Metaheuristic Optimization Review

♦ ISSN: 2692-4048 Vol: 02, No. 01, 2024 ♦ PP. 11–15

Article DOI: https://doi.org/10.54216/MOR.020102



REVIEW ARTICLE

A Review on the Role of Machine Learning in Predicting the Spread of Infectious Diseases

S. K. Towfek[1,*] **Mohamed Elkanzi**[2]

[1] Computer Science and Intelligent Systems Research Center, Blacksburg 24060, Virginia, USA

[2] School of ICT, Faculty of Engineering, Design and Information & Communications Technology (EDICT), Bahrain Polytechnic, Isa Town, Bahrain

Emails: sktowfek@jcsis.org · mohamed.elkanzi@polytechnic.bh

Received: April 20, 2024 Revised: July 14, 2024 Accepted: November 09, 2024 ⋆ Corresponding author

ABSTRACT

AI and the development of machine-learning (ML) systems are expected to play a crucial role in preventing and controlling infectious diseases as part of global health issues. Conventional epidemic models typically provide a narrow perspective on disease distribution and causation, which motivates the use of AI/ML solutions. Some of these tools use genomic, environmental, and patient information to improve forecast accuracy and facilitate real-time disease surveillance. Human-driven models of pandemic identification are increasingly being replaced by sophisticated artificial-intelligence models, including deep learning and advanced neural networks, that identify patterns, estimate future outbreaks, and guide public-health interventions. Examples demonstrate the efficiency of ML approaches in combating antimicrobial resistance, tuberculosis relapse, and spatiotemporal modeling of diseases such as measles and COVID-19. Nevertheless, data standardization, scalability, ethics, and bias remain major limitations. Addressing unfairness requires transparency, patient-data confidentiality, and equitable deployment of AI systems. Practical implementation also requires cross-sector cooperation and global data sharing across varied healthcare conditions. Future developments point to opportunities for enriching epidemic prediction models by combining genomic precision systems, explainable AI, and interdisciplinary studies. This review shows how AI/ML can transform infectious-disease management, calls for responsible innovation and ethical deployment, and encourages international collaboration to strengthen global health preparedness against emerging diseases.

Keywords: Artificial intelligence ▪ Machine learning ▪ Epidemic prediction ▪ Infectious diseases ▪ Public health ▪ Data integration

1. INTRODUCTION

Epidemics and pandemics have created considerable challenges for health systems worldwide and have left a major imprint on global health dynamics. COVID-19, Zika, and Ebola illustrate the need for reliable models that can predict disease spread and support control strategies. Traditional disease-spread studies based on linear epidemiological models have limitations in explaining how diseases evolve. Consequently, researchers have introduced more efficient technologies, including artificial intelligence (AI), machine learning (ML), and mathematical modeling, to predict epidemic trends. By analyzing communities and simulating disease trajectories, these methods create opportunities for early identification and more adaptive public-health interventions. This review explores the use of AI and ML in infectious-disease forecasting and emphasizes their contribution to predicting global health threats.

© 2024 The Author(s). Published by ASPG. This article is licensed under CC BY 4.0.

1.1 Introduction to Epidemic Forecasting and AI Integration

Epidemic forecasting is one of the healthcare areas that has benefited substantially from AI because of the large volume of available data. AI improves predictions about disease occurrence and progression by using historical and real-time data, including disease history, environmental conditions, population movement, and health-sector indicators. Recurrent neural networks and graph-based algorithms, for example, can estimate disease dissemination and guide resource distribution. Such tools are useful for public-policy design, particularly in emerging cases such as COVID-19 and influenza outbreaks. AI-driven models can also process large datasets with spatiotemporal resolution, enabling detailed monitoring of epidemics across regions [1].

A major advantage of AI in epidemic forecasting is its ability to handle new diseases through data revision, real-time monitoring, and uncertainty characterization. This contrasts with traditional models that rely on fixed parameters, equations, and assumptions. During COVID-19, AI models were used to estimate hospitalizations and mortality, thereby supporting rationing and planning. Important challenges remain, including data normalization, privacy protection, bias mitigation, and the integration of heterogeneous sources. These challenges require careful governance if AI is to be deployed safely in public health [2, 3].

1.2 Application of Machine Learning and Big Data in Disease Prediction

Machine learning and big-data analytics allow public-health systems to exploit diverse datasets, including clinical records, mobility data, climate information, genomic sequences, and social-media signals. These technologies can detect patterns that are difficult to capture with conventional methods and can support earlier warnings of outbreaks. Supervised learning, unsupervised learning, reinforcement learning, and deep learning have all been applied to disease prediction, diagnosis, and response planning. Their value is especially evident when data arrive continuously and must be interpreted quickly for operational decisions [4, 5].

Big data also improves the scale and granularity of epidemiological analysis. By combining patient-level information with environmental and demographic data, ML models can identify risk factors, estimate local vulnerability, and improve allocation of medical resources. However, the use of large datasets raises concerns about privacy, representativeness, algorithmic fairness, and the reliability of predictions when models are transferred from one population to another.

1.3 Case Studies of Epidemic Prediction Models

Several infectious-disease case studies show the potential of AI and ML. In antimicrobial resistance, ML models have been combined with clinical decision-support systems and mass-spectrometry data to predict resistance profiles and support treatment selection. In tuberculosis, predictive models such as logistic regression, random forests, and boosting have been used to estimate relapse risk and reduce dependence on resource-intensive experimental assays. In COVID-19, deep-learning and reinforcement-learning approaches have supported transmission forecasting, early screening, and evaluation of intervention policies [12, 13, 18, 19].

Spatiotemporal models have also been applied to measles and other infectious diseases. Neural-network models can capture nonlinear temporal patterns and spatial hierarchies, while hybrid approaches combine mechanistic epidemiological models with data-driven learning. These examples indicate that AI/ML can supplement classical SIR and SEIR models by improving adaptability and predictive accuracy under complex real-world conditions [14, 16].

1.4 Challenges in Modeling Disease Dynamics

Disease dynamics are shaped by biological, social, environmental, and behavioral processes. This complexity makes modeling difficult, especially when data are incomplete, biased, delayed, or collected using inconsistent standards. Models trained on one region may not generalize to another because of differences in healthcare access, reporting practices, population density, intervention policies, and pathogen evolution. These problems can lead to misleading forecasts and poor public-health decisions if uncertainty is not properly communicated.

Another challenge is interpretability. Many high-performing AI models, particularly deep neural networks, behave as black boxes. Public-health authorities need transparent explanations to justify interventions that affect communities. Therefore, model performance must be balanced with explainability, ethical review, and accountability.

1.5 Future Directions and Implications for Public Health

Future epidemic forecasting should integrate genomic, environmental, clinical, and socioeconomic data within explainable and update-friendly models. Federated learning and privacy-preserving analytics may help institutions collaborate without exposing sensitive patient data. Interdisciplinary cooperation among epidemiologists, clinicians, data scientists, policymakers, and ethicists is essential for building models that are accurate, fair, and usable in practice.

The One Health perspective also offers an important direction by linking human, animal, and environmental health. Such integration can improve pandemic preparedness by identifying ecological drivers of infectious-disease emergence and by supporting coordinated prevention strategies [9, 10].

2. LITERATURE REVIEW

Global epidemics and pandemics have emphasized the need for organized control systems with accurate prediction techniques at their core. Understanding how infectious diseases emerge and spread is crucial for formulating intervention strategies that reduce social and economic burdens. Recent advances in AI, ML, big data, and mathematical modeling provide new opportunities to address the complexity of infectious-disease dynamics.

AI has advanced infection prediction through ML and deep-learning approaches. Existing work can be grouped into three main categories: public-health data used to prevent regional disease spread, patient medical data used for infection detection, and combined public and patient data used to estimate propagation within populations [11]. For multidrug-resistant gram-negative bacteria, AI-based clinical decision-support systems using MALDI-TOF MS data and random forests

Table 1. Summary of Literature Review

Ref.	Study Focus	Methodology/Technology Used	Key Findings/Contributions
[11]	AI and ML applications in infection prediction	AI, ML, DL, public-health data, patient data	AI/ML improved disease prediction and public-health interventions.
[12]	AI-based CDSS for antimicrobial-resistance prediction	ML with MALDI-TOF MS; random forest	High accuracy in predicting antibiotic resistance; random forest AUC of 0.95.
[13]	ML techniques for TB relapse prediction	Logistic regression, random forests, boosting	Efficient prediction of TB relapse with reduced resource use.
[14]	AI in infectious-disease modeling	SIR, SEIR, stochastic models, big-data analytics	Enhanced epidemic prediction; ethical considerations highlighted.
[15]	Real-time epidemic prediction	Cloud computing, wearable-device data, epidemiological models	Improved real-time outbreak prediction and resource allocation.
[16]	Spatiotemporal measles transmission	High-dimensional neural networks, TSIR model	SFNN outperformed TSIR in long-term forecasting; spatial hierarchies were observed.
[17]	Climate-driven marine pathogens	Genomic and environmental data integration	Revealed climate influence on *Vibrio* spread; predicted infections with high sensitivity.
[18]	Dynamic COVID-19 modeling in Medan	DEEPCOV, deep learning, transmission modeling	Enhanced forecasting and public-health responses to COVID-19.
[19]	Early COVID-19 detection from tracker data	Reinforcement learning, multilayer perceptron	Reduced data needs and improved early detection.
[20]	ML for predicting TB incidence	Predictive modeling, visualization, ML techniques	Accurate TB-incidence prediction; visualization enhanced insights.
[21]	AI/ML in infectious-disease prediction	Public-health data, patient data, combined datasets	Promising applications for managing disease spread.
[22]	Early COVID-19 detection	Reinforcement learning, time-series classification	Earlier predictions with reduced data needs.
[23]	SIR modeling of lockdown in India	SIR modeling, simulation, Microsoft Excel	Lockdowns reduced infection rates; scenarios clarified intervention effects.
[24]	Explainable AI for infectious-disease prediction	Explainable AI, deep learning	XAI can make predictions transparent and applicable in health systems.
[25]	Ecological factors in Oropouche fever	Ecological predictors, midges, landscape change	Ecological factors influence spread; improved models are needed.
[26]	AI detection of foodborne pathogens	Self-transfer deep learning, Raman spectra	Accurate pathogen detection with practical significance.
[27]	Eco-evolutionary *Vibrio* dynamics	ML, genomic–environmental data, sensitivity analysis	El Niño effects on *Vibrio* dynamics; integrated modeling demonstrated.
[28]	SFNN model for measles forecasting	Feed-forward neural networks, mechanistic models	SFNN identified drivers; hybrid models proposed for accuracy.
[29]	Deep learning for COVID-19 in Medan	DEEPCOV, isolation measures, task-force data	Advanced modeling improved understanding of COVID-19 dynamics.
[30]	AI/ML for global epidemic prediction	Predictive modeling, mathematical models, big-data analytics	AI/ML improved prediction and control of infectious-disease outbreaks.

achieved strong predictive performance for antibiotic resistance [12]. In tuberculosis, ML techniques have been studied to predict relapse rates using experimental and treatment-specific variables, offering efficient alternatives to costly recurrence evaluations [13].

Infectious-disease modeling has evolved from classical SIR and SEIR models toward stochastic, network-based, and big-data methods. These advances improve predictive capacity but also introduce challenges in parameter estimation, validation, real-time data use, privacy, and risk communication [14]. Cloud computing, wearable devices, and real-time analytics further support outbreak prediction and resource allocation [15]. Deep neural networks and hybrid mechanistic models have improved long-term forecasting for measles dynamics, while climate-sensitive pathogen studies show the importance of integrating genomic and environmental data [16, 17].

Artificial intelligence, machine learning, and mathematical models have transformed the understanding and prediction of infectious-disease outbreaks. More precise forecasting can trigger pre-emptive actions to control transmission, support early detection, improve response strategies, and strengthen outbreak management. Although these technologies have substantial potential, optimization challenges remain in real-life settings, especially under diverse and dynamic environmental conditions. Future research should refine predictive algorithms and improve data integration to strengthen global responses to infectious-disease threats.

3. DISCUSSION

The analysis of genomics, environmental data, and novel ML algorithms advances epidemic prediction and public-health planning. AI/ML systems can identify hidden correlations, forecast disease trajectories, and support intervention design. Their ability to integrate heterogeneous data sources makes them valuable for both routine surveillance and emergency response.

3.1 Integration of AI and ML in Epidemic Predictions

AI and ML support epidemic prediction by combining diverse data streams, including mobility, clinical records, genomic information, weather, and social behavior. These systems improve the speed and granularity of forecasting and can update predictions when new information becomes available. They are particularly useful for emerging diseases, where mechanistic knowledge may be incomplete and rapid decisions are required.

3.2 Key Achievements in Forecasting Accuracy

Important achievements include improved antimicrobial-resistance prediction, TB relapse estimation, COVID-19 screening, and measles forecasting. Random-forest algorithms have performed well in resistance prediction, while deep-learning architectures such as ForecastNet and spatiotemporal neural networks have strengthened multi-step time-series forecasting and disease-dynamics modeling [33, 34]. These advances show that AI/ML can improve both predictive accuracy and operational decision-making.

3.3 Challenges and Ethical Considerations

Despite major progress, AI and ML in healthcare face challenges related to dataset bias, scalability, fairness, and trust. Bias can enter through training data and distort treatment or intervention decisions, especially for disadvantaged groups. Models that perform well in laboratories may fail in clinical or public-health settings because of data shifts and implementation constraints. Ethical issues include patient privacy, informed consent, transparency, accountability, and equitable access to AI-enabled health benefits [35, 36].

3.4 Future Directions for AI in Public Health

Future systems should combine external factors, genomic information, and next-generation ML to create accurate alert systems. Integrating genetic, climate, and socioeconomic data can capture relationships that influence transmission dynamics and can enable targeted public-health action. Tools such as learnMET illustrate how genomic and environmental data can be integrated for prediction across multiple environments [38].

The deployment of these approaches requires cooperation across sectors, strong ethical compliance, and international data-sharing programs. Healthcare, environmental, governmental, and research institutions must share data responsibly to reduce fragmentation. Trust, clear governance, privacy protection, and transparent standards are necessary for effective collaboration [37].

4. CONCLUSION

AI and ML have become revolutionary methodologies in infectious-disease forecasting and management. These technologies facilitate more accurate outbreak detection, tracking, and control through advanced computational procedures. AI/ML models can extract important information from genomic and environmental data, health monitoring systems, and other sources, enabling better public-health intervention. Their application to antimicrobial resistance, tuberculosis relapse, and pandemic dynamics demonstrates their potential to transform disease-control models.

However, adoption of AI/ML systems also has drawbacks. Data standardization, scalability, and inherent dataset bias can compromise prediction reliability. Ethical constraints involving patient privacy, algorithmic transparency, equity of access, and integration into healthcare facilities must also be addressed. Infectious-disease transmission requires models that are easy to update, interpretable, and capable of accommodating multiple data sources and types.

The most important solutions are interdisciplinary teamwork, strict compliance with ethical rules, and the global dissemination of large standardized databases. Collaboration among governments, healthcare institutions, researchers, and technology providers can help build robust, fair, and sustainable AI/ML systems. Responsible innovation, explainable models, and international cooperation are therefore essential for strengthening preparedness against future epidemics and pandemics.

REFERENCES

[1] A. Adadi, M. Berrada, et al., “AI-driven epidemic monitoring with spatiotemporal resolution,” 2024.

[2] Y. Hu, J. Jacob, G. J. Parker, D. J. Hawkes, J. R. Hurst, and D. Stoyanov, “The challenges of deploying artificial intelligence models in a rapidly evolving pandemic,” *Nature Machine Intelligence*, vol. 2, no. 6, pp. 298–300, 2020.

[3] B. Murdoch, “Privacy and artificial intelligence: challenges for protecting health information in a new era,” *BMC Medical Ethics*, vol. 22, no. 1, pp. 1–5, 2021.

[4] H. S. B. et al., “Machine learning and big data analytics for disease prediction,” 2019.

[5] S. S. et al., “Machine learning algorithms for disease prediction,” *BMC Medical Informatics and Decision Making*, vol. 19, no. 1, pp. 1–16, 2019.

[6] X. B. Lin et al., “All few-mode fiber spatiotemporal mode-locked figure-eight laser,” *Journal of Lightwave Technology*, 2024.

[7] A. A. et al., “Big-data approaches for infectious disease prediction,” 2024.

[8] M. M. et al., “Machine learning, mathematical modeling, and infectious-disease preparedness,” 2024.

[9] A. M. Campbell, C. Hauton, C. Baker-Austin, R. van Aerle, and J. Martinez-Urtaza, “An integrated eco-evolutionary framework to predict population-level responses of climate-sensitive pathogens,” *Current Opinion in Biotechnology*, vol. 80, 2023.

[10] S. B. et al., “One Health approaches and pandemic planning using AI,” 2024.

[11] M. M. Siddique, M. M. Bin Seraj, M. N. Adnan, and S. M. Galib, “Artificial Intelligence for Infectious Disease Detection: Prospects and Challenges,” in *Surveillance, Prevention, and Control of Infectious Diseases*, pp. 1–22, 2024.

[12] F. P. et al., “Artificial intelligence clinical decision-support system for antimicrobial resistance prediction,” 2024.

[13] A. P. et al., “Sviluppo di modelli di Machine Learning per prevedere la relapse al fine di valutare il tempo di guarigione dopo un’infezione da Mycobacterium Tuberculosis usando dati del Relapsing Mouse Model,” 2024.

[14] J. Huang and J. S. Morris, “Infectious Disease Modeling,” *Annual Review of Statistics and Its Application*, 2024.

[15] S. Bi, Y. Lian, and Z. Wang, “Research and Design of a Financial Intelligent Risk Control Platform Based on Big Data Analysis and Deep Machine Learning,” 2024.

[16] W. G. Madden et al., "Deep neural networks for endemic measles dynamics: Comparative analysis and integration with mechanistic models," *PLoS Computational Biology*, vol. 20, no. 11, e1012616, 2024.

[17] A. M. Campbell, C. Hauton, C. Baker-Austin, R. van Aerle, and J. Martinez-Urtaza, "An integrated eco-evolutionary framework to predict population-level responses of climate-sensitive pathogens," *Current Opinion in Biotechnology*, vol. 80, 2023.

[18] R. U. Ginting, "Model Transmisi Dinamis dari Covid-19 dengan Adanya Isolasi Menggunakan Pendekatan Deep Learning," 2024.

[19] A. Sarwar, A. Almadani, and E. O. Agu, "Early Time Series Classification Using Reinforcement Learning for Pre-Symptomatic Covid-19 Screening From Imbalanced Health Tracker Data," *IEEE Journal of Biomedical and Health Informatics*, pp. 1–11, 2024.

[20] H. M. Hamna, S. A. Jose, A. Jirawattanapanit, and K. Mathew, "A comprehensive study on tuberculosis prediction models: Integrating machine learning into epidemiological analysis," *Journal of Theoretical Biology*, vol. 597, 111988, 2025.

[21] S. Merchant, Y. Elmir, F. Kacimi, and L. Boubchir, "Artificial Intelligence for Infectious Disease Prediction and Prevention: A Comprehensive Review," 2024.

[22] A. Sarwar, A. Almadani, and E. O. Agu, "Early Time Series Classification Using Reinforcement Learning for Pre-Symptomatic Covid-19 Screening," 2024.

[23] R. S. Yadav, "A Study of the Effectiveness of Lockdown During COVID-19 Using Machine Learning Methods: A Case Study of India," 2024.

[24] B. Addaali, R. Latif, and A. Saddik, "Addressing the Spread of Infectious Diseases in the Era of Explainable AI," in *World Conference on Complex Systems*, pp. 1–6, 2024.

[25] I. Postigo-Hidalgo and J. F. Drexler, "Probing Oropouche fever ecology beyond the Amazon," *The Lancet Infectious Diseases*, 2024.

[26] D. Li, Y. Zhu, A. Mehmood, Y. Liu, X. Qin, and Q. Dong, "Intelligent identification of foodborne pathogenic bacteria by self-transfer deep learning and ensemble prediction based on single-cell Raman spectrum," *Talanta*, vol. 285, 127268, 2025.

[27] A. M. Campbell, C. Hauton, C. Baker-Austin, R. van Aerle, and J. Martinez-Urtaza, "Eco-evolutionary dynamics of climate-sensitive pathogens," 2023.

[28] W. G. Madden et al., "Deep neural networks for endemic measles dynamics," *PLoS Computational Biology*, 2024.

[29] R. U. Ginting, "Deep learning for COVID-19 transmission modeling in Medan," 2024.

[30] S. Towfek and M. Elkanzi, "AI/ML for predicting global epidemics and pandemics," 2024.

[31] M. M. Siddique et al., "Artificial Intelligence for Infectious Disease Detection: Prospects and Challenges," 2024.

[32] Y. Hu et al., "The challenges of deploying artificial intelligence models in a rapidly evolving pandemic," *Nature Machine Intelligence*, 2020.

[33] X. Xie, X. Li, Y. Sheng, and B. Gu, "Prediction of Antimicrobial Resistance Based on Random Forest Algorithms," 2019.

[34] J. J. Dabrowski, Y. F. Zhang, and A. Rahman, "ForecastNet: A Time-Variant Deep Feed-Forward Neural Network Architecture for Multi-step-Ahead Time-Series Forecasting," in *Lecture Notes in Computer Science*, vol. 12534, pp. 579–591, 2020.

[35] B. Murdoch, "Privacy and artificial intelligence: challenges for protecting health information in a new era," *BMC Medical Ethics*, vol. 22, no. 1, pp. 1–5, 2021.

[36] R. M. Ratwani, K. Sutton, and J. E. Galarraga, "Addressing AI Algorithmic Bias in Health Care," *JAMA*, vol. 332, no. 13, pp. 1051–1052, 2024.

[37] E. H. Allen, H. Samuel-Jakubos, and T. A. Waidmann, "Data Sharing in Cross-Sector Collaborations," Urban Institute, 2021.

[38] C. C. Westhues, H. Simianer, and T. M. Beissinger, "learnMET: an R package to apply machine learning methods for genomic prediction using multi-environment trial data," *G3 Genes|Genomes|Genetics*, vol. 12, no. 11, 2022.