrent uveitis: the viewpoint from the USA. Equine Vet J. 2010;42(S37):57–61.
4. May A, Gesell-May S, Müller T, Erte W. Artificial intelligence as a tool to aid in the differentiation of equine ophthalmic diseases with an emphasis on equine uveitis. Equine Vet J. 2022;54(5):847–55.
5. Yala A, Lehman C, Schuster T, Portnoi T, Barzilay R. A deep learning mammography-based model for improved breast cancer risk prediction. Radiology. 2019;292:60–6.
6. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA. 2016;316:2402–10.
7. Quellec G, Charrière K, Boudi Y, Cochener B, Lamard M. Deep image mining for diabetic retinopathy screening. Med Image Anal. 2017;39:178–93.
8. Alyoubi WL, Shalash WM, Abulkhair MF. Diabetic retinopathy detection through deep learning techniques: a review. Inform Med Unlocked. 2020;20:100377.
9. Goodfellow IBY, Courville A. Deep learning. Vol 329ff. Cambridge, MA: MIT Press; 2016. Accessed 17 September 2023. www.deeplearningbook.org.
10. Li Z, He Y, Keel S, Meng W, Chang RT, He M. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. Ophthalmology. 2018;125:1199–206.
11. Lee CS, Baughman DM, Lee AY. Deep learning is effective for the classification of OCT images of normal versus age-related macular degeneration. Ophthalmol Retina. 2017;1:322–7.
12. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542:115–8.
13. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Brunseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health. 2019;1:e271–97.
14. Becker AS, Marcon M, Ghafoor S, Wurnig MC, Frauenfelder T, Boss A. Deep learning in mammography: diagnostic accuracy of a multipurpose image analysis software in the detection of breast cancer. Invest Radiol. 2017;52:434–40.
15. Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. BMJ. 2020;368:m689.
16. Bertram CA, Marzahl C, Bartel A, Stayt J, Bonsembiante F, Beeler-Marfisi J, et al. Cytologic scoring of equine exercise-induced pulmonary hemorrhage: performance of human experts and a deep learning-based algorithm. Vet Pathol. 2023;60:75–85.
17. Kim E, Fischetti AJ, Sreetharan P, Weltman JG, Fox PR. Comparison of artificial intelligence to the veterinary radiologist's diagnosis of canine cardiogenic pulmonary edema. Vet Radiol Ultrasound. 2022;63:292–7.
18. Zhai S, Wang H, Sun L, Zhang B, Huo F, Qiu S, et al. Artificial intelligence (AI) versus expert: a comparison of left ventricular outflow tract velocity time integral (LVOT-VTI) assessment between ICU doctors and an AI tool. J Appl Clin Med Phys. 2022;23:e13724.
19. Graf M, Knitza J, Leipe J, Krusche M, Welcker M, Kuhn S, et al. Comparison of physician and artificial intelligence-based symptom checker diagnostic accuracy. Rheumatol Int. 2022;42:2167–76.
20. Aubreville M, Bertram CA, Marzahl C, Gurtner C, Dettwiler M, Schmidt A, et al. Deep learning algorithms out-perform veterinary pathologists in detecting the mitotically most active tumor region. Sci Rep. 2020;10:16447.
21. Ter Riet G, Bachmann LM, Kessels AG, Khan KS. Individual patient data meta-analysis of diagnostic studies: opportunities and challenges. Evid Based Med. 2013;18:165–9.

22. Li Z, Jiang J, Qiang W, Guo L, Liu X, Weng H, et al. Comparison of deep learning systems and cornea specialists in detecting corneal diseases from low-quality images. iScience. 2021;24:103317.
23. Tamori H, Yamashina H, Mukai M, Morii Y, Suzuki T, Ogasawara K. Acceptance of the use of artificial intelligence in medicine among Japan's doctors and the public: a questionnaire survey. JMIR Hum Factors. 2022;9:e24680.
24. Bisdas S, Topriceanu CC, Zakrzewska Z, Irimia A-V, Shakallis L, Subhash J, et al. Artificial intelligence in medicine: a multinational multi-center survey on the medical and dental students' perception. Front Public Health. 2021;9:795284.
25. Fraiwan MA, Abutarbush SM. Using artificial intelligence to predict survivability likelihood and need for surgery in horses presented with acute abdomen (colic). J Equine Vet. 2020;90:102973.
26. Alexeenko V, Jeevaratnam K. Artificial intelligence: is it wizardry, witchcraft, or a helping hand for an equine veterinarian? Equine Vet J. 2023;55:719–22.
27. Darcy AM, Louie AK, Roberts LW. Machine learning and the profession of medicine. JAMA. 2016;315:551–2.
28. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017;42:60–88.

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: Scharre A, Scholler D, Gesell-May S, Müller T, Zablotski Y, Ertel W, et al. Comparison of veterinarians and a deep learning tool in the diagnosis of equine ophthalmic diseases. Equine Vet J. 2024. https://doi.org/10.1111/evj.14087