UNIVERSITY OF CALGARY
FACULTY OF VETERINARY MEDICINE

This review accompanies the relevant episode of the Cutting Edge veterinary podcast. In each episode of this podcast, 3rd year students in the University of Calgary's veterinary medicine program fill you in on the most up-to-date literature and evidence-based practices on topics that matter to you, the practising veterinarian.

The Use of Artificial Intelligence in Small Animal General Practice

STUDENT: Katherine Fleury
FACULTY MENTOR: Dr. Patrick Whelan

Introduction

The technologies used in veterinary medicine are constantly changing and adapting to the world around us. From flushing a mare's uterus with kerosene to POCUS, how we approach animal health and well-being continues to evolve as new tools become available. Some examples include wearable medical devices for dairy cows that use an increase in activity as a proxy for estrus timing, motion detection to monitor respiration in dogs, and artificial intelligence (AI) to recognize patterns that we may miss (1). The use of AI is slowly becoming mainstream in the small-animal veterinary medicine world, resulting in earlier detection of abnormalities (1). AI can be described as a branch of computer science that uses systems to perform tasks that normally require human intelligence (2). While exiting, our increased reliance on this technology means it is important to understand what these tools are and how they work.

To understand how AI is being used in veterinary medicine, it is important to understand the difference between machine learning, neural networks, narrow AI, and generative AI. To start, machine learning is a subfield of AI that can perform tasks without being programmed to do so. Unlike a software system that is static and requires updates to change how it works, machine learning is trained on large data sets and then uses what it has been taught to predict patterns in other, future data sets (2). This data can be anything: images, genomics, administrative data, or any other collection of large data where patterns can be distinguished (2,3). The algorithms are trained on this data, "taught" to notice changes to the patterns, and then use this learning to create classifications or predictions when presented with new data (3). The main objectives of machine learning in medical fields are to aid in classification, prediction, and discovery (e.g. new drugs or new disease phenotypes) (3). Machine learning can analyze large amounts of data much faster than humans while identifying patterns. Importantly, and unlike explicitly coded program algorithms, tools that utilize machine learning can dynamically adjust decision making steps. (4).

Machine learning encompasses three different types of learning: supervised learning, unsupervised learning, and semi-supervised learning (2). Supervised learning is the most common in medical fields and uses data sets that have been previously labelled to teach their algorithms. For example, radiographs that have been assessed by a human radiologist will be used to teach an AI radiograph technology the difference between normal and abnormal. The human-annotated data set is considered the gold standard in this case, and the outcome (i.e. diagnosis) is known before the model is trained (2). To help visualize this, imagine you are looking at two radiographs. One has a pulmonary nodule, and is labelled as "abnormal", while a second is clear, and is labelled "normal". Then, you are presented with a third, unlabelled radiograph with a pulmonary nodule. Based on the date you have just seen, you label it "abnormal". Supervised learning does this on a much larger scale, and to a more precise diagnosis.

On the other end of the spectrum, unsupervised learning occurs when a model is not pre-trained, but instead creates its own criteria to sort the data. This means that no labels are prescribed to the data before the model assesses it. This type of learning is helpful when trying to make sense of large data sets. Finally, semi-supervised learning is a combo of the aforementioned learning styles that can be helpful if some data is missing the outcome/diagnosis when the model is being trained (2).

Machine learning also encompasses neural networks. Neural networks are computer systems that have nodes which can be switched on and off to reach an outcome (2). It can be compared to the human brain: for example, imagine you are again looking at a radiograph. When you start scanning the image, your brain uses all your training from vet school and experience in the field to say yes or no to abnormalities present, and then further differentiates the types of abnormalities, what they may mean, and differential diagnoses (2). A neural network does the same thing, where interconnected prompts convert the data given into an output (2). The more layers the network can go through, the more the value of the output increases (2). Again, this can be related to the brain of a veterinarian: in the first year of veterinary school, you may be able to differentiate between normal and abnormal, but over the years you are able to follow the "abnormal" branch to a much more specific diagnosis. Neural networks follow their own branches based on how they are trained to come up with a final output.

Lastly, it is important to know the difference between narrow AI and generative AI: narrow AI is AI that has been trained for a specific task (reading radiographs, scanning blood smears, etc) while generative AI is a type of machine learning that can create data following text prompts (e.g. ChatGBT) (2,5).

Significance to the field

With all those definitions that fall outside of the Latin terminology we as veterinarians already had to learn, why care? As a veterinarian, you are not expected to know the ins and outs of an entirely different branch of science, but you are expected to know how tests within the field work. This is similar to any other diagnostic test; you do not need to know how to develop an ELISA test but you do need to know about antigens, antibodies, sensitivity and specificity (2). Beyond knowing how the tools work, you should be aware of increasing demands and future career opportunities, as well as increased time efficiency and insight.

As the amount of information available online grows, veterinary clients are becoming more efficient at researching potential diagnoses, tests, and diseases. With this increase in information, clients are also expecting that their veterinarians will be able to answer questions about the things they've read, including new technologies. One thing that is unique to AI is that the implementation has preceded scientific review. Scott et al. summarize this well: “Interestingly, clinical implementation of AI-based image interpretation technology has preceded many of the reviewed publications, and claimed capabilities are being scaled larger than anything described in the peer-reviewed literature. This rapid implementation prior to availability of validating results speaks to the ever-growing demand for expert image interpretation, the perception of business opportunity, and the rapid nature by which competition in the ML space can develop” (3). With the ever-growing field of AI just beginning to be adopted in general practice, these tools are not going away anytime soon. It can be assumed that there will be continuous learning opportunities about AI-based technology for veterinarians in the future, something vital for it’s safe implementation and proper interpretation (6). AI should not be feared as a tool that can replace medical practitioners, but instead as a tool that will allow veterinarians to work smarter (6).

Working smarter includes allocating your time effectively. With implemented AI technologies, there is the potential to free up time for both doctors and technicians. With increasing workloads causing burnout in the veterinary industry, having additional tools that decrease the amount of time spent completing diagnostic tests may improve veterinarian welfare (6). Moreover, time can be spent elsewhere, including talking with the client and having your hands on the patient. This helps clients see the value of veterinary services. For example, while an assessment is being run, you can stay in the room with your client and discuss next steps. Importantly, you may reach a diagnosis faster if you have in-clinic technologies; not having to wait for a formal radiology or cytology report can increase the time to diagnosis (7). Many common AI technologies also offer input from boarded specialists within their program design, allowing for feedback if any issues in the AI-generated reports are found. This is just one way in which AI may increase insight as it pertains to veterinary medicine, but we must also consider the data analysis functions provided by generative AI. If you have questions about common diseases you are seeing, from clinical presentations to prevalence, you may be able to use a generative program to assess trends in your own clientele much faster than if you had to analyze the data yourself.

Beyond demand and time efficiency, an important aspect of AI technologies is how they may impact your clinic and your clients financially. The following sections summarize some of the current uses of AI technologies in small animal medicine as well as future applications, providing the financial breakdown where applicable.

Current uses in the small animal setting

Diagnostic imaging

One of the most common uses for AI as it pertains to diagnostic imagining is in computer-aided diagnostics (8). This application falls into the field of radiomics: the use of machine learning to perform quantitative analysis using computer algorithms (4). AI trained to analyze radiographs can be used to detect abnormalities or to diagnose based on criteria provided (4). These technologies are designed to aid in the efficiency and accuracy of medical imaging and can be used with a veterinarian interpreting the image in tandem (4). Hennessey et al. completed a review of diagnostic imaging AI technologies in veterinary medicine in 2022 which included studies that assessed the application of these tools to musculoskeletal (MSK), thoracic, cardiac, pulmonary, abdominal, and central nervous system (CNS) images. As a whole, the performance of the algorithms reviewed had sensitivities, specificities, and accuracies ranging from 60-95% (4). These studies are important as they consider what types of cases AI can handle, the level of agreement an algorithm may have with a specialist, and how they can be applied for other uses.

The first use of machine learning in veterinary medicine was completed in 2013 by McEvoy et al., where it was used to identify the coxofemoral joint in ventrodorsal canine pelvic radiographs (9). Since then, several studies have been completed on the MSK system. Notably, Gomes et al. found that machine learning was suitable to complete the initial classification of canine coxofemoral joints when hip dysplasia was of concern (10). This is an important milestone as the researchers were able to discern that radiologist resources may be better allocated to more challenging classifications/mid-grade hips (10). Knowing the limits of any model is important, and when presented with a case, these kinds of studies help clinicians know when they may want to ask for a consult beyond what an AI computer diagnostic can generate.

AI models have also been assessed for their efficacy when trained and presented with thoracic and cardio radiographs. Boissady et al. found that the model they tested on vertebral heart scoring (VHS) had an intraclass correlation coefficient of 0.998 when compared to the average of VHS assigned by two cardiologists (11). This result shows applicability to reduce interobserver variability when performing serial VHS measurements and may be extended to other fields in the future (11).

In the abdomen, Shaker et al. utilized AI trained in texture-analysis (12). This model assessed the heterogeneity of CT images acquired from dogs with liver masses and found that decreased uniformity was associated with malignancy (12). Moving forward, this study has important implications as it serves as a potential non-invasive technique to assess for malignancy (12).

While many different studies have been completed on the use of AI to assist in diagnostic imaging assessment, clinical implementation has also been rampant. One tool already in place in many small animal clinics throughout North America is SignalPet, an AI model that provides real-time interpretation of any radiograph submitted through their platform.

Signal Pet (13)

SignalPet does not provide any hardware, rather clinicians can submit any radiograph they take at their practice to SignalPet's online portal for AI assessment. The cost of this assessment varies as the company offers a sliding scale model per study submitted. For example, if you run 20 studies, each assessment costs $14.50. However, if you run 31-35 studies, then the cost is reduced to $13.50/study. This will also change if you have any corporate discount. SignalPet has panels to detect pathologies of the thorax, abdomen, MSK, and oral radiographs which are run on every image it is sent. The program will circle any abnormalities found, tell you what the abnormality is, and provide a list of differentials. Currently, this system is only available for dogs and cats, but the company is looking into exotic applications in the future. Every radiograph submitted is added to the SignalPet database to further train their model, and the results can be viewed on their app or website. One thing that is important to note is that you cannot submit a history to this program, so the results do not consider medical history. This tool has been described as a "backup camera" – it will not park the car for you, but it can help see where you are going. The company hopes that this tool will work as a second set of eyes, especially for those who are working in solo or rural practices. SignalPet also offers a consulting service for $75 USD, where you can submit up to 30 images for a board-certified radiologist to review. There is also a chat function, where you can ask clinicians about any issues or concerns you have with your AI interpretation.

One of the benefits of this technology is that it is instantaneous. You do not have to wait for a radiologist to consult on images that you may be uncertain about; in comparison, Antech's radiologist review provides results within 24 hours. Moreover, as multiple panels are run on each image, the technology has the capacity to look for abnormalities outside any presenting complaints that a physician may narrow in on. While potentially flagging non-urgent issues, this tool may increase revenue if other disease processes are found. Finally, SignalPet also provides a digital client report that has the patient's radiographs attached. This allows clients to see what they are paying for, as the cost of the service can easily be added to your clinic's base cost for radiographs.

Digital pathology

Digital pathology is "a subfield of pathology focus[ed] on generating, managing, and interpreting pathology information from digitized glass slides" (14). Rather than analyzing the glass slides directly, images of the slides are captured and uploaded for specialists to review. This use of technology is not new; In fact, IDEXX and Antech have had entirely digital workflows for their anatomic pathology services since 2014 (14). This allows pathologists to review preparations regardless of their physical location and instantaneous sharing of samples (14). What has changed in the past few years is the possibility of having in-house scanners with machine learning software. The Vet Scan Imagyst, produced by Zoetis, has AI assessments available for blood, fecal, urinary, and dermatological samples. This kind of AI tool has the potential benefit of speeding up diagnoses as well as freeing up doctor and technician time. If a machine learning model can assess your slide for you, then you have one more set of hands and eyes that can see a patient.

Vet Scan Imagyst

The Zoetis Vet Scan Imagyst has two main parts: the scanner and the software. The scanner is an in-house analyzer, while the software can be downloaded to your clinic servers (15). For any sample, you would prepare the slide as indicated for the appropriate pathological assessment and then load it into the scanner. From there, you would log into your account and select your patient information; this is possible as the software works with select PIMS systems (15). Once the slide is in view on your screen, you can preview the area of interest and a report will be generated for you. Zoetis currently has a "no contract" kit where they own and maintain the scanner for the first year and the clinic pays for the tests that are run. It is expected to meet a target of $700 a month in tests with this model, which equates to ~40 tests. Currently, it costs $17.00 to run a blood smear or dermatology slide, and $19.00 for a fecal analysis (16). Zoetis will also reimburse the cost of any tests where the results had an error.

AI Blood Smear Analysis

The AI blood smear analysis is completed in a few minutes and includes abnormal white blood cell count, low platelet count, platelet clumping, and changes associated with anemia (17). You can scan the slide yourself and flag the results if something seems amiss. This will tell Zoetis to look at your image and use it for further algorithm training. A sharable PDF is generated, and you can select how many photos are included. This report is available to anyone with a Vet Scan Imagyst to keep data open-access. At any point in this process, you can request an "expert review". Zoetis has stated that the accuracy in diagnosis is comparable to board-certified clinical pathologists (17).

AI Fecal Analysis

The AI fecal analysis is completed in 3.5 minutes after a 2-minute centrifugal time and is aimed at limiting the number of pests missed due to poor staff training and inconsistencies in sample preparation (18). It can detect parasitic ova, cysts, and oocysts (18). Table 1 summarizes the sensitivity and specificity of this analyzer to detect common parasites compared to board-certified pathologists (18). It should be noted that the data Zoetis uses is not open access, and the data from both dogs and cats have been combined (with the exception of Whipworms which is from canine samples only).

Table 1: Sensitivity and Specificity of Zoetis AI fecal analysis as compared to board-certified pathologists

Parasite	Sensitivity	Specificity
Hookworms (Ancylostoma spp)	90.7 (86.4-94%)	97.5% (95.8-98.6)
Roundworms (toxocara spp.)	95.5% (91.7-97.8)	98.3% (96.8-99.2)
Whipworms (Trichuris vulpis)	93.1% (87.8-96.5)	99.7% (98.7-100%)
Tapeworms (Taeniidae)	100% (78.3-100%)	97.8% (93.1-99.5)
Coccidia (Cystoisospora spp.)	94.9% (90.6-97.6)	96% (94.1-97.4)
Giardia	92.1% (83.5-96.9)	98.8% (94.6-99.9)

AI Urinary

The AI Urine sediment analysis aims to inform diagnoses and allow for rapid treatment decisions (19). The system scans ~1000 fields of view and can identify red blood cells, white blood cells, epithelial cells, struvite and calcium oxalate dihydrate crystals, hyaline and non-hyaline casts and cocci and rod bacteria (19). Again, Zoetis claims these results are highly accurate based on internal studies.

AI dermatology

The AI dermatological assessment can detect yeast, inflammatory cells, and differentiate between cocci and rods (20). As with the other Zoetis AI technologies, you can add on a pathologist review or consult with remote specialists if needed. Zoetis claims that the technology is accurate when compared with clinical pathologists, based on non-public data from 143 dogs and 75 cat samples (skin impression smears, ear swabs, and skin swabs) (20).

New advances

With increasing demand and capacity, a variety of AI tools are being developed in the human and veterinary medical fields. Having more data gives more opportunities, and they are being used in many different aspects of medicine (8). Some of the technologies that may soon be implemented in clinics include AI-trained grimace scales, dictation with organization, and generative AI.

Grimace scales

While grimace scales are not new to veterinary medicine, the use of neural networks trained to assess photo data is. Pain Face is a neural network and cloud-based software that can score pain in mice with the hopes that it will reduce subjective scoring (21). Standard users can upload their own videos to the program, which can identify specific 'action units' and generate a grimace score (21). The technology utilized structured machine learning to train the algorithm, including >475,000 images from 270 videos that were annotated by humans (21). Currently, it can score up to 300 images per minute, significantly more than a human observer. While developed to work on mice, the goal is to scale the model and adapt it to other species (21). This type of technology may be used in research to determine when pain medication should be administered and how effective drugs are in comparison to each other. These results may limit the amount of time animals spend in pain post-surgery.

Dictation with organization

The National Academy of Medicine in the United States found that ~half of a clinician's day was spent doing burnout-inducing administrative tasks (22). To try and combat the effects of this work, technologies have been developed that utilize natural language processing. These machine-learning models can "decipher and attribute meaning to text and spoken word" (2). Van Buchem et al. completed a scoping review in 2021 that summarised digital scribe tools available on the market and published results on their technical and clinical validity. The goal of these tools was to record a conversation, transcribe it, and use a model that can extract relevant information (22). This is an incredibly difficult task in veterinary medicine, as medical records are not written with the same standardized disease codes that human physicians use (23). The value of a tool that can summarize information is in its ability to do so correctly, however, this is an issue even with human scribes. The studies included in the review drew attention to how symptoms, medications, or properties were often difficult to interpret by human annotators. Thus, how we hope to use these types of tools will be highly reliant on how well they are trained, and what we can consider a gold standard. Van Buchem et al. found that the models which were not trained on clinical conversations had word error rates (WER) as high as 65%. In contrast, the systems trained on big data sets of clinical conversations had WERs as low as 18%. Moving forward, we must ensure we have the data available to create such tools.

One veterinary model that is currently in development is VetTag. VetTag is a neural network trained to predict Systemized Nomenclature of Medicine – Clinical Terms (SNOMED-

CT) codes based on free-text notes (23). The model was trained with labelled and unlabelled veterinary notes; the use of unsupervised learning was used to help the algorithm predict what the next word may be while the supervised data was used to predict the diagnostic codes (23). This model is still in its infancy and is biased by disease prevalence, but may be utilized in the future to limit the amount of time clinicians spend completing medical records.

Generative AI

Generative AI can create data based on text prompts; currently, the most well-known form of this technology is ChatGBT. ChatGBT uses neural networks that have been trained on large text bodies to comprehend language patterns, making it a Large Language Model (LLM) (24). This type of technology can be used to easily access trends without requiring humans with specific statistical training. For example, you could ask the LLM about all of the information on clinical symptoms when presented with an odd case, or upload data from your clinic to instantly assess disease rates, understand what types of appointments you are seeing the most, or how much money various services are generating. This will dramatically decrease the amount of time data assessment takes but should not be done without considering how the model is trained. How information inputted in the text prompt is protected is still being discussed, specifically as it pertains to medical safety data (24). Moreover, there “is currently no safeguard in place to effectively monitor the quality of the learning experience and to direct the future ChatGBT output towards a verifiable and accurate level of medical information” (24). That being said, as ChatGBT develops, fewer “hallucinations” are produced; hallucinations are non-existent papers that the tool has been known to present (2). The tool can make logical, factual, and statistical mistakes, all of which are worked on with every new release. This type of technology will allow clinicians to summarize disease states while providing instantaneous statistical analysis.

Concerns

The use of AI in small animal veterinary medicine is increasing, and many powerful tools aim to make our lives as clinicians easier. From help assessing radiographs to analyzing data trends, AI is going to be a part of clinical practice for the foreseeable future. Though helpful, there are concerns with any new tool developed, and machine learning is no exception. We must question how models are trained, who owns the data that is presented, who monitors the implemented algorithms, and how it will impact practicing veterinarians if things go wrong.

Model Training

As previously mentioned, AI technologies are being implemented faster than scientific review can process. With all of these available products, we must question if the models have been trained with appropriate types and amounts of data. This is important because it is their training that determines how well they can recognize patterns, and the more complex the pattern, the more data is required (3). This is especially hard in veterinary medicine where getting large data sets is a continuous problem. Moreover, the quality of the data that is used to train these models has the potential to influence outcomes; “Data sets that are incomplete, inaccurate, poorly labelled, unrepresentative of populations of interest or of small volume will generate erroneous models" (3). For example, a low prevalence of disease within a data set can cause the algorithms to artificially report "no abnormalities" when presented with that disease state in the future (4). On the other hand, Boissady et al., who completed a study on AI evaluation of thoracic radiographs, used images from a referral institution to train their neural network. The authors noted that this may cause errors when the model is presented with lower-quality images, something that would be more common in a general practice (25). It is also important to note that with supervised learning, the labelling that represents the "ground truth" or gold standard must be in uniform language and completed by veterinary specialists. The saying "garbage in, garbage out" succinctly summarizes all of the concerns with AI model training. We require high-quality data to train models, but it must not be so exclusive that practitioners hoping to utilize the technology have to undergo additional training to submit samples/images.

To combat these challenges, we must consider machine learning developmental stages: understanding the context of a problem, creating an AI problem, planning development, gathering and cleaning the data, exploratory data analysis, engineering, model building, and deployment (26). The fourth phase, gathering and cleaning the data, is the most prone to human bias as we select which data will be used to train models. By lengthening this period, we can evaluate the data for inherent bias, and ask ourselves why some pieces are included while others are not (26). Additionally, we must strive for open-access data. As presented above, the Zoetis Imagyst was trained and reported to be highly accurate, but the studies are not available for public review. This is a major concern moving forward as we must hold companies accountable for the tools they are developing, especially as they may play roles in future diagnoses. The term Responsible AI is used to describe the actions taken to ensure that artificial intelligence does not cause harm (26) and leads to a conversation around data stewardship.

Data stewardship

Statistics Canada defines data stewardship as data governance in action, meaning "the exercise of decision-making and authority for data-related matters" (27). It includes data protection and privacy as well as adherence to data quality standards (27). These definitions are important as we must ensure that AI is created and used ethically. We must ask ourselves 'how was the data gathered when the tool was trained? whose data was it?'. The information within veterinary medical records is considered privileged and confidential (28) and yet we have tools such as the VetScan Imagyst and SignalPet that use the data submitted by practitioners to further train their models. In such cases, how do we ensure that patient information is not identifiable? How do we make sure that a cat diagnosed with lymphoma after their veterinarian assessed a cytology sample scanned on an Imagyst machine is not identifiable to the clinic down the road? It is the protection and privacy aspect of data stewardship that falls onto the creators of each neural network used. Scrubbing data of identifiable information is something that must be ensured if we hope to use these technologies in the future. This poses a challenge for technologies aimed at summarizing clinical conversations, as they must have access to private information. When it comes to dictation with organization, the security of the information entered must be ensured by the company who designs the model. Future considerations include how long the raw data (conversation recordings) are held, how they are stored, and by whom.

Beyond supervised learning models, other issues exist for generative AI technologies. Currently, we cannot assess the quality of information that generative AI uses (5). While continuously worked on with each new rendition, this is something that practitioners must keep in mind when using services such as ChatGBT. The data gathered and summarized each time you pose a question is not quality-controlled and thus the adherence to data quality standards is unknown.

Algorithm monitoring

Big data is the new normal, from mortgage rates to road checks, and comes with the responsibility to monitor any machine learning model used to automate processes. These math-powered machines are created by humans with their own biases, regardless of intention (29). This means that models can encode "human prejudice, misunderstanding, and bias into the software systems" (29). This is an important aspect of ethical AI use, as automated systems do not change their thinking process the way humans do; humans will make incorrect decisions, however, we can learn and adapt when presented with new information. We will also make these decisions at a much slower rate than any machine learning model will (29). To prevent AI technology from encoding biases, we must ensure that models are frequently checked for discrepancies in their results (26). This follows model deployment and is already being done by some of the tools previously presented. For example, if the report generated by the Zoetis Imagyst has an error, the practitioner reviewing the results can flag the study. This will send it back to Zoetis for algorithm training, helping to ensure that the same mistakes are not made in the future (17). It is the responsibility of the developers to document how these evaluations are done, however, it must be noted that the selling organization is not responsible for what happens after their tool is used (26).

Liability

In human medicine, artificial intelligence/ machine learning tools that support medical specialties are considered medical devices and are thus subject to the approval of Health Canada (and the FDA in the United States) (30,31). There are currently no protocols for the use of AI in veterinary medicine. Despite this lack of formal regulations, specialists such as the Canadian Association of Radiologists have issued position papers to help ensure these tools are being used properly (28). The Canadian Veterinary Medical Association (CVMA) published their position statement on Artificial Intelligence in Veterinary Medicine in October 2023, which states "Veterinary professionals should refer to existing and new provincial or territorial regulations concerning AI systems and their applications in practice"(32). The Alberta Veterinary Medical Association does not yet have a position statement, however, the Vice President has noted that the "ABVMA cannot really regulate AI and it is up to each practitioner, in their duty as veterinarians, to use their judgement" (33). This is echoed by the CVMA, who state that "veterinary professionals should be aware that lack of structured regulatory oversight could result in liability for veterinary practitioners in the event that errors occur as a result of reliance on AI systems…Veterinarians must use AI in a responsible and ethical manner and that its benefits are balanced with consideration for bias and misdiagnosis, risk, animal welfare and privacy"(32). The company that owns the tools we will be using are unlikely to be liable for misdiagnoses (8) so we must ask ourselves how much we trust these tools. As practitioners, we must understand the sensitivity and specificity of any test we run, and AI-trained models are no exception. We must use our clinical knowledge to evaluate the reliability of any new tool. Moving forward, we must assess open-access data and understand the purpose of any AI technology that we hope to employ. It is bad practice to run any test without a reason, something more easily forgotten when a whole radiology assessment is just one button away.

Discussion

Many new tools on the market are aimed at making a practitioner's life easier, including automated assessment of radiographs, pathological samples, pain scales, medical records, and data analysis. These are wonderful tools with the capacity to increase the time veterinarians can spend with their clients while increasing revenue. This paper is meant as a background to better understand the tools that will be marketed to veterinarians as well as prepare future veterinarians to question the reliability of AI technology. There are a few main takeaways to consider as you prepare yourself for a future of veterinary medicine that utilizes Artificial Intelligence: informed consent, tolerance, the role of veterinarians in the development of these tools, and what we can do to protect ourselves.

As with any treatment or procedure, the tests that we utilize in veterinary medicine require informed consent. When it comes to artificial intelligence, it's our responsibility as to inform animal owners of all of the diagnostic options available and provide a risk-benefit analysis for each tool as well as the associated costs (28). No owner should be subject to any diagnostic tests that they do not feel they understand or have questions about. Importantly, as these tests become mainstream, their demand is going to increase. This further bolsters the need for informed consent, as reading about new technology online does not ensure an owner understands the purpose of the test or what the results may mean.

Kim et al. raise an important point when considering the design and implementation of AI systems: the tolerance for false negatives and false positives will be determined by the consequence of the diagnosis (34). We do not expect these tools to faulter, and yet specialists with a patient's entire medical history still make mistakes from time to time. If the mistake these tools make result in incorrect treatment or misdiagnosis, we may become much less tolerant of the technology itself. Thus, we must use our clinical reasoning to determine when confidence in an AI tool is appropriate while continuously monitoring and assessing diagnosis and treatment. This leads to another discussion on the future of veterinary medicine as it pertains to AI, which is that veterinarians are vital to their development. The systems must be designed with an understanding of the implications for both false positives and false negatives, something only clinicians can truly appreciate (4). We are also needed to ensure that labelled data sets are accurate and to report and incorrect findings. Veterinarians will not be replaced by artificial intelligence technologies, but those who use them will become more valuable than those who do not (2).

Finally, we must ask what we can do to protect ourselves. As our governing bodies are leaving the liability of mistakes on the veterinarian, there is currently no entity to protect veterinarians. There are a few routes we could consider: academic institutions, continuous education (CE), or a new, special division dedicated to AI duty of care. For new veterinarians entering the field, we could suggest that academic institutions implement courses on AI technologies and when they should be used. For vets already working, we might ask that more

CE credits area available on the topic. More forcefully, we could ask our governing bodies to develop a special division and re-consider the position statement to provide coverage when it comes to tools that utilize machine learning. There is no clear-cut answer, but it is important to consider what liability looks like now and in the future as these tools increase in popularity.

REFERENCES

1. Ogilvie T, Kastelic J. Technology is rapidly changing our world, including veterinary medicine. Can Vet J. 2022 Dec;63(12):1177–8.

2. Appleby RB, Basran PS. Artificial intelligence in veterinary medicine. Journal of the American Veterinary Medical Association. 2022 May 1;260(8):819–24.

3. Scott IA. Demystifying machine learning: a primer for physicians. Internal Medicine Journal. 2021;51(9):1388–400.

4. Hennessey E, DiFazio M, Hennessey R, Cassel N. Artificial intelligence in veterinary diagnostic imaging: A literature review. Vet Radiology Ultrasound. 2022 Dec;63(S1):851–70.

5. Morley J, DeVito NJ, Zhang J. Generative AI for medical research. BMJ. 2023 Jul 12;382:p1551.

6. Basran PS, Appleby RB. The unmet potential of artificial intelligence in veterinary medicine. American Journal of Veterinary Research. 2022 May 1;83(5):385–92.

7. Ezanno P, Picault S, Beaunée G, Bailly X, Muñoz F, Duboz R, et al. Research perspectives on animal health in the era of artificial intelligence. Vet Res. 2021 Mar 6;52(1):40.

8. Appleby R, Gowland S. Vet Sessions: AI and Diagnostic Imaging in Vet Med. (Vet Sessions).

9. McEvoy FJ, Amigo JM. Using machine learning to classify image features from canine pelvic radiographs: evaluation of partial least squares discriminant analysis and artificial neural network models. Vet Radiol Ultrasound. 2013;54(2):122–6.

10. Gomes DA, Alves-Pimenta MS, Ginja M, Filipe V. Predicting Canine Hip Dysplasia in X-Ray Images Using Deep Learning. Conference Paper: In: Optimization, Learning Algorithms and Applications. Springer International Publishing; 2021. p. 393–400.

11. Boissady E, De La Comble A, Zhu X, Abbott J, Adrien-Maxence H. Comparison of a Deep Learning Algorithm vs. Humans for Vertebral Heart Scale Measurements in Cats and Dogs Shows a High Degree of Agreement Among Readers. Front Vet Sci. 2021 Dec 9;8:764570.

12. Shaker R, Wilke C, Ober C, Lawrence J. Machine learning model development for quantitative analysis of CT heterogeneity in canine hepatic masses may predict histologic malignancy. Vet Radiol Ultrasound. 2021 Nov;62(6):711–9.

13. Hammond P. SignalPet overview. 2023.

14. Zuraw A, Aeffner F. Whole-slide imaging, tissue image analysis, and artificial intelligence in veterinary pathology: An updated introduction and review. Vet Pathol. 2022 Jan 1;59(1):6–25.

15. Zoetis USA. Zoetis VetScan Imagyst. 2023 [cited 2023 Nov 15]. VETSCAN IMAGYST. Available from: https://www.vetscanimagyst.com/

16. Simeon Aa. Zoetis Vet Scan Imagyst: Call with Sales Representative. 2023.

17. Zoetis USA. Vetscan Imagyst AI Blood Smear [Internet]. 2023 [cited 2024 Jan 19]. Available from: https://www.zoetisdiagnostics.com/point-of-care/vetscan-imagyst/imagyst-ai-blood-smear

18. Zoetis USA. Vetscan Imagyst AI Fecal [Internet]. 2023 [cited 2024 Jan 19]. Available from: https://www.zoetisdiagnostics.com/point-of-care/vetscan-imagyst/imagyst-ai-fecal

19. Zoetis USA. Vetscan Imagyst AI Urine Sediment [Internet]. [cited 2024 Jan 19]. Available from: https://www.zoetisdiagnostics.com/point-of-care/vetscan-imagyst/imagyst-ai-urine-sediment

20. Zoetis USA. Vetscan Imagyst AI Dermatology [Internet]. 2023 [cited 2024 Jan 19]. Available from: https://www.zoetisdiagnostics.com/point-of-care/vetscan-imagyst/imagyst-ai-derm

21. McCoy ES, Park SK, Patel RP, Ryan DF, Mullen ZJ, Nesbitt JJ, et al. Development and validation of PainFace, a software platform that simplifies and standardizes mouse grimace analyses [Internet]. bioRxiv; 2022 [cited 2023 Nov 15]. p. 2022.08.12.503790. Available from: https://www.biorxiv.org/content/10.1101/2022.08.12.503790v1

22. van Buchem MM, Boosman H, Bauer MP, Kant IMJ, Cammel SA, Steyerberg EW. The digital scribe in clinical practice: a scoping review and research agenda. npj Digit Med. 2021 Mar 26;4(1):1–8.

23. Zhang Y, Nie A, Zehnder A, Page RL, Zou J. VetTag: improving automated veterinary diagnosis coding via large-scale language modeling. NPJ Digit Med. 2019 May 8;2:35.

24. Varghese J, Chapiro J. ChatGPT: The transformative influence of generative AI on science and healthcare. Journal of Hepatology [Internet]. 2023 Aug 5 [cited 2023 Dec 11]; Available from: https://www.sciencedirect.com/science/article/pii/S0168827823050390

25. Boissady E, de La Comble A, Zhu X, Hespel AM. Artificial intelligence evaluating primary thoracic lesions has an overall lower error rate compared to veterinarians or veterinarians in conjunction with the artificial intelligence. Vet Radiol Ultrasound. 2020 Nov;61(6):619–27.

26. Odubela A. Foundations of Responsible AI [Internet]. LinkedIn Learning presented at; 2022 Aug. Available from: https://www.linkedin.com/learning/foundations-of-responsible-ai/understanding-responsible-ai

27. Statistics Canada. Data stewardship: An introduction [Internet]. 2022 Jun 9 [cited 2024 Feb 8]. Available from: https://www.statcan.gc.ca/en/wtc/data-literacy/catalogue/892000062020013

28. Bourque DT, Horney DB. Principles of Veterinary Medical Ethics of the CVMA. 2016 Jul;

29. O'Neil C. Weapons of math destruction : how big data increases inequality and threatens democracy. First edition. Crown,; 2016.

30. Da Silva M, Flood CM, Goldenberg A, Singh D. Regulating the Safety of Health-Related Artificial Intelligence. Healthc Policy. 2022 May;17(4):63–77.

31. Benjamens S, Dhunnoo P, Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. npj Digit Med. 2020 Sep 11;3(1):1–8.

32. Canadian Veterinary Medical Association. Artificial Intelligence in Veterinary Medicine Position Statement [Internet]. 2023 Oct [cited 2024 Feb 8]. Available from: https://www.canadianveterinarians.net/policy-and-outreach/position-statements/statements/artificial-intelligence-in-veterinary-medicine/

33. Chalhoub S. Artificial intelligence and the ABVMA. Personal Communication. 2023.

34. 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 May;63(3):292–7.