Computational Molecular Biology 2025, Vol.15, No.6, 291-298
http://bioscipublisher.com/index.php/cmb

Research Insight Open Access

AI-Powered Prediction of Animal Disease Outbreaks Using Genomic Surveillance Data

Qiqi Zhou, Shiqiang Huang
Tropical Animal Resources Research Center, Hainan Institute of Tropical Agricultural Resources, Sanya, 572025, Hainan, China
Corresponding author: shiqiang.huang@hitar.org

Computational Molecular Biology, 2025, Vol.15, No.6 doi: 10.5376/cmb.2025.15.0029
Received: 17 Oct., 2025
Accepted: 27 Nov., 2025
Published: 16 Dec., 2025

Copyright © 2025 Zhou and Huang, This is an open access article published under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preferred citation for this article:
Zhou Q.Q., and Huang S.Q., 2025, AI-powered prediction of animal disease outbreaks using genomic surveillance data, Computational Molecular Biology, 15(6): 291-298 (doi: 10.5376/cmb.2025.15.0029)

Abstract The outbreak of animal epidemics poses a significant threat to livestock production, public health security and global food supply. Traditional epidemic surveillance methods have problems such as delayed response and low resolution of pathogen recognition. With the rapid development of genomics and artificial intelligence technologies, the integration of genomic monitoring data and AI models has provided a brand-new path for the early and high-precision prediction of animal epidemics. This study reviews the main approaches for collecting genomic data of animal diseases, including whole-genome sequencing (WGS) of pathogens, metagenomics and metagenomic analysis, etc. It systematically explores the AI algorithm systems used for epidemic modeling, such as supervised learning, deep learning and graph neural networks, with a focus on analyzing their advantages in temporal pattern recognition and spatial transmission path modeling. Through the case analysis of the avian influenza epidemic, this study constructed a high-resolution genomic monitoring dataset and combined feature engineering with model comparison and evaluation to verify the superiority of the AI model in prediction accuracy and response speed. This study demonstrates the huge potential of AI-enabled genomic monitoring technology in the proactive management of animal health, providing rapid response support for emerging epidemics and promoting the construction of an intelligent and data-driven animal disease prevention and control system.

Keywords Animal epidemic; Genomic monitoring; Artificial intelligence; Epidemic prediction; Machine learning

1 Introduction

Before discussing animal diseases, people often first think of the troubles they bring to livestock production, but the problems usually do not end there. Some pathogens can also spread across species. Once they enter the human population, they may evolve into a larger public health event. The outbreak of the epidemic is often accompanied by large-scale culling and trade stagnation, and its impact on the agricultural economy often lasts longer. What is more difficult is that the manifestations of such diseases are not determined by a single factor. The environment, genetic background and human activities may all make the transmission more complex and increase the difficulty of early detection and control (Yoon et al., 2025).

Meanwhile, the progress of high-throughput sequencing technology enables researchers to track the changes of pathogens more clearly, whether it is the evolutionary path or the transmission chain, which is more transparent than before. The integration of artificial intelligence - including machine learning and protein language models - has further enhanced the efficiency of data utilization, enabling this genomic information to predict epidemics faster and more accurately than traditional methods. Today's AI models usually analyze genomic, environmental indicators and epidemiological data together in order to assess risks more timely and assist in formulating early intervention strategies (Zhao et al., 2024; Lytras et al., 2025; Shafi et al., 2025).

This study will review the challenges currently faced in disease prediction, introduce the integration methods of artificial intelligence and genomics, and explore their impact on future monitoring systems. This study aims to develop an AI-based framework that uses genomic monitoring data to predict animal disease outbreaks. By integrating multivariable inputs and adaptive learning capabilities, it improves existing models to promote proactive disease management, reduce economic losses, and lower public health risks associated with zoonoses.

2 Genomic Data Acquisition in Animal Disease Surveillance

2.1 Whole-genome sequencing (WGS) of pathogens and real-time tracking platforms

In animal disease surveillance, many tasks often start with whole genome sequencing (WGS), as this method can provide very detailed genetic information, which can be used to identify the source of infection, determine changes in pathogen populations, and even track transmission chains. Although traditional post-culture sequencing has been in use for many years, it always encounters some problems, such as long time consumption, high biosafety requirements, and it is also difficult to obtain pure enough DNA once there is mixed infection or contamination in the sample. In recent years, the emergence of non-culture long-read sequencing technologies has changed this. For instance, the adaptive sampling of Oxford nanopores enables direct measurement of tissue samples without culture, which is faster and the quality can be maintained within an acceptable range. Current real-time tracking platforms connect WGS data with automated bioinformatics processes, enabling rapid completion of typing, clustering, and phylogenetic analysis, making it easier to identify epidemic outbreaks early (Bautista et al., 2023; Ghielmetti et al., 2023; Knijn et al., 2023).

2.2 Application of metagenomic and microbiome data in outbreak early warning

In terms of epidemic early warning, metagenomic sequencing is often used to observe the complete microbial community in animals. It can simultaneously detect known pathogens and also capture novel viruses that are usually not easily noticed. Previous studies have relied on this method to identify long-term latent viral infections in wild animals and also to find bacterial species related to respiratory diseases, indicating that its value in early warning is not only theoretical. By leveraging publicly available sequencing data, metagenomics can also identify emerging viruses with potential zoonotic risks, providing clues for proactive monitoring. After aggregating metagenomic data from different sources, researchers can more clearly observe multiple infections and changes in the ecological structure of pathogens, and this information is particularly crucial for epidemic risk assessment (Kawasaki et al., 2021; Prentice et al., 2024).

2.3 Sample sources and standardized data processing pipelines

In the actual process of conducting genomic monitoring, the types of samples are often more complex than imagined, including not only culture isolates but also tissue biopsies, environmental samples, and public sequencing banks. Problems often arise here as well, such as too low DNA content, severe contamination or the presence of mixed infections, all of which may affect the sequencing effect. Therefore, optimizing the sample processing procedure and sequencing strategy is particularly important. To enable data obtained from different laboratories to be compared with each other, standardized processes usually incorporate whole genome amplification, strict quality control and computational analysis. Meanwhile, if the monitoring and diagnostic procedures can be agreed upon among different institutions and follow the FAIR data principles, it will be more conducive to cross-departmental sharing and integration, thereby supporting the advancement of the "One Health" framework for animal diseases (Stärk et al., 2019; Pinto et al., 2024; Struelens et al., 2024).

3 AI Frameworks for Outbreak Prediction

3.1 Supervised learning and time-series modeling (e.g., LSTM, random forest)

When dealing with epidemic data, researchers often take time variations into account, so time series models like LSTM are frequently employed. It is good at capturing the correlations between sequences and is also relatively adaptable to complex nonlinear changes, so it performs well in predicting the trend of infection. The advantages of random forests, however, are somewhat different. They enhance stability through the combination of multiple decision trees and are less sensitive to common noise or heterogeneous variables in monitoring data. Both types of methods rely on diverse data for training, such as the number of cases, demographic data, and various environmental factors, which enables the model to detect some signs of an outbreak in the early stage of the epidemic. However, in different studies, the performance of the algorithm may also vary, specifically depending on the data characteristics and the target task (Ardabili et al., 2020; Alwakeel, 2025; Gao et al., 2025).

3.2 Deep learning and graph neural networks for transmission pathway modeling

When simulating transmission paths, some teams prefer to use graph structures to represent hosts, pathogens, and their connections. Therefore, graph neural networks (GNNS) have become commonly used tools. It can handle spatial dependencies and capture the interaction intensity between different nodes, thereby identifying propagation clusters or potential super-propagation points. Deep learning models themselves have an advantage in handling complex structures. If real-time genomic or epidemiological data is added, the details of the propagation process will be presented more clearly, and the judgment of hotspots will also be faster. In the face of the dynamic changes in communication patterns, such models can be said to be more flexible and more suitable for constantly updated data scenarios (Dubey et al., 2025; Kaur and Butt, 2025).

3.3 Model training, cross-validation, and performance evaluation metrics (e.g., AUC, F1-Score)

When building a predictive system, training and validation are often more crucial than the model itself. Researchers usually adopt methods such as K-fold cross-validation or time series segmentation to first observe how the model performs on unseen data, so as to avoid overfitting. AUC and F1-score are commonly used evaluation metrics. They can simultaneously reflect the accuracy, sensitivity and precision of classification. They are particularly important for epidemic prediction because incorrect judgments often bring actual risks. In recent years, some studies have tended to use integrated or hybrid models to combine the advantages of different AI algorithms, making the system more robust in different regions or different types of outbreak scenarios (Figure 1). This also highlights a fact: when applying artificial intelligence to disease surveillance, the evaluation framework itself is equally worthy of attention (Santosh, 2020; Jin et al., 2022).

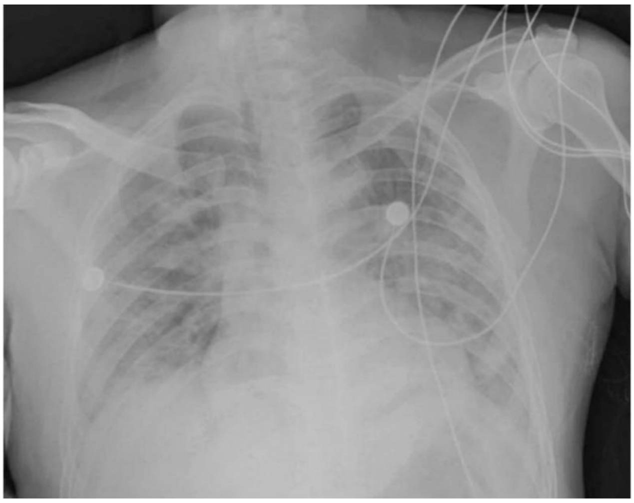

Figure 1 Chest X-ray: Bilateral focal consolidation, lobar consolidation, and patchy consolidation are clearly observed (Adopted from Santosh, 2020)

4 Data Integration and Feature Engineering Strategies

4.1 Integration of genotype-phenotype data and host-pathogen interactions

When studying the mechanisms of animal disease outbreaks, many teams will first look at the data of genotype and phenotype together, and then fill in the interaction information between the host and the pathogen. The purpose of doing this is not to build the model in one go, but to identify genetic variations related to virulence or host susceptibility earlier. When genomic information, phenotypic characteristics and interactive networks are combined, the details in disease dynamics that are not easily captured by a single data point will be revealed, and it can also make the explanations of artificial intelligence models closer to biological reality. However, the most attractive aspect of this holistic integration is that it can help identify the critical paths or markers that affect transmission and disease severity, which are valuable for subsequent monitoring and intervention (Baker et al., 2019).

4.2 Spatiotemporal, environmental, and host behavior data fusion

When building predictive models, people often do not start directly from the algorithm but first look at the background conditions of disease occurrence, including climate, land use, habitat conditions, as well as the migration and behavioral patterns of animals, etc. Once these spatio-temporal and ecological factors are combined with host behavior, the locations of epidemic hotspots and transmission risks become easier to identify. After multiple types of data are placed in the same framework, the model's simulation of propagation speed, dynamic changes and regional differences will be more realistic, and the spatiotemporal accuracy of the early warning system will also be higher. The integration of such data links ecological, epidemiological and behavioral factors together, which is more in line with the concept of "One Health" (Mumuni and Mumuni, 2024; Sultana and Rozony, 2025).

4.3 Feature selection and dimensionality reduction techniques (e.g., PCA, SHAP-based interpretability)

When dealing with high-dimensional and heterogeneous data, researchers often need to perform feature selection or dimensionality reduction first; otherwise, the model will run slowly and may even be disturbed by noise. Methods like Principal Component Analysis (PCA) can recombine variables with strong correlations into fewer components, reducing complexity while ensuring that necessary variance information is not lost. Meanwhile, explanatory tools such as SHAP help researchers determine which features are more critical in the model, making the prediction process more transparent and easier to screen out truly valuable predictors. Automated or hybrid feature engineering usually plays a role at this stage, making the model more robust and widely applicable by eliminating redundant features and retaining variables with large amounts of information (Mehra et al., 2025).

5 Case Study: AI-Driven Prediction of Avian Influenza Outbreaks

5.1 Construction and preprocessing of genomic surveillance datasets for avian influenza

Before starting the modeling, the research team usually has to first organize various monitoring information of avian influenza (AI) together, such as the virus sequence, sampling location, species category and sampling time. The data is not directly put into use. It still needs to be cleaned, and anomalies need to be picked out by methods such as isolating forests. Then, some classification information is converted into a numerical format that the model can recognize. For datasets of highly pathogenic avian influenza (HPAI), environmental indicators such as poultry density, temperature, water quality, and vegetation, as well as some features reflecting the interaction between the host and the pathogen, are sometimes added additionally. These pieces of information can more comprehensively present the multi-factor structure of the epidemic formation. Such comprehensive datasets enable the model to capture the complex correlations that affect the outbreak and spread of avian influenza (Madhava et al., 2024).

5.2 Comparative analysis of model performance and key feature identification

When comparing different AI models, researchers found that not all models performed the same. Some ensemble learning methods, such as random forests or gradient boosters, often achieve very high accuracy rates, with AUC frequently exceeding 0.85. However, deep learning models such as convolutional neural networks (CNNS) can still improve a bit on some datasets. The key features frequently "named" by the model are mostly concentrated on factors related to the environment and poultry farming, such as temperature, water quality index, vegetation index, as well as variables like poultry density and sampling time, because they can affect the spread of the virus and the way hosts gather. In addition, some machine learning methods can also identify genetic markers related to pathogenicity, making the prediction results more well-grounded in biological interpretation (Figure 2). These models have a relatively high spatial resolution and can also dynamically provide epidemic risk assessment, offering practical support for subsequent monitoring and intervention (Chadha et al., 2023).

5.3 Comparison with traditional surveillance in terms of accuracy and response time

Compared with traditional monitoring methods, artificial intelligence systems are often faster and more accurate in detecting epidemics. For instance, some research has utilized social media for digital symptom monitoring, and the results show that approximately 75% of real epidemics can be detected in advance, and one third of them are even earlier than the official announcements. Machine learning models that integrate environmental, biosafety and genomic data can also provide daily risk scores for individual poultry farms, with an AUC of approximately 0.88, and can identify potential high-risk locations several days in advance. These capabilities enable prevention and control measures to be initiated earlier and also make up for the common problems of delay and limited predictive ability in traditional diagnosis and reporting processes (Yousefinaghani et al., 2019; Yoo et al., 2021).

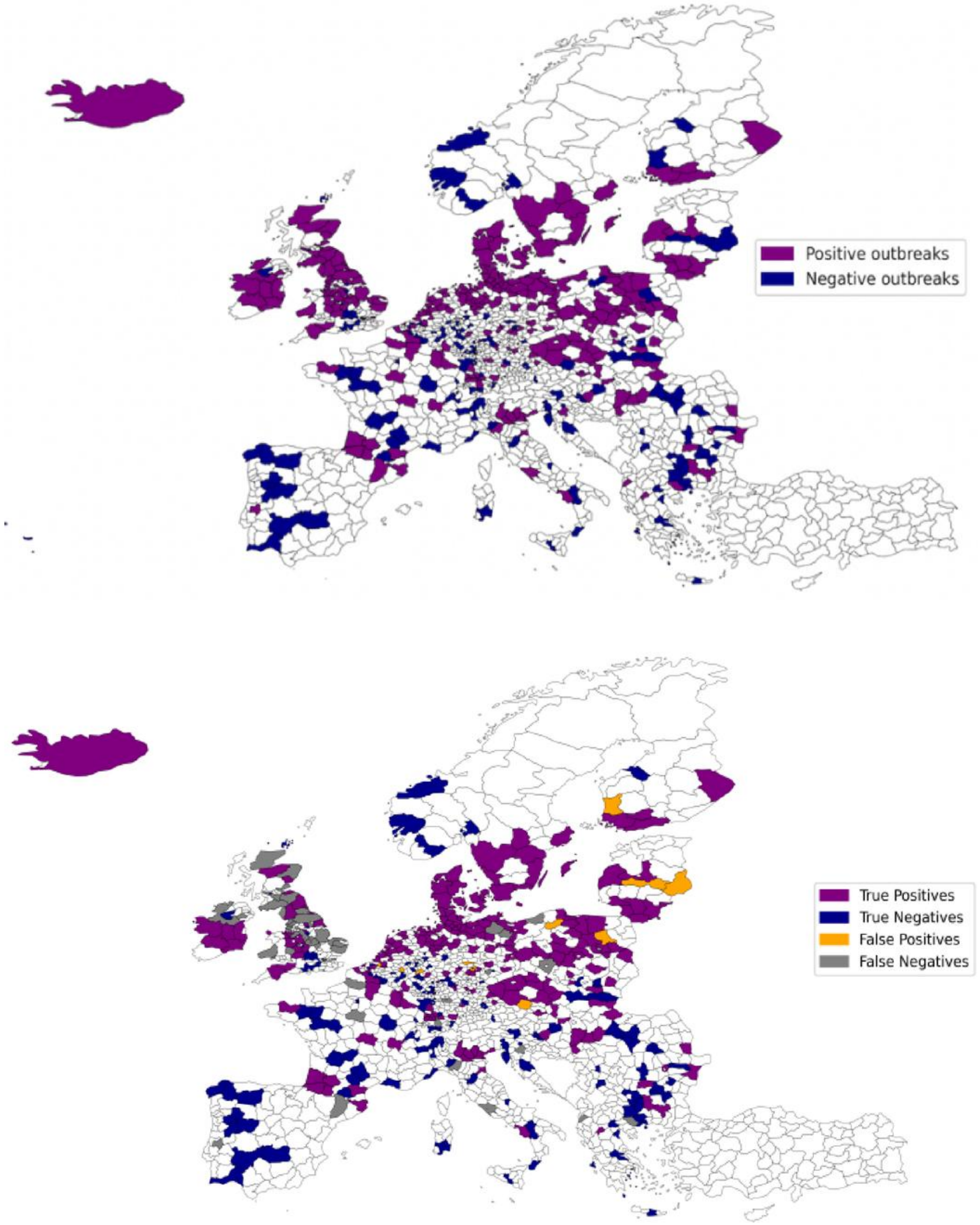

Figure 2 Test dataset labeled positive and negative used in testing the model on data from 2021 (Left). Predicted AIV outbreaks binned into True positives (Actual positive outbreaks predicted), True negatives (Actual negative outbreaks predicted), False positives (Negative outbreaks predicted as positives) and False negatives (Positive outbreaks predicted as negatives) to the right based on the test data from 2021 (Adopted from Opata et al., 2025)

6 Challenges and Limitations

6.1 Data imbalance and cross-regional data sharing barriers

Before discussing model performance, many studies have actually pointed out a fundamental issue long ago: the available outbreak data itself is not balanced enough. Although many AI models rely on historical samples, these samples are often concentrated in only a few regions or specific disease types, and the quantity is also limited, resulting in their poor performance in areas lacking data. In some places, although data has been accumulated, legal, institutional or technical restrictions make cross-regional sharing difficult, which also makes it less realistic to build more comprehensive data sets. Multi-source information such as genomic, environmental and epidemiological information is thus difficult to truly integrate, and the adaptability and predictive ability of the model will naturally be constrained (Ezanno et al., 2021; Keshavamurthy et al., 2022).

6.2 Model generalization capabilities and overfitting risks

Even if the data is sufficient, it is no easy task for the model to maintain stable performance across different regions, different host species, and even constantly changing pathogen strains. Especially on small-scale or imbalanced datasets, the model tends to remember the training samples easily. As a result, it performs averagingly in new scenarios, and the problem of overfitting is exposed. Infectious diseases are inherently dynamic and changing. Pathogens evolve rapidly, and environmental conditions may also change over time. All these require models to have a certain degree of adaptability. To avoid the model "failing" during application, means such as richer high-quality data, diverse validation methods and regularization are particularly crucial (Meslamani et al., 2024).

6.3 Ethical, privacy, and regulatory issues in cross-sectoral data integration

Cross-departmental data integration sounds very attractive, but when it comes to actually promoting it, various ethical and privacy concerns often arise first. Especially for sensitive data involving genomes or locations, the fields of animal health, environmental monitoring, and human public health themselves have different requirements and restrictions. Moreover, due to the significant differences in regulatory systems among countries and institutions, it is not easy to formulate a unified data usage agreement. To ensure smoother cooperation, transparency, data protection measures and ethical norms must all be clearly defined; otherwise, trust will be difficult to establish. All these indicate that for the rational use of artificial intelligence in animal disease surveillance, a clear and enforceable governance framework is indispensable, while also ensuring that privacy and ethical principles are truly implemented (Guo et al., 2023).

7 Future Directions and Application Prospects

Before discussing the future, many researchers have actually realized that relying solely on data from a single country or institution is difficult to support high-quality animal disease predictions. Therefore, a more practical approach would be to first place genomic information, environmental records, and epidemiological data from different regions on a shared platform, and then consider how to train the model. Once such a global collaboration framework is put into operation, real-time data intercommunication and joint model updates will no longer be difficult problems, and it can also alleviate the current old issues such as data being stored separately and significant regional differences. By leveraging international technological and experience exchanges, the speed and accuracy of prediction results will be more guaranteed, and it will be easier to form a consistent response to sudden risks. Moreover, this platform enables the model to continuously absorb new changes, whether it is pathogen mutations or fluctuations in the ecological environment, and adapt more quickly, ultimately making global animal health monitoring more stable.

As for the combination of artificial intelligence and edge computing, although it sounds highly technical, in practical scenarios, its significance lies more in "on-site processing". For instance, at the farm or community level, devices can directly process genomic or environmental data without having to upload it layer by layer to the cloud and wait for results. This will significantly reduce latency and decrease reliance on remote servers. Especially when the epidemic just emerged and many situations were still unclear, local rapid judgment was often more crucial than high computing power. In this way, risk assessment can be closer to the scene, safety management measures can be implemented more quickly, and farmers and local administrative departments can obtain more direct and immediately usable information, thereby improving response speed and resource utilization efficiency.

Another trend in the future application of artificial intelligence is to attempt to consider cross-species disease prediction, vaccine development, and policy planning under the same train of thought. As the genomic data of animals and humans can already be mutually verified, AI has the opportunity to identify the potential risks of certain zoonotic diseases in advance and provide references for the selection of vaccine targets, thus making the formulation of protective measures less lagging behind. If this predictive ability is further combined with the policy framework, the decision-making process can be more flexible, more evidence-based, and more conducive to the rational allocation of resources. This approach of integrating veterinary science, public health and regulatory systems is actually more in line with the concept of "One Health" and can better help society deal with those complex and rapidly changing epidemic threats.

Acknowledgments

We extend our sincere thanks to two anonymous peer reviewers for their invaluable feedback on the manuscript of this paper.

Conflict of Interest Disclosure

The authors affirm that this research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Alwakeel M., 2025, AI-assisted real-time monitoring of infectious diseases in urban areas, Mathematics, 13(12): 1911.
https://doi.org/10.3390/math13121911

Ardabili S., Mosavi A., Ghamisi P., Ferdinand F., Várkonyi-Kóczy A., Reuter U., Rabczuk T., and Atkinson P., 2020, COVID-19 outbreak prediction with machine learning, Algorithms, 13(10): 249.
https://doi.org/10.1101/2020.04.17.20070094

Baker Y., Tang T., and Allen G., 2019, Feature selection for data integration with mixed multiview data, Annals of Applied Statistics, 14(4): 1676-1698.
https://doi.org/10.1214/20-aoas1389

Bautista C., Jaswant G., French H., Campbell K., Durrant R., Gifford R., Kia G., Ogoti B., Hampson K., and Brunker K., 2023, Whole genome sequencing for rapid characterization of rabies virus using nanopore technology, Journal of Visualized Experiments, 191: e65414.
https://doi.org/10.3791/65414

Chadha A., Dara R., Pearl D., Sharif S., and Poljak Z., 2023, Predictive analysis for pathogenicity classification of H5Nx avian influenza strains using machine learning techniques, Preventive Veterinary Medicine, 216: 105924.
https://doi.org/10.1016/j.prevetmed.2023.105924

Dubey A., Gupta K., S., Sankari A., Namani S., and Hemalatha R., 2025, Predicting disease outbreaks with AI: an in-depth analysis of infectious diseases surveillance, In: 2025 International Conference on Frontier Technologies and Solutions (ICFTS), IEEE, pp.1-7.
https://doi.org/10.1109/icfts62006.2025.11031614

Ezanno P., Picault S., Beaunée G., Bailly X., Muñoz F., Duboz R., Monod H., and Guégan J., 2021, Research perspectives on animal health in the era of artificial intelligence, Veterinary Research, 52(1): 40.
https://doi.org/10.1186/s13567-021-00902-4

Gao S., Chakraborty A., Greiner R., Lewis M., and Wang H., 2025, Early detection of disease outbreaks and non-outbreaks using incidence data based on feature-based time series classification and machine learning, PLOS Computational Biology, 21(2): e1012782.
https://doi.org/10.1371/journal.pcbi.1012782

Ghielmetti G., Loubser J., Kerr T., Stuber T., Thacker T., Martin L., O'Hare M., Mhlophe S., Okunola A., Loxton A., Warren R., Moseley M., Miller M., and Goosen W., 2023, Advancing animal tuberculosis surveillance using culture-independent long-read whole-genome sequencing, Frontiers in Microbiology, 14: 1307440.
https://doi.org/10.3389/fmicb.2023.1307440

Guo W., Lv C., Guo M., Zhao Q., Yin X., and Zhang L., 2023, Innovative applications of artificial intelligence in zoonotic disease management, Science in One Health, 2: 100045.
https://doi.org/10.1016/j.soh.2023.100045

Jin W., Dong S., Yu C., and Luo Q., 2022, A data-driven hybrid ensemble AI model for COVID-19 infection forecast using multiple neural networks and reinforced learning, Computers in Biology and Medicine, 146: 105560.
https://doi.org/10.1016/j.compbiomed.2022.105560

Kaur J., and Butt Z., 2025, AI-driven epidemic intelligence: the future of outbreak detection and response, Frontiers in Artificial Intelligence, 8: 1645467.
https://doi.org/10.3389/frai.2025.1645467

Kawasaki J., Kojima S., Tomonaga K., and Horie M., 2021, Hidden viral sequences in public sequencing data and warning for future emerging diseases, mBio, 12(4): 10.1128/mbio. 01638-21.
https://doi.org/10.1128/mbio.01638-21

Keshavamurthy R., Dixon S., Pazdernik K., and Charles L., 2022, Predicting infectious disease for biopreparedness and response: a systematic review of machine learning and deep learning approaches, One Health, 15: 100439.
https://doi.org/10.1016/j.onehlt.2022.100439

Knijn A., Michelacci V., Gigliucci F., Tozzoli R., Chiani P., Minelli F., Scavia G., Ventola E., and Morabito S., 2023, IRIDA-ARIES Genomics: a key player in the One Health surveillance of infectious diseases in Italy, Frontiers in Public Health, 11: 1151568.
https://doi.org/10.3389/fpubh.2023.1151568

Lytras S., Lamb K., Ito J., Grove J., Yuan K., Sato K., Hughes J., and Robertson D., 2025, Pathogen genomic surveillance and the AI revolution, Journal of Virology, 99(2): e01601-24.
https://doi.org/10.1128/jvi.01601-24

Madhava G., Yousefinaghani S., Dara R., Poljak Z., Sharif S., Laera M., Sangiorgi F., Verdi F., Roversi R., Garuti M., Calzolari M., Gibertoni M., Martello S., and Gherardi M., 2024, Advanced machine learning and deep learning approaches for predicting avian influenza outbreaks, EPRA International Journal of Research & Development (IJRD), 9(9): 21-30.
https://doi.org/10.36713/epra18191

Mehra S., Mohan S., and Agarwal N., 2025, Case study on the effectiveness of feature engineering techniques for performance improvement of data mining applications, In: 2025 International Conference on Intelligent Control, Computing and Communications (IC3), IEEE, pp.248-252.
https://doi.org/10.1109/ic363308.2025.10956338

Meslamani A., Sobrino I., and De La Fuente J., 2024, Machine learning in infectious diseases: potential applications and limitations, Annals of Medicine, 56(1): 2869.
https://doi.org/10.1080/07853890.2024.2362869

Mumuni A., and Mumuni F., 2024, Automated data processing and feature engineering for deep learning and big data applications: a survey, Journal of Intelligent Information Systems, 3(2): 113-153.
https://doi.org/10.1016/j.jiixd.2024.01.002

Opata M., Lavarello-Schettini A., Semenza J., and Rocklöv J., 2025, Predictiveness and drivers of highly pathogenic avian influenza outbreaks in Europe, Scientific Reports, 15(1): 20286.
https://doi.org/10.1038/s41598-025-04624-x

Pinto D., Themudo G., Pereira A., Botelho A., and Cunha M., 2024, Rescue of *Mycobacterium bovis* DNA from cultured samples for robust whole-genome sequencing during animal TB surveillance, International Journal of Molecular Sciences, 25(7): 3869.
https://doi.org/10.3390/ijms25073869

Prentice M., Gilbertson M., Storm D., Turner W., Walsh D., Pinkerton M., and Kamath P., 2024, Metagenomic sequencing identifies microbes associated with pneumonia-related fatalities of white-tailed deer (*Odocoileus virginianus*), Microbial Genomics, 10(3): 001214.
https://doi.org/10.1099/mgen.0.001214

Santosh K., 2020, AI-driven tools for coronavirus outbreak: need of active learning and cross-population models on multimodal data, Journal of Medical Systems, 44(5): 93.
https://doi.org/10.1007/s10916-020-01562-1

Shafi M., Shabir S., Jan S., Wani Z., Rather M., Beigh Y., Kamil S., Mir M., Rafiq A., and Shah S., 2025, The role of artificial intelligence in detecting avian influenza virus outbreaks: a review, Open Veterinary Journal, 15(5): 1880-1894.
https://doi.org/10.5455/ovj.2025.v15.i5.4

Stärk K., Pękala A., and Muellner P., 2019, Use of molecular and genomic data for disease surveillance in aquaculture: toward improved evidence-based decisions, Preventive Veterinary Medicine, 167: 190-195.
https://doi.org/10.1016/j.prevetmed.2018.04.011

Struelens M., Ludden C., Werner G., Sintchenko V., Jokelainen P., and Ip M., 2024, Real-time genomic surveillance for enhanced control of infectious diseases and antimicrobial resistance, Frontiers in Science, 2: 1298248.
https://doi.org/10.3389/fsci.2024.1298248

Sultana R., and Rozony F., 2025, A meta-analysis of artificial intelligence-driven data engineering: evaluating cloud-based integration models, ASRC Procedia: Global Perspectives in Science and Scholarship, 1(01): 193-214.
https://doi.org/10.63125/8a5k2j16

Yoo D., Song Y., Choi D., Lim J., Lee K., and Kang T., 2021, Machine learning-driven dynamic risk prediction for highly pathogenic avian influenza at poultry farms in Korea, Transboundary and Emerging Diseases, 69(5): 2667-2681.
https://doi.org/10.1111/tbed.14419

Yoon H., Kim K., and Son Y., 2025, Enhancing AI-based risk prediction for animal disease outbreaks in Korea using KAHIS big data, Journal of Veterinary Science and Research, 12(1): 4968.
https://doi.org/10.52338/jovsr.2025.4968

Yousefinaghani S., Dara R., Poljak Z., Bernardo T., and Sharif S., 2019, Assessment of Twitter's potential for outbreak detection: avian influenza case study, Scientific Reports, 9(1): 54388.
https://doi.org/10.1038/s41598-019-54388-4

Zhao A., Li S., Cao Z., Hu P., Wang J., Xiang Y., Xie D., and Lu X., 2024, AI for science: predicting infectious diseases, Journal of Safety Science and Resilience, 5(2): 130-146.
https://doi.org/10.1016/j.jnlssr.2024.02.002

Disclaimer/Publisher's Note

The statements, opinions, and data contained in all publications are solely those of the individual authors and contributors and do not represent the views of the publishing house and/or its editors. The publisher and/or its editors disclaim all responsibility for any harm or damage to persons or property that may result from the application of ideas, methods, instructions, or products discussed in the content. Publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.