Disease Prediction and Precision Veterinary Medicine

Applications, Opportunities, and Limitations of Artificial Intelligence in Small Animal Practice

Audrey Ruple, DVM, MS, PhD, MRCVS, FNAP[a,*],
Stuart W.J. Reid, BVMS, PhD, DVM, FRSE, FRCVS, CBE[b]

KEYWORDS

• Artificial intelligence • Machine learning • Clinical prediction models
• Precision veterinary medicine • Veterinary big data

KEY POINTS

- Disease prediction models estimate future risk rather than diagnose current disease, enabling earlier intervention and more proactive veterinary care.
- Predictive performance depends heavily on the underlying data sources, including electronic medical records, insurance claims, owner-generated data, wearable devices, and environmental datasets.
- Artificial intelligence–based models can detect complex patterns in large veterinary datasets, but their predictions must be interpreted within the clinical context.
- Predictive tools should complement, not replace, clinical expertise, and their outputs should be integrated with physical examination findings and diagnostic testing.
- Responsible implementation requires attention to model validation, bias, and transparency, particularly when tools are applied across different populations or practice settings.

INTRODUCTION

Artificial intelligence (AI) and machine learning (ML) are increasingly being integrated into small animal veterinary medicine, with applications ranging from image interpretation and clinical documentation to disease prediction and population-level surveillance. Among these emerging uses, predictive modeling represents one of the most

[a] Department of Population Sciences, Virginia-Maryland College of Veterinary Medicine, 205 Duck Pond Drive, Blacksburg, VA 24060, USA; [b] Royal Veterinary College, Hawkshead Lane, North Mymms, Hatfield, Hertfordshire, AL9 7TA, England
* Corresponding author.
E-mail address: aruple@vt.edu

Vet Clin Small Anim 56 (2026) 1165–1180
https://doi.org/10.1016/j.cvsm.2026.04.003 **vetsmall.theclinics.com**
0195-5616/26/© 2026 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

Abbreviations	
AI	artificial intelligence
LLM	large language models
ML	machine learning
MMVD	myxomatous mitral valve disease
NLP	natural language processing
VEMRs	veterinary electronic medical records

promising, and potentially disruptive, developments for clinical practice. Rather than reacting to disease once it becomes clinically apparent, predictive tools aim to estimate future risk, enabling earlier intervention, individualized prevention strategies, and more efficient allocation of clinical resources.

Disease prediction and precision veterinary medicine sit at the intersection of data science and clinical care. These approaches leverage large, heterogeneous datasets to identify patterns that are often imperceptible to human clinicians alone. While similar methods have been adopted widely in human medicine, veterinary applications are only beginning to mature. Importantly, predictive models do not replace clinical judgment; instead, they provide probabilistic information that can support decision-making when used thoughtfully and in appropriate clinical contexts. Both generative and nongenerative AI have found utility in the handling analysis and interpretation of medical records and clinical data; natural language processing (NLP) and large language models (LLM) offer different advantages depending upon the required outcome, viz, interpretation of human language versus precision clinical decision support.

What Disease Prediction and Precision Veterinary Medicine Mean in Practice

Disease prediction refers to the use of statistical or ML models to estimate the likelihood that an animal will develop a specific condition within a defined time horizon.[1] These models draw on combinations of clinical records, demographic information, diagnostic histories, environmental exposures, and other data streams to generate individualized or population-level risk estimates. Precision veterinary medicine builds on this concept by tailoring preventive care, diagnostics, and treatment strategies to the individual patient rather than relying solely on population averages.[2]

Distinguishing Risk Prediction from Diagnosis

Importantly, prediction does not equate to diagnosis of whether a disease is currently present. Predictive models estimate probabilities and future outcomes[3]; they do not confirm disease. A high-risk score signals the need for closer monitoring or further evaluation, not definitive pathology. Similarly, low predicted risk does not guarantee absence of disease. A predictive model may estimate, for example, the likelihood that a dog will develop Cushing's syndrome or an apparently healthy adult cat will develop chronic kidney disease over the next year.[4,5] Prediction shifts veterinary medicine toward a proactive paradigm, emphasizing prevention and early detection rather than reactive management of diseases and conditions alone.[6]

Individual-Level versus Population-Level Prediction

Predictive modeling operates at multiple scales. Individual-level prediction estimates risk for a specific patient and can guide personalized care plans.[7] Population-level prediction aggregates risk across groups or geographic regions, supporting surveillance, resource planning, and public health preparedness.[8] Both approaches are

valuable but serve different purposes. Individual predictions inform clinical conversations with pet owners about their specific animal, while population models can identify emerging disease trends, geographic areas at high-risk of disease incursions, or seasonal patterns.[9]

Why These Tools Are Emerging Now

Several converging factors have accelerated the adoption of predictive analytics in veterinary medicine. First, the digitization of veterinary records and the growth of large-scale datasets including electronic health records, pet insurance claims data, wearable devices, litter box monitors, and open-source datasets containing owner-reported information have created unprecedented opportunities for data-driven insights.[10] Second, advances in computational power and ML algorithms have made it feasible to analyze complex, high-dimensional data at scale.[11] Finally, commercialization of AI-enabled products has brought predictive tools directly into clinical workflows.[12]

At the same time, growing recognition of the interconnectedness of animal, human, and environmental health has expanded interest in predictive models that operate beyond the individual patient. Dogs share many environmental exposures with people, positioning companion animals as potential sentinels for broader health risks and reinforcing the relevance of predictive approaches within a One Health framework.[13]

Together, these developments have laid the groundwork for disease prediction and precision veterinary medicine to move from theoretic concepts to practical clinical tools. However, their successful integration into practice depends on careful attention to data quality, model limitations, and responsible clinical use.

DATA FOUNDATIONS FOR DISEASE PREDICTION

Successful disease prediction depends fundamentally on the quality, scope, and structure of the underlying data. In veterinary medicine, predictive models increasingly draw on diverse sources summarized with strengths and weaknesses in **Table 1**. Each data stream contributes unique strengths while also introducing distinct limitations that clinicians should understand when interpreting predictive outputs.

Veterinary Electronic Medical Records

Veterinary electronic medical records (VEMRs) provide the clinical backbone for most veterinary predictive models. Large-scale initiatives such as VetCompass in the United Kingdom have demonstrated the value of aggregating primary-care VEMRs for population-level research and disease surveillance, providing insights into breed-associated disease risk, morbidity patterns, and clinical outcomes across large companion animal populations.[14–16]

However, VEMRs also present challenges. Clinical notes are often unstructured, diagnostic certainty varies across practices, and behavioral or environmental variables are inconsistently captured.[17] Both traditional (nongenerative) AI approaches and newer generative AI systems are increasingly being incorporated into clinical decision-support tools.[11,18] The NLP methods are commonly used to extract structured information from unstructured clinical text within medical records.[17,19] In contrast, LLMs, a recent class of NLP systems trained on large text corpora, can further support clinical workflows by summarizing records, synthesizing information, and assisting with clinical documentation.[20] While these tools primarily support information management, they may complement predictive modeling systems that estimate disease risk using structured clinical data.[21] Referral-based datasets may

Table 1
Key features, strengths, limitations, biases, and application uses of data sources available for use in disease prediction in veterinary medicine

Data Source	Key Features	Strengths for Prediction	Limitations and Biases	Representative Applications
VEMRs	Clinical records from primary and referral care, including diagnoses, laboratory results, imaging, treatments, and demographics	Rich clinical detail; longitudinal disease trajectories; large-scale aggregation through networks (eg, VetCompass)	Variable data quality; unstructured clinical notes; inconsistent diagnostic coding; referral bias; influenced by care-seeking behavior	Disease risk modeling; epidemiologic surveillance; prognostic modeling (eg, Cushing's disease and chronic kidney disease)
Pet Insurance Claims Data	Administrative datasets capturing billed diagnoses and procedures across large insured populations	Large sample sizes; geographic diversity; longitudinal follow-up; more complete diagnostic workups due to reduced financial constraints	Not representative of all pets (bias toward insured populations, higher socioeconomic status, single-breed dogs); limited diagnostic granularity; coding constraints	Population-level disease trends; morbidity and mortality studies; One Health applications linking animal and human disease patterns
Owner-Generated Data	Survey-based or app-based information reported by owners, including health history, diet, behavior, and household environment	Captures behavioral and environmental exposures not recorded clinically; scalable across large cohorts; validated concordance with clinical data for many conditions	Recall bias; misclassification; underreporting of certain conditions; variability in owner interpretation	Longitudinal cohort studies (eg, Dog Aging Project); risk factor identification; integration with clinical, and environmental datasets
Wearable and Sensor Data	Device-generated continuous data (eg, activity, rest, posture) collected in real-world settings using accelerometers and related sensors	Continuous, objective monitoring; detection of subtle behavioral or mobility changes; potential early indicators of disease progression or recovery	Variable validation across devices and contexts; performance depends on environment (eg, on- vs off-leash); data processing and clinical integration challenges; reliance on owner adherence	Monitoring of osteoarthritis, obesity, and postsurgical recovery; assessment of activity trends and quality of life; early detection of functional decline

Environmental and Geospatial Data	Data linked to location (eg., climate, air quality, land use, socioeconomic indicators) via geocoding and public datasets	Captures external exposures relevant to infectious, environmental, and chronic diseases; supports One Health analyses; enables spatial and temporal risk modeling	Spatial and temporal misalignment; variable resolution; potential confounding; requires linkage to clinical or cohort data	Prediction of vector-borne and environmentally mediated diseases; exposure assessment in longitudinal cohorts; regional disease forecasting
Integrated Multimodal Data	Combined datasets linking clinical, owner-reported, wearable, and environmental data streams	More comprehensive representation of disease risk; improved predictive performance for complex, multifactorial conditions; supports precision medicine approaches	Data harmonization challenges; missing data; computational complexity; requires robust infrastructure and validation	Precision risk stratification; longitudinal aging studies; integrated clinical decision support; population health surveillance

overrepresent severe or complex disease, while primary care records may lack standardized diagnostic confirmation.[22] These factors introduce heterogeneity that can influence model performance and generalizability.

Pet Insurance Claims Data

Pet insurance datasets represent an increasingly valuable complement to VEMRs. Claims databases typically include large, geographically diverse populations and offer longitudinal follow-up at the individual animal level.[9,10,23] A key advantage is that insured pets are more likely to receive recommended diagnostics and treatments and less likely to experience economic euthanasia, resulting in more complete clinical documentation.[23]

Despite these strengths, insurance datasets are not fully representative of the general pet population. Insured dogs tend to be younger and more often of a single breed as opposed to mixed breed, and coverage varies substantially by region and socioeconomic factors.[17] Additionally, insurance claims systems may limit the number of diagnostic codes per visit, reducing clinical granularity.

Owner-Generated Data and Wearable Technologies

Owner-reported information and wearable technologies provide access to behavioral, lifestyle, and environmental exposures that are largely absent from traditional clinical records, including activity patterns, diet, household characteristics, and perceived health status.[24] Historically, owner-generated data have been viewed with caution due to concerns about recall bias and misclassification. However, recent validation studies from large companion animal cohorts demonstrates substantial concordance between owner-reported data and information obtained from VEMRs or laboratory testing.[25,26]

Wearable devices further extend owner-generated data by capturing continuous, real-world data on animal behavior and physiology.[27] These technologies encompass a wide array of sensor types, including accelerometers, gyroscopes, and other embedded instruments, that can quantify activity patterns, rest and sleep states, posture, and, in some cases, vital sign proxies.[27] Accelerometer-based activity monitors have been validated in domestic dogs to track physical activity with moderate to strong correlation against direct behavioral observations, illustrating their potential for detecting changes in movement that may precede clinical concerns such as osteoarthritis, obesity, and postsurgical recovery, as well as for assessing overall welfare and behavior outside structured clinical assessments.[28–32] Despite this promise, clinical integration of wearable technology remains uneven and device performance can vary depending on use case and environment.[33,34] In addition, continuous streams of data must be processed, visualized, and interpreted in ways that align with veterinary decision-making; this often involves cloud-based analytical platforms and, in some cases, integration with practice management systems.[27,34]

Environmental and Geospatial Data: a One Health Perspective

Environmental exposures play a critical role in many disease processes, yet are rarely captured directly in clinical encounters or in insurance claims. Advances in geocoding and linkage to publicly available datasets now allow VEMRs and other datasets to be enriched with neighborhood-level measures of air quality, climate, walkability, and socioeconomic context.[13,35] Incorporating environmental variables into predictive models can improve performance for conditions influenced by climate, pollution, or vector ecology, such as tick-borne and fungal diseases.[10] At the same time, environmental data introduce additional complexity related to spatial resolution, temporal

alignment, and potential confounding, underscoring the importance of careful model design and validation.

Integrating Data Streams for Clinical Prediction

No single data source is sufficient for robust disease prediction. Precision veterinary medicine emerges from the thoughtful integration of these complementary streams. For clinicians, the critical takeaway is that predictive outputs reflect the strengths and limitations of their underlying data. Understanding where information originates, and what may be missing from that data source, helps contextualize risk estimates and supports responsible use of AI-enabled tools in practice.

PREDICTIVE MODELING APPROACHES USED IN VETERINARY MEDICINE

Predictive modeling encompasses a suite of analytical approaches that learn patterns in historical data to estimate future outcomes. In veterinary medicine, these approaches range from traditional statistical models to modern ML and AI based methods. The purpose of this section is to provide clinicians with a conceptual overview of the major modeling paradigms, highlight how they have been applied in veterinary contexts, and discuss considerations for their interpretation and implementation.

Traditional Statistical Models versus Machine Learning

Traditional statistical models such as linear regression and logistic regression have long been used in veterinary epidemiology to estimate associations between predictors and health outcomes.[36] These models rely on predefined functional forms and assumptions about linearity, independence, and distribution of predictor variables.

ML methods extend this framework by focusing more explicitly on predictive performance than on inference about individual predictors. ML approaches such as decision trees, random forests, support vector machines, and gradient boosting iteratively learn patterns in complex, high-dimensional data without requiring explicit specification of relationships between predictors and outcomes.[37] Compared with traditional models, ML methods often tolerate nonlinear relationships and interactions more naturally and can handle large numbers of predictors more effectively.[38]

Veterinary ML applications have demonstrated this capability in specific clinical contexts. For instance, machine-learning models trained on primary-care veterinary clinical records have been used to predict the future diagnosis of hyperadrenocorticism (Cushing's syndrome) in dogs based on demographic, clinical, and laboratory variables, achieving strong discriminatory performance.[4] In veterinary diagnostic imaging, deep learning models have demonstrated high accuracy for detecting conditions such as canine hip dysplasia from radiographs, highlighting the potential of AI-assisted diagnostic tools in clinical practice.[39,40] Supervised ML classification models have been used to estimate risk of heart failure in dogs diagnosed with myxomatous mitral valve disease (MMVD), with demographic data, radiographic measurements, echocardiographic values, and laboratory results contained within VEMRs shown to have excellent performance in estimating the prognosis of patients.[41] In another example, clinical features and laboratory measured biomarkers were used to predict clinical outcome and time to recovery in dogs infected with parvovirus.[42]

Integration with Veterinary Practice

Despite methodological advancements, the integration of predictive models into everyday veterinary practice is still evolving. Barriers include variability in data quality

across practices, lack of standardized coding of clinical records, and disparities in algorithm performance across patient populations.[43] As veterinary datasets grow in size and diversity there is increasing potential for models that support personalized risk stratification and optimized preventive care. Continued collaboration between veterinary clinicians and data scientists is essential to develop, validate, and implement models that improve animal health while maintaining clinical relevance and transparency.

CLINICAL USE CASES AND EXAMPLES

Predictive modeling in veterinary medicine is most compelling when it translates into clinically meaningful action. While algorithm development is an essential first step, the ultimate value of disease prediction lies in its application to patient care, surveillance, and preventive strategy. The following sections highlight key domains in which predictive approaches are already being explored in companion animal medicine.

Infectious Disease Prediction and Early Warning

Companion animals can serve as sentinels for infectious and zoonotic diseases that affect both animal and human populations.[9] Leveraging large-scale veterinary datasets, predictive models have been developed to estimate risk of vector-borne and environmentally mediated diseases across geographic regions.[37] Predictive infectious disease modeling is particularly well suited to conditions influenced by environmental variables such as climate, vector ecology, and population density. Integration of geospatial and environmental data layers with clinical records may further enhance regional forecasting of tick-borne and fungal diseases.[35,44]

At the clinical level, predictive modeling has also been applied to individual infectious diseases. In dogs with parvoviral enteritis, ML models trained on clinical and laboratory variables have been used to predict survival and recovery time, demonstrating the potential to inform triage and client communication.[42] Such models may assist clinicians in identifying high-risk patients early in hospitalization, although external validation and prospective testing remain essential before widespread adoption.

Dogs as Sentinels: a One Health Perspective

The concept of dogs as sentinels for environmental exposure and disease risk is increasingly supported by large longitudinal cohorts. Companion dogs share household environments with humans, experience similar air and water exposures, and develop many comparable chronic and infectious diseases.[35,41,42] Within the Dog Aging Project, integration of owner-reported environmental data, VEMRs, and geospatial exposure variables has enabled the construction of comprehensive environmental profiles linked to canine health outcomes.[35,45,46] This infrastructure supports predictive modeling that operates not only at the individual patient level but also at neighborhood and regional scales. Incorporating companion animal health trends into public health modeling frameworks may improve early detection of emerging disease clusters and support coordinated response strategies across veterinary and human health systems.

Chronic Disease Risk Prediction

Chronic, progressive diseases represent a substantial portion of morbidity in companion animals. Predictive modeling and ML algorithms have been applied in several chronic disease domains, such as dogs with MMVD, to estimate future risk or

progression.[41] Similarly, activity monitor–derived metrics have been proposed as objective markers of mobility and disease progression in osteoarthritis, offering potential integration with predictive frameworks for chronic musculoskeletal disease.[29] Predictive use of wearable-derived longitudinal data could, in theory, allow clinicians to identify deterioration before overt clinical decline is observed, though real-world clinical implementation remains limited. Predictive modeling may also assist in stratifying risk for metabolic or endocrine disorders, particularly when longitudinal weight, activity, and laboratory data are available. While such models require further external validation, they illustrate how predictive analytics may support earlier intervention or more intensive monitoring in high-risk patients.

Preventive Care Stratification and Individualized Screening

A central promise of precision veterinary medicine is tailoring preventive care recommendations to individual risk profiles rather than applying uniform schedules across all patients. Predictive models may help identify animals that would benefit from earlier screening, more frequent monitoring, or targeted preventive interventions. For example, population-level analyses of insurance and VEMR datasets have identified breed- and age-associated patterns of disease incidence, which could inform more individualized screening strategies.[47–51] When combined with validated breed ancestry data and environmental exposure profiles, such risk estimates may refine preventive recommendations in ways that move beyond broad breed stereotypes. However, implementing risk-stratified screening requires careful evaluation of predictive performance, cost-effectiveness, and potential unintended consequences such as overtesting or increased client anxiety.[52]

LIMITATIONS, BIAS, AND CLINICAL RISK

While predictive modeling holds substantial promise for advancing precision veterinary medicine, its clinical utility depends on careful consideration of limitations, sources of bias, potential risks, and ethical and professional considerations associated with implementation. Predictive models influence clinical decisions, client perceptions of risk, and allocation of care resources; therefore, their development and deployment carry responsibilities that extend beyond algorithmic performance. Understanding these challenges is essential for interpreting model outputs responsibly and avoiding unintended consequences in patient care.

Dataset Shift and Generalizability

Predictive models are inherently dependent on the data used for their development. When models are applied to populations that differ from their training data by geography, breed composition, socioeconomic context, or care-seeking behavior, performance may degrade, a phenomenon commonly referred to as dataset shift.[53] In veterinary medicine, this risk is amplified by variability in practice management systems, diagnostic coding conventions, and regional disease prevalence.

Models trained on referral hospital populations may not generalize well to primary care settings, and algorithms derived from insured or highly engaged owner populations may perform differently when applied to broader community cohorts. For example, pet insurance datasets and large longitudinal research cohorts often overrepresent younger dogs and owners with higher socioeconomic status, which may influence both disease detection and treatment patterns.[24] These factors underscore the importance of external validation across diverse clinical settings before predictive tools are deployed widely.

Confounding versus Prediction

A critical conceptual distinction in predictive modeling is the difference between identifying causal relationships and optimizing prediction. ML algorithms prioritize accuracy in forecasting outcomes but do not inherently distinguish between causal drivers and correlated features.[37] As a result, variables that improve predictive performance may not represent modifiable risk factors.

This distinction has practical implications. A model may accurately predict disease risk based on demographic or environmental proxies without revealing actionable intervention points. Clinicians should therefore avoid interpreting model-derived predictors as evidence of causality unless supported by independent epidemiologic or experimental data.[37]

Overreliance on Probabilistic Outputs and Communicating Risk

Predictive models generate probabilities rather than definitive diagnoses.[52] There is a risk that clinicians or clients may overinterpret these outputs as certainties, particularly when presented as single numerical risk scores. Communicating these probabilities effectively to pet owners presents practical and ethical challenges.[54] Numerical risk scores may be misunderstood as certainties, potentially increasing anxiety or prompting unnecessary interventions.[55] Evidence from human health care demonstrates that framing and presentation of risk information strongly influence patient perception and decision-making.[55] Veterinary clinicians similarly need strategies to contextualize predictive outputs, emphasizing uncertainty and avoiding deterministic language. Shared decision-making frameworks that incorporate owner values and goals are particularly important when discussing predictive information related to chronic disease progression or future health risks.[56]

Equity, Representation, and Stewardship in Veterinary Data

Ethical implementation also requires attention to data governance, privacy, and representation.[57] Veterinary datasets increasingly incorporate owner-generated information, geospatial exposures, and longitudinal health records, raising questions about consent, data security, and appropriate secondary use.[43]

Bias can arise when certain populations are underrepresented or misrepresented in training datasets.[58] In companion animal medicine, this may include uninsured pets, animals from rural or underserved communities, or owners with limited access to veterinary care.[13,59] Such disparities can influence model accuracy and may exacerbate existing inequities if predictive tools preferentially benefit well-represented groups. Ongoing evaluation of model performance across diverse populations is therefore critical to avoid exacerbating existing inequities.[40]

Clinical Risk and Model Transparency

A central ethical and clinical challenge is the interpretability of predictive models. Complex ML systems particularly those involving ensemble methods and deep learning architectures can function as *black boxes*, producing risk estimates without clear explanations of how individual inputs contributed to the output.[60] Poorly calibrated models may lead to unnecessary diagnostic testing, overtreatment, or increased client anxiety, while underperforming models may fail to identify animals at genuine risk.[52] Transparency regarding model development, validation metrics, and intended use is therefore essential.[60]

Professional organizations in human medicine increasingly emphasize the importance of explainability and clinical oversight in AI deployment, enabling clinicians to

contextualize predictions and identify potential errors or biases in model outputs.[53,60–62] Although comparable veterinary guidelines are still emerging, similar principles apply: clinicians must understand and be able to communicate how predictions are generated, what data are used, and where uncertainty exists.[62] Models that provide interpretable features or clinically intuitive risk factors may therefore be more readily adopted in practice, even if their raw predictive performance is modestly lower.[53,60]

Toward Responsible Implementation

Mitigating these risks requires rigorous model evaluation, including external validation, monitoring for performance drift over time, and clear communication of uncertainty.[52,53] Integration into clinical workflows should prioritize interpretability and alignment with veterinary decision-making processes.[62] Ongoing feedback from real-world outcomes can support iterative model refinement within a learning health system framework.[63]

Ultimately, predictive tools are most effective when paired with clinical expertise, shared decision-making with pet owners, and an appreciation of the broader social and environmental context in which veterinary care occurs. Responsible implementation depends not only on technical innovation but also on ethical stewardship and thoughtful clinical integration.

Responsibility for Clinical Decisions

Predictive tools are designed to support, not replace, clinical judgment. Ultimately, veterinarians remain responsible for diagnostic and therapeutic decisions.[61] Overreliance on algorithmic outputs risks diminishing professional autonomy and may contribute to automation bias, whereby users defer excessively to model recommendations even when they conflict with clinical intuition or contextual information.[64] In human medicine, concerns about automation bias where users defer excessively to algorithmic recommendations, have been well documented.[64] Similar risks exist in veterinary contexts, especially as AI-enabled tools become embedded in practice management platforms.

Clear delineation of responsibility is especially important when predictive outputs influence high-stakes decisions, such as aggressive diagnostics, long-term treatment commitments, or end-of-life discussions. Veterinarians must retain authority over care decisions and integrate model-derived insights with physical examination findings, client preferences, and the broader clinical picture.

FUTURE DIRECTIONS IN PRECISION VETERINARY MEDICINE

The future of disease prediction in veterinary medicine lies in increasingly integrated, longitudinal, and adaptive systems that combine multiple data streams to support individualized care and population health.

Integration Across Data Streams

Large longitudinal initiatives such as the Dog Aging Project demonstrate the feasibility of linking clinical, behavioral, and environmental variables at scale, enabling more comprehensive modeling of disease risk and aging trajectories.[24] However, technical challenges remain, including harmonization of data formats, management of missing information, and alignment of temporal scales across disparate sources.[43]

Learning Health Systems

A longer-term vision for precision veterinary medicine involves development of learning health systems, in which clinical data are continuously analyzed to refine

predictive models and inform practice in near real time.[63] Human health care systems have begun to explore this approach, leveraging EMR data to support dynamic clinical decision support.[18]

Regulatory and Validation Needs

As predictive tools become more embedded in clinical workflows, standards for validation, reporting, and oversight will become increasingly important.[53,65] External validation across independent populations should be considered a minimum requirement before clinical deployment.[52,66]

Toward Clinically Meaningful Precision Care

Ultimately, the success of predictive modeling in veterinary medicine will depend on its ability to improve outcomes that matter to animals, owners, and clinicians. This includes not only disease prevention and early detection but also quality of life, client satisfaction, and efficient use of clinical resources. Precision veterinary medicine is unlikely to be realized through algorithms alone. Instead, it will emerge through thoughtful integration of predictive tools with clinical expertise, transparent communication, and ongoing evaluation of impact in real-world practice. By embracing both technological innovation and professional stewardship, veterinary medicine can harness predictive analytics to support more proactive, personalized, and equitable care.

SUMMARY

Disease prediction and precision veterinary medicine are emerging applications of artificial intelligence that aim to identify future disease risk and support more proactive, individualized patient care. By integrating data from veterinary electronic medical records, insurance claims, owner-reported information, wearable devices, and environmental sources, predictive models can estimate risk for a wide range of conditions at both individual and population levels. Although these tools have demonstrated promise in areas such as chronic disease management, infectious disease surveillance, and preventive care stratification, their performance depends heavily on data quality, representativeness, and appropriate validation. Predictive outputs should be interpreted as probabilistic estimates rather than definitive diagnoses and used to complement, not replace, clinical expertise. As veterinary datasets continue to expand and predictive technologies mature, responsible implementation of artificial intelligence has the potential to advance precision veterinary medicine, improve preventive care, and support more personalized and equitable health outcomes for companion animals.

CLINICS CARE POINTS

- Predictive models estimate risk, not diagnosis. A high-risk prediction should prompt closer monitoring or targeted diagnostics but does not confirm disease presence.
- Model outputs reflect the data used to build them. Predictions derived from referral populations, insured pets, or highly engaged owner cohorts may not generalize to all clinical settings.
- Artificial intelligence tools are most effective when used as decision-support aids. Predictive models should complement clinical judgment rather than replace physical examination, diagnostic testing, or clinician experience.
- Data quality strongly influences predictive performance. Incomplete clinical records, inconsistent coding, or missing environmental and behavioral variables can reduce model reliability.

- External validation is essential before clinical deployment. Predictive tools should demonstrate consistent performance across independent populations and practice environments.
- Risk communication with clients requires careful framing. Probabilistic predictions should be presented as estimates with uncertainty, not as definitive outcomes.
- Integrating multiple data sources improves prediction. Combining electronic medical records, insurance claims, wearable device data, and environmental exposures can enhance risk stratification for complex diseases.
- Companion animals can serve as sentinels for environmental and infectious disease risk. Veterinary data may provide early signals of shared health threats within a One Health framework.

DISCLOSURE

The authors have nothing to disclose.

REFERENCES

1. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15(4):361–87.
2. Chen YM, Hsiao TH, Lin CH, et al. Unlocking precision medicine: clinical applications of integrating health records, genetics, and immunology through artificial intelligence. J Biomed Sci 2025;32(1):16.
3. Kengne AP, Masconi K, Mbanya VN, et al. Risk predictive modelling for diabetes and cardiovascular disease. Crit Rev Clin Lab Sci 2014;51(1):1–12.
4. Schofield I, Brodbelt DC, Kennedy N, et al. Machine-learning based prediction of Cushing's syndrome in dogs attending UK primary-care veterinary practice. Sci Rep 2021;11(1):9035.
5. Biourge V, Delmotte S, Feugier A, et al. An artificial neural network-based model to predict chronic kidney disease in aged cats. J Vet Intern Med 2020;34(5):1920–31.
6. Golubnitschaja O, Baban B, Boniolo G, et al. Medicine in the early twenty-first century: paradigm and anticipation - EPMA position paper 2016. EPMA J 2016;7(1):23.
7. van Smeden M, Reitsma JB, Riley RD, et al. Clinical prediction models: diagnosis versus prognosis. J Clin Epidemiol 2021;132:142–5.
8. Baldwin JR, Caspi A, Meehan AJ, et al. Population vs individual prediction of poor health from results of adverse childhood experiences screening. JAMA Pediatr 2021;175(4):385–93.
9. O'Brien JS, McCullough A, Debes C, et al. Using pet insurance claims to predict occurrence of vector-borne and zoonotic disease in humans in the United States. Sci Rep 2024;15(1):14407.
10. Paynter AN, Dunbar MD, Creevy KE, et al. Veterinary big data: when data goes to the dogs. Animals (Basel) 2021;11(7). https://doi.org/10.3390/ani11071872.
11. Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. N Engl J Med 2016;375(13):1216–9.
12. Bajwa J, Munir U, Nori A, et al. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J 2021;8(2):e188–94.
13. Sexton C, Ruple A. Canine sentinels and our shared exposome. Science 2024;384(6701):1170–2.

14. O Neill DG, Church DB, McGreevy PD, et al. Prevalence of disorders recorded in dogs attending primary-care veterinary practices in England. PLoS One 2014; 9(3):e90501.
15. Pittaway C, Schofield I, Dobson J, et al. Incidence and risk factors for the diagnosis of lymphoma in dogs in UK primary-care practice. J Small Anim Pract 2019; 60(10):581–8.
16. O'Neill DG, Edmunds GL, Urquhart-Gilmore J, et al. Dog breeds and conformations predisposed to osteosarcoma in the UK: a VetCompass study. Canine Med Genet 2023;10(1):8.
17. Farrell S, Appleton C, Noble PJM, et al. PetBERT: automated ICD-11 syndromic disease coding for outbreak detection in first opinion veterinary electronic health records. Sci Rep 2023;13(1):18015.
18. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med 2019; 380(14):1347–58.
19. Wang Y, Wang L, Rastegar-Mojarad M, et al. Clinical information extraction applications: a literature review. J Biomed Inform 2018;77:34–49.
20. Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge. Nature 2023;620(7972):172–80.
21. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2017;2(4):230–43.
22. O'Neill DG, Church DB, McGreevy PD, et al. Approaches to canine health surveillance. Canine Genet Epidemiol 2014;1:2.
23. Egenvall A, Nødtvedt A, Penell J, et al. Insurance data for research in companion animals: benefits and limitations. Acta Vet Scand 2009;51(1):42.
24. Creevy KE, Akey JM, Kaeberlein M, et al, Dog Aging Project Consortium. An open science study of ageing in companion dogs. Nature 2022;602(7895):51–7.
25. Schmid SM, Sexton CL, Yoerger A, et al. Accuracy of owner-reported diagnoses for dogs enrolled in the dog aging project as compared to veterinary electronic medical records. PLoS One 2026;21(3):e0342427.
26. Sexton CL, Reichel C, Keyser A, et al, Dog Aging Project Consortium. Comparing owner reported and genetic breed identification reveals high concordance in a large cohort from the dog aging project. Sci Rep 2025;15(1):30493.
27. Mitek A, Jones D, Newell A, et al. Wearable devices in veterinary health care. Vet Clin North Am Small Anim Pract 2022;52(5):1087–98.
28. Colpoys J, DeCock D. Evaluation of the FitBark activity monitor for measuring physical activity in dogs. Animals (Basel) 2021;11(3):781.
29. Brown DC, Boston RC, Farrar JT. Use of an activity monitor to detect response to treatment in dogs with osteoarthritis. J Am Vet Med Assoc 2010;237(1):66–70.
30. Sacoor C, Leitão S, Domingues C, et al. Continuous activity monitoring using a wearable sensor in dogs with osteoarthritis: an exploratory case series. Animals (Basel) 2025;15(18):2639.
31. Thonen-Fleck C, Sharon KP, Enomoto M, et al. Physical activity monitors in companion animal chronic pain research-A review focused on osteoarthritis pain. Animals (Basel) 2025;15(14):2025.
32. Schuster LAH, de Carvalho AL, Dos Santos EAR, et al. Physical activity measured with an accelerometer in dogs following extracapsular stabilisation to treat cranial cruciate ligament rupture. J Small Anim Pract 2023;64(10):619–25.
33. Karimjee K, Harron RCM, Piercy RJ, et al. A standardised approach to quantifying activity in domestic dogs. R Soc Open Sci 2024;11(7):240119.
34. Martin KW, Olsen AM, Duncan CG, et al. The method of attachment influences accelerometer-based activity data in dogs. BMC Vet Res 2017;13(1):48.

35. Xue D, Collins D, Kauffman M, et al. Big data from small animals: integrating multi-level environmental data into the dog aging project. Rev Sci Tech 2023; 42:65–74.
36. Hosmer DW, Lemeshow S, Sturdivant RX. Applied logistic regression. 3rd edition. Wiley-Blackwell; 2013.
37. Breiman L. Statistical modeling: the two cultures (with comments and a rejoinder by the author). Stat Sci 2001;16(3):199–231.
38. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction. 2nd edition. Springer; 2009.
39. McEvoy FJ, Proschowsky HF, Müller AV, et al. Deep transfer learning can be used for the detection of hip joints in pelvis radiographs and the classification of their hip dysplasia status. Vet Radiol Ultrasound 2021;62(4):387–93.
40. Franco-Gonçalo P, Leite P, Alves-Pimenta S, et al. A computer-aided approach to canine hip dysplasia assessment: measuring femoral head–acetabulum distance with deep learning. Appl Sci (Basel) 2025;15(9):5087.
41. Kim Y, Kim J, Kim S, et al. Machine learning-based risk prediction model for canine myxomatous mitral valve disease using electronic health record data. Front Vet Sci 2023;10:1189157.
42. Sanaei N, Zamani-Ahmadmahmudi M, Nassiri SM. Development of machine learning models to predict clinical outcome and recovery time in dogs with parvovirus enteritis. Front Vet Sci 2025;12:1555714.
43. Lustgarten JL, Zehnder A, Shipman W, et al. Veterinary informatics: forging the future between veterinary medicine, human medicine, and one health initiatives-a joint paper by the association for veterinary informatics (AVI) and the CTSA One Health Alliance (COHA). JAMIA Open 2020;3(2):306–17.
44. Liu Y, Nordone SK, Yabsley MJ, et al. Quantifying the relationship between human lyme disease and Borrelia burgdorferi exposure in domestic dogs. Geospat Health 2019;14(1). https://doi.org/10.4081/gh.2019.750.
45. Matheson R, Sexton CL, Wise CF, et al. Silicone tags as an effective method of monitoring environmental contaminant exposures in a geographically diverse sample of dogs from the dog aging project. Front Vet Sci 2024;11:1394061.
46. Sexton CL, O'Brien J, Lytle J, et al. Testing for heavy metals in drinking water collected from dog aging project participants. PLOS Water 2025;4(8):e0000296.
47. Bonnett BN, Egenvall A, Olson P, et al. Mortality in insured Swedish dogs: rates and causes of death in various breeds. Vet Rec 1997;141(2):40–4.
48. Egenvall A, Bonnett BN, Häggström J, et al. Morbidity of insured Swedish cats during 1999-2006 by age, breed, sex, and diagnosis. J Feline Med Surg 2010;12-(12):948–59.
49. Bonnett BN, Egenvall A, Hedhammar A, et al. Mortality in over 350,000 insured Swedish dogs from 1995-2000: I. Breed-gender-age- and cause-specific rates. Acta Vet Scand 2005;46(3):105–20.
50. Egenvall A, Bonnett BN, Hedhammar Å, et al. Mortality in over 350,000 insured Swedish dogs from 1995–2000: II. breed-specific age and survival patterns and relative risk for causes of death. Acta Vet Scand 2005;46(3):121.
51. Egenvall A, Nødtvedt A, Häggström J, et al. Mortality of life-insured Swedish cats during 1999â2006: age, breed, sex, and diagnosis. J Vet Intern Med 2009;23(6): 1175–83.
52. Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. 2nd edition. Springer Nature; 2019.
53. Kelly CJ, Karthikesalingam A, Suleyman M, et al. Key challenges for delivering clinical impact with artificial intelligence. BMC Med 2019;17(1):195.

54. Gigerenzer G, Gaissmaier W, Kurz-Milcke E, et al. Helping doctors and patients make sense of health statistics. Psychol Sci Public Interest 2007;8(2):53–96.
55. Edwards A, Elwyn G, Covey J, et al. Presenting risk information–a review of the effects of "framing" and other manipulations on patient outcomes. J Health Commun 2001;6(1):61–82.
56. Elwyn G, Frosch D, Thomson R, et al. Shared decision making: a model for clinical practice. J Gen Intern Med 2012;27(10):1361–7.
57. Price WN 2nd, Cohen IG. Privacy in the age of medical big data. Nat Med 2019;25(1):37–43.
58. Obermeyer Z, Powers B, Vogeli C, et al. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019;366(6464):447–53.
59. Pasteur K, Diana A, Yatcilla JK, et al. Access to veterinary care: evaluating working definitions, barriers, and implications for animal welfare. Front Vet Sci 2024;11:1335410.
60. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 2019;1(5):206–15.
61. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25(1):44–56.
62. Amann J, Blasimme A, Vayena E, et al, Precise4Q Consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak 2020;20(1):310.
63. Friedman CP, Rubin JC, Sullivan KJ. Toward an information infrastructure for global health improvement. Yearb Med Inform 2017;26(1):16–23.
64. Goddard K, Roudsari A, Wyatt JC. Automation bias: empirical results assessing influencing factors. Int J Med Inform 2014;83(5):368–75.
65. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ 2015;350(jan07 4):g7594.
66. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med 2019;170(1):51–8.