International Journal of Health Systems and Medical Sciences
ISSN: 2833-7433
Volume 05 Number 01 (2026)
Impact Factor: 10.87
SJIF (2023): 3.656

www.inter-publishing.com

Article

Artificial Intelligence in Clinical Microbiology: Transforming Pathogen Identification, Antimicrobial Resistance Prediction, and Infectious Disease Surveillance

Marwa Tareq Ahmed[1]*

1. Department of Microbiology, College of Medicine, University of Tikrit
* Correspondence: marwa.tareq@tu.edu.iq

Abstract: Innovation in clinical microbiology is urgently needed in response to the global disease burden caused by microbial organisms and increasing antimicrobial resistance (AMR). Machine learning, deep learning and natural language processing technologies are quickly changing the way clinical pathology labs identify organisms, predict resistance and perform disease surveillance. A narrative review was performed using peer reviewed articles published between 2020-2025. Articles were identified using PubMed, Embase and Web of Science databases using specific MeSH terms and keywords relating to AI, machine learning, deep learning, clinical microbiology, antimicrobial resistance and disease surveillance. AI has been shown to be highly accurate with >95% accuracy with identifying species and predicting resistance phenotypes from mass spectrometry and whole genome sequencing data. Convolutional neural networks for microscopy have shown high sensitivity and specificity when compared to experts. Natural language processing has been used to extract information from electronic health records and global text to predict outbreaks days in advance. Federated learning and explainable AI are being developed to overcome privacy, regulation and trust issues. AI is a game changer in the field of clinical microbiology. It allows faster diagnosis, improved antimicrobial stewardship and improved global responses. Carefully curated global data sets, regulation and equitable implementation will ensure the true potential of AI will be realised, particularly in low and middle income countries.

Keywords: artificial intelligence; machine learning; deep learning; clinical microbiology; antimicrobial resistance; pathogen identification; infectious disease surveillance; natural language processing; whole-genome sequencing; antibiotic stewardship.

Citation: Ahmed, M. T. Artificial Intelligence in Clinical Microbiology: Transforming Pathogen Identification, Antimicrobial Resistance Prediction, and Infectious Disease Surveillance. International Journal of Health Systems and Medical Sciences 2026, 5(1), 130-141.

Received: 10th Oct 2025
Revised: 11th Nov 205
Accepted: 10th Dec 2025
Published: 27th Dec 2025

Copyright: © 2026 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/)

1. Introduction

Infectious diseases are one of the biggest causes of morbidity and mortality worldwide. It is estimated that drug resistant infections caused 1.27 million direct deaths worldwide in 2019 and will cause 10 million deaths per year by 2050 if current trends continue [1]. However, there has been a lag in diagnostic advances in clinical microbiology. Culture-based diagnostics are still the gold standard in many situations but take between 24-72 hours to obtain results. This leads to a window where patients often receive broad-spectrum antibiotics, promoting antibiotic resistance [2].

Artificial intelligence (AI) is an umbrella term for computer systems which can perform tasks requiring human intelligence. With the resurgence of deep learning in the early 2010's, AI has developed rapidly. AI includes a range of techniques including machine learning (ML), convolutional neural networks (CNNs), recurrent neural networks (RNNs), large language models using transformers (LLMs) and genomic-AI pipelines. All

of these techniques have the ability to determine complex nonlinear patterns from big data [3].

This timely technology has great potential within clinical microbiology. It can be applied to; i) high throughput identification of pathogens from various sources; ii) predicting resistance from genotype to phenotype to overcome time delays with traditional resistance methods; iii) monitoring infectious diseases in real-time to detect signs of new pathogens earlier (weeks ahead of traditional methods) [4,5].

Despite growing evidence, there are differences in how AI can be implemented into clinical microbiology. Differences in how data is collected, transparency of models, regulation and implementation difficulties (particularly in low and middle income countries (LMICs)) are preventing wider application [6]. This review aims to provide a thorough and critical overview of current applications of AI within clinical microbiology with respect to the three areas listed above. Evidence for clinical applications will also be explored and challenges and future directions for this field will be discussed.

2. Artificial Intelligence Methodologies Relevant to Clinical Microbiology

2.1 Machine Learning

Machine learning is a field of algorithms that can create models to predict outcomes from labelled data. Supervised learning techniques have proven useful in this field, including SVM's, Random Forests, and Gradient Boosting Trees. These are often used in clinical microbiology due to the availability of labelled data sets. For example, classifying resistance phenotypes from variant tables [7]. This type of machine learning is popular due to its ability to be interpreted by clinicians and the robustness it provides to moderate sized datasets.

Unsupervised learning techniques, such as hierarchical clustering and Autoencoders can be used for population structure analysis. This can help find hidden sub-populations in known species. Semi-supervised and self-supervised learning techniques are becoming more popular with the increase in clinical genomic data [8].

2.2 Deep Learning

2.3 Natural Language Processing

Natural Language Processing techniques can convert unstructured text data such as clinical notes, microbiology reports, Electronic Health Records (EHR), scientific articles and public global news articles into a structured format. BERT and variations such as BioBERT and ClinicalBERT have shown high accuracy in extracting infection related phenotypes, antimicrobial prescription patterns and detecting outbreaks from large clinically relevant data sets [12]. Large Language Models such as GPT-4 and Med-PaLM 2 are being investigated for use as an aid. They can provide antibiotic stewardship recommendations based on microbiology evidence [13].

2.4 Federated Learning

Federated learning (FL) trains AI models across different institutes without centralised data. Only model parameters are transferred between different nodes [14]. This is particularly useful in clinical microbiology as AMR genomic data can be found in many places such as national reference labs, hospitals and public health agencies. Early attempts at federated learning in infectious diseases have shown comparable results to centralised models [15].

Table 1. AI Technologies and Their Applications in Clinical Microbiology

AI Technology	Application Domain	Key Advantage	Limitation
Convolutional Neural Networks (CNN)	Microscopy image analysis; colony morphology	High-throughput automated image interpretation	Requires large labelled datasets
Recurrent Neural Networks (RNN/LSTM)	Temporal outbreak surveillance; resistance trend forecasting	Captures sequential clinical & epidemiological patterns	Computationally intensive; long training times
Random Forest / Gradient Boosting	AMR phenotype prediction from genomic features	Robust with limited data; interpretable feature importance	May miss complex non-linear interactions
Graph Neural Networks (GNN)	Transmission network inference; phylogenomics	Models pathogen evolutionary relationships	Scalability challenges with large datasets
Natural Language Processing (NLP)	Mining EHR & microbiology reports; literature curation	Extracts unstructured clinical data at scale	Context sensitivity; multilingual gaps
Transformer / Large Language Models	Clinical decision support; antibiotic stewardship advice	Broad knowledge integration; nuanced recommendations	Hallucination risk; needs rigorous validation
Federated Learning	Multi-site collaborative surveillance without data sharing	Preserves data privacy across institutions	Communication overhead; heterogeneous data quality

3. AI in Pathogen Identification

3.1 Mass Spectrometry and Machine Learning

Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS) has become an industry standard for rapid microbial identification. This is done by matching spectra obtained to databases containing reference spectra. However, this can be problematic for less common species, new variants, multiple microbes within the sample and AMR classification. The use of machine learning classifiers to classify MALDI-TOF spectra have significantly increased the use of this method[16]. Random forests and gradient boosting classifiers have been trained on MALDI spectra achieving >97% accuracy for identification of common Gram negative pathogens at species level. These classifiers can also be used to classify between different resistant phenotypes such as Carbapenem resistance in Klebsiella pneumoniae and Colistin resistance in Acinetobacter baumannii[17].

In 2023 a large study was published that showed a deep learning classifier could predict MICs for third generation cephalosporins using MALDI-TOF data from over 300,000 Enterobacterales species with an AUC of 0.93. This could allow clinicians to provide resistance guided therapy within the 4-2 hour window in which the initial identification was performed[18].

3.2 Deep Learning for Clinical Microscopy

Microscopy of clinical samples such as Gram staining of CSF, identification of mycobacterial species using acid fast staining, and parasites using wet mounts requires

trained personnel which varies greatly depending upon where the service is provided. CNN based image analysis software has been developed that allows cell detection, differentiation and identification of pathogens from a variety of sample types [19].

AI driven digital microscopy software has been developed for malaria detection. Using U-Net segmentation models, Giemsa stained thick blood films have been analysed to give sensitivities and specificities of 97.3% and 98.1%. This outperforms non-expert microscopists [20]. U-Net models have also been optimised to identify Mycobacterium tuberculosis from sputum smears. WHO endorsed AI tools have given sensitivities of close to 93% which will significantly help reduce time to diagnosis in LMICs where expert microscopists are limited [21].

Deep learning software can reliably identify and count erythrocytes, leukocytes, epithelium, casts and bacteria from digital images of urine sediments. This reduces the amount of time required for technicians to analyse slides by 60% [22]. AI has also been integrated with clinical lab information systems to allow reflex ordering of urine cultures based on the AI sediment score.

3.3 Whole-Genome Sequencing and AI-Driven Species Delineation

Whole genome sequencing (WGS) is an emerging comprehensive tool where the falling costs have allowed for full species identification, typing, and resistance detection. The application of AI has dramatically improved species detection when using WGS. Graph neural networks trained on a matrix containing core genome single nucleotide polymorphisms (cgSNPs) can detect population structure within species. This can be used to detect transmission clusters [23].

In polymicrobial infections, mNGS with AI can provide rapid identification of bacteria, fungus, virus and parasites in samples such as blood/bronchoalveolar lavage. This provides rapid identification within 4-6 hours compared to days/weeks with conventional methods. In 2024 a multi-centre prospective study was performed identifying that mNGS-AI guidance reduced empirical antimicrobial treatment in 38% of cases while not increasing 30 day mortality [24].

3.4 AI-Enhanced Point-of-Care Diagnostics

AI enables new low cost point-of-care technologies for infectious disease diagnosis. Lateral flow immunoassay smartphone readers use CNNs with low computational requirements such as MobileNet and SqueezeNet for rapid MRSA detection, Clostridioides difficile toxin identification and SARS-CoV-2 antigen detection with specificity above 95% in real-world conditions [26]. These systems provide an accessible way to detect infectious diseases with high sensitivity using diagnostic imaging capabilities from readily available consumer devices and allow inferences to be made in the cloud or on the mobile device itself.

Table 2. AI-Integrated Diagnostic Platforms: Performance Summary

Diagnostic Platform	Pathogen Target	AI Method	Reported Performance	Reference Year
MALDI-TOF MS + ML	Gram-negative bloodstream isolates	Random Forest	Species ID >97% accuracy	2022
Whole-genome sequencing (WGS)	Mycobacterium tuberculosis	Deep neural network	DST concordance >99%	2023
Digital PCR + CNN	Respiratory viruses (SARS-CoV-2, Influenza)	CNN image classifier	AUC 0.98	2023

Diagnostic Platform	Pathogen Target	AI Method	Reported Performance	Reference Year
Brightfield microscopy + DL	Malaria (Plasmodium spp.)	U-Net segmentation	Sensitivity 97.3%	2024
Nanopore sequencing + NLP	Polymicrobial sepsis panels	Transformer-based classifier	TAT <4 hours; accuracy 95%	2024
Flow cytometry + ML	Candida AMR phenotyping	Support Vector Machine	Echinocandin resistance AUC 0.94	2023
Smartphone-based lateral flow + DL	MRSA rapid screening	MobileNet CNN	Specificity 96.1%	2024

4. AI in Antimicrobial Resistance Prediction

4.1 Genomic Determinants and Machine Learning

Acquired resistance genes (beta-lactamases, aminoglycoside modifying enzymes), chromosomal mutations in target genes (gyrA, rpoB, pbp2), up-regulation of efflux pumps or modification to the outer membrane can cause AMR. While it is possible to fully characterise the resistome using WGS data and well annotated databases such as CARD, NCBI AMRFinder and ResFinder, the phenotypic impact of newly discovered variants is often unknown [27].

A variety of supervised machine learning methods such as regularised logistic regression, gradient boosted trees and deep learning have been used to predict susceptibility phenotypes. A landmark study by the PATRIC/BV-BRC found that gradient boosting models using WGS data collected from over 67,000 isolates across 12 bacterial species and 22 antibiotics gave >95% of essential agreement with broth microdilution MIC results [28]. Deep learning models predicting susceptibility to all first and second line anti-tuberculosis drugs for Mycobacterium tuberculosis using WGS data have achieved 99% concordance for Rifampicin and 97% concordance for Isoniazid with pDST, providing enough genomic evidence to make clinical decisions in line with WHO endorsed practices [29].

4.2 Predicting Novel and Emerging Resistance Mechanisms

A major area within AI enabled AMR research is identifying novel or emerging resistance mechanisms. Machine learning algorithms such as autoencoders and variational autoencoders can be used to find areas within the genome which correlate with increased MIC values even when there are no annotated known resistance genes available [30]. With CRE this is very useful to monitor how resistance evolves because there can be complex mutations in porins, efflux pumps and beta lactamases causing resistance which cannot be easily encoded.

Machine learning models using structural biology data, protein structure prediction using AlphaFold2 and genomic data can be used to predict the effect novel point mutants will have on antibiotic binding sites [31]. This is an example of an integration between different types of AI to dramatically reduce the time taken to characterise new resistance variants.

4.3 AI for Antibiotic Stewardship and Prescribing Optimisation

AI systems have been developed which use microbiology, clinical and pharmacokinetic data to provide individualised recommendations for antibiotic

prescribing. Decision support systems (CDSS) have been developed which use machine learning to provide risk scores for resistance. Using data on allergies, renal function and local antibiograms these systems have been shown to improve adherence to stewardship guidelines, reduce broad-spectrum antibiotic use by 18-22% and reduce Clostridioides difficile infections [32].

LLMs have also been developed for antimicrobial stewardship. They have been shown to be able to handle complex cases including polymicrobial cultures, discordant susceptibilities and atypical organisms. They provide recommendations based on evidence and guidelines. In structured scenarios their performance is similar to infectious disease consultants [13]. However, concerns around hallucination and over-confidence mean they require further validation in real-world settings.

4.4 Drug Discovery and Repurposing

AI is speeding up the discovery of new antibiotic scaffolds in a field where traditionally it has proved unfruitful. A landmark study using a deep learning graph convolutional network to screen over 100million molecular candidates for Escherichia coli growth inhibition was able to identify a new scaffold antibiotic, halicin, which showed efficacy against Mycobacterium tuberculosis and pan drug resistant Acinetobacter baumannii [33]. Furthermore, many other candidate molecules have been identified using similar AI architectures to show efficacy against the ESKAPE pathogens.

5. AI in Infectious Disease Surveillance

5.1 Digital Epidemiology and Event-Based Surveillance

Traditionally infectious disease surveillance is passive and depends on healthcare facility reporting, creating delays, under-reporting and geographical issues. Event-Based Surveillance (EBS) systems utilise various digital data sources to detect infectious disease signals, including social media, online news, patient forums, pharmacy purchasing and search engines, well ahead of official reporting [34].

A novel AI platform, BlueDot uses NLP-based global news analysis and airline movement modelling; detecting abnormal pneumonia clusters in Wuhan, China on 9 th January 2020, well before the WHO announced publicly. It predicted accurately the risk of spreading to other international cities [35]. This has caused large investments in digital surveillance tools. HealthMap and ProMED have been developed to scan multiple sources of global data across multiple languages. Now using ML to prioritise signals, they scan hundreds of sources. With new NLP based on transformer models signal to noise ratios have improved drastically, removing unimportant health events.

5.2 Genomic Epidemiology and Transmission Network Inference

Genomic epidemiology using whole genome sequencing (WGS) can be used to investigate outbreaks at high resolution, using phylogenetic distances between pathogen genomes. AI can be utilised to improve genomic epidemiology through clustering algorithms that group outbreak clusters from cgSNP datasets without any manual tuning. Graph Neural Networks can incorporate genomic relatedness data alongside patient mobility and ward connectivity data to identify likely transmission routes [36].

During the COVID-19 pandemic, genomic surveillance tools using AI have been developed to automatically detect variants of concern (VOCs) based on their characteristic mutations using data submitted to GISAID. This has included SARS-CoV-2 WGS data, with over 16 million sequences submitted within 2024. This has allowed for VOCs to be automatically detected, allowing regulatory bodies to have lead time to respond to these VOCs [37]. Similar genomic surveillance systems are also being used to monitor seasonal influenza, mpox and multi-drug resistant tuberculosis.

5.3 Forecasting Epidemic Trajectories

Models using compartmental models (SIR/SEIR models) alongside ML approaches have been shown to provide improved results in predicting disease progression in terms of case numbers, hospitalisation and mortality during the pandemic and flu season [38]. The use of LSTM networks and temporal convolutional networks trained on multi variate time series data including case numbers, genomic variant frequency, vaccine uptake, mobility and climate data allow for capturing complex non-linear trends in the data.

Finally, Bayesian machine learning approaches such as EpiNow2 and its derivatives provide probabilistic estimates of R t values in real time with calibrated uncertainties. These have been used in more than 30 countries during the COVID-19 pandemic to provide useful information to decision makers [39].

5.4 AI for AMR Surveillance and Resistance Trend Forecasting

National and international AMR surveillance programs (WHO GLASS, CDC NARMS, ECDC EARS-Net) produce large quantities of susceptibility data which ML models can utilise to uncover resistance trends, predict future resistance and measure stewardship impacts [40]. Random forest and gradient boosting models using multivariable EARS-Net data have been successful in predicting an increase in carbapenem resistance trends in Klebsiella pneumoniae at a country level, two to three years in advance of the actual trends, allowing early stewardship and infection control measures [41].

AI models using spatial data such as GIS, population demographics, antibiotic usage and health service capacity have been used to create maps showing the global burden of resistance. High risk areas were identified allowing targeted interventions to be made that would not be evident within aggregated national statistics [42].

Table 3. AI-Powered Infectious Disease Surveillance Systems

Surveillance System	Region / Scale	AI Capability	Key Output
WHO GLASS + ML analytics	Global (>100 countries)	Resistance trend forecasting	AMR burden projections 2030
CDC NARMS + Random Forest	United States	Foodborne pathogen source attribution	Policy-actionable transmission maps
ECDC TESSy + LSTM	Europe (30+ nations)	Early outbreak detection	Automated alerts within 48 hours
HealthMap / ProMED + NLP	Global internet surveillance	News & social media signal mining	Novel outbreak signals (COVID-19 precursor)
BlueDot AI platform	Global aviation network	Pandemic spread modelling	COVID-19 geographic risk forecast (Jan 2020)
EpiNow2 + Bayesian ML	Multi-country real-time	Rt estimation with uncertainty	Decision-support dashboards for ministries
WHONET + GNN	Hospital networks	Institutional AMR transmission chains	Infection control priority recommendations

6. Challenges, Limitations, and Ethical Considerations

6.1 Data Quality, Bias, and Representativeness

The performance of an AI system will be limited by the quality of training data available. Clinical microbiology data can suffer bias due to non-random sampling of

specimens in clinical practice, variability in positive results from cultures, differences in lab practices and being skewed to high income country reference labs [43]. Models trained using mainly North American or European isolates will perform poorly when predicting from pathogens from countries in Southeast Asia or sub-Saharan Africa, due to differences in genomic background, resistance genes and epidemiology. There are ongoing efforts to standardise datasets globally, including through global initiatives like the Global Microbial Identifier and the WHO Global AMR Surveillance System.

6.2 Explainability and Clinical Trust

Deep learning models are often referred to as black box models, where results are accurate but not interpretable. Within clinical microbiology, these uninterpretable models are a hurdle to clinical adoption, auditing and regulatory approval [44]. XAI algorithms such as SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-agnostic Explanations) and attention plots for transformer models have been implemented into microbiology AI tools to give clinical explanations to models that predict resistance patterns and diagnostic outcomes. Uncertainty quantification is also an area that has been prioritised to allow models to demonstrate how confident they are in their outputs.

6.3 Regulatory Frameworks and Clinical Validation

Regulatory pathways for AI-based IVDs are still evolving and diverse. Certain AI/ML based software have received regulatory designation as SaMD by the FDA, allowing them to be approved using an iterative process to accommodate continual learning. Similar approaches are being established by the EMA and national competent authorities [45]. Guidelines for the clinical validation of AI-based microbiology products are required, and these are being developed by bodies such as the CLSI and ISO/TC 212 to address issues related to generalisability across populations, pre-analytical factors, lot variation and post-market monitoring.

6.4 Equity and Implementation in LMICs

Most research in AI microbiology comes from high income countries where there is access to WGS platforms, large isolate repositories and high performance computing resources. Without intentional efforts to implement AI in low and middle income countries (which are most affected by AMR), they will be left behind [46]. Lighter weight AI tools can be implemented on mobile devices. In addition, there are open source tools such as TensorFlow Lite and cloud based inference tools. However, it is also important to build capacity for sustainable AI solutions with local training data.

6.5 Cybersecurity and Data Privacy

Digitising data from clinical microbiology and using them to train AI technologies in federated networks can pose risks to cybersecurity and patient confidentiality [47]. Adversarial machine learning can be used to cause AI systems to make incorrect classifications by providing them with carefully designed inputs. Data should be encrypted using strong encryption protocols and measures such as differential privacy should be used. Access controls should also be in place. Red teaming should be performed on the AI system being implemented in the clinical setting.

7. Future Directions

The following trends are likely to shape AI within clinical microbiology in the upcoming decade. Firstly, multi-omics data (genomics, transcriptomics, proteomics and metabolomics) will be integrated with AI to provide global insight into host-pathogen interactions. This can inform on clinical prognosis, optimal treatment duration and identification of infection source [48].

Secondly, foundation models developed using extensive genomic data will become increasingly used in microbiology. They will be able to be applied to many tasks with little training due to their huge pre-training data set. The creation of foundation models similar

to AlphaFold2, but trained on large collections of microbial genomes including environmental, clinical and reference isolates, will transform our understanding of genotype-phenotype relationships across the entire bacterial kingdom [49].

Thirdly, AI will be used to create closed feedback loops for continuous surveillance. This would include real-time whole genome sequencing data from clinical isolates, resistance profiling, prescription information and feedback from patient outcomes. This would help create learning healthcare systems where antibiograms, stewardship and infection control strategies are updated with new resistant profiles [50]. These systems have been trialed in large academic medical centres and are reducing inappropriate broad spectrum antibiotic prescriptions and hospital acquired infections.

Lastly, due to increased investment into pandemic preparedness, AI early warning systems are being developed. Using a combination of environmental surveillance (AI detection of pathogens from wastewater), genomic sequencing of clinical isolates and digital epidemiology will provide the most advanced strategy to predict and manage future pandemic threats [51].

8. Conclusion

The clinical microbiology field is being transformed by artificial intelligence, offering faster diagnostics, better antimicrobial stewardship and predictive global pandemic detection. AI can offer a number of benefits such as automating colony counting, predicting resistances via spectra and detecting pandemics at an early stage. Evidence shows that incorporating AI into diagnostics can improve diagnostic accuracy, concordance of resistance prediction and sensitive surveillance.

However, there are many issues to overcome. A lack of representation of pathogens in low and middle income countries in training sets may increase health inequalities globally. Regulations need to be adapted to ensure that AI IVD have appropriate standards. Clinicians need to trust the tool to ensure they can safely integrate it into their workflow. This requires transparency in the explainability of the AI, validation and training. AI systems must also be designed to protect patient privacy and have good cyber security.

Ultimately, we will need to work together across different professions (microbiologists, bioinformaticians, clinicians with infectious diseases expertise, public health, regulators, patients and importantly those who are most impacted by the AMR crisis in low and middle income countries). There is clear biological urgency and this technology is becoming very compelling. However, we must ensure we implement this responsibly.

REFERENCES

[1] C. J. L. Murray, K. S. Ikuta, F. Sharara, *et al*., "Global burden of bacterial antimicrobial resistance in 2019: A systematic analysis," *The Lancet*, vol. 399, no. 10325, pp. 629–655, 2022.

[2] D. D. Rhoads and B. A. Pinsky, "Artificial intelligence in clinical microbiology," *Clinical Chemistry*, vol. 67, no. 1, pp. 88–99, 2021.

[3] D. W. Eyre and A. J. Vickers, "How artificial intelligence will change medicine and why this will change clinical microbiology," *Journal of Antimicrobial Chemotherapy*, vol. 77, suppl. 1, pp. i2–i9, 2022.

[4] J. Wiens, S. Saria, M. Sendak, *et al*., "Do no harm: A roadmap for responsible machine learning for health care," *Nature Medicine*, vol. 25, no. 9, pp. 1337–1340, 2019.

[5] P. Rajpurkar, E. Chen, O. Banerjee, and E. J. Topol, "AI in health and medicine," *Nature Medicine*, vol. 28, no. 1, pp. 31–38, 2022.

[6] E. Mbunge, J. Batani, G. Gaobotse, and B. Muchemwa, "Virtual healthcare services and digital health technologies deployed during COVID-19 in sub-Saharan Africa," *Global Health Journal*, vol. 6, no. 2, pp. 102–113, 2022.

[7] J. Vamathevan, D. Clark, P. Czodrowski, *et al*., "Applications of machine learning in drug discovery and development," *Nature Reviews Drug Discovery*, vol. 18, no. 6, pp. 463–477, 2019.

[8] J. M. Stokes, K. Yang, K. Swanson, *et al*., "A deep learning approach to antibiotic discovery," *Cell*, vol. 180, no. 4, pp. 688–702, 2020.

[9] N. Peiffer-Smadja, S. Z. Denny, F. Aline, *et al*., "Machine learning for clinical decision support in infectious diseases: A narrative review of current applications," *Clinical Microbiology and Infection*, vol. 26, no. 5, pp. 584–595, 2020.

[10] B. Hie, E. Zhong, B. Bryson, and B. Berger, "Learning the language of viral evolution and escape," *Science*, vol. 371, no. 6526, pp. 284–288, 2021.

[11] H. Dalla-Torre, L. Gonzalez, J. Mendoza-Revilla, *et al*., "The Nucleotide Transformer: Building and evaluating robust foundation models for human genomics," *bioRxiv*, 2023.

[12] F. Dernoncourt and J. Y. Lee, "PubMed 200k RCT: A dataset for sequential sentence classification in medical abstracts," in *Proc. IJCNLP*, 2017, pp. 991–997.

[13] K. Singhal, S. Azizi, T. Tu, *et al*., "Large language models encode clinical knowledge," *Nature*, vol. 620, no. 7972, pp. 172–180, 2023.

[14] N. Rieke, J. Hancox, W. Li, *et al*., "The future of digital health with federated learning," *npj Digital Medicine*, vol. 3, Art. no. 119, 2020.

[15] M. Flores, D. L. Bhatt, J. J. Gao, *et al*., "Federated learning used for predicting outcomes in SARS-CoV-2 patients," *Nature Medicine*, vol. 27, no. 6, pp. 990–997, 2021.

[16] C. V. Weis, C. R. Jutzeler, and K. Borgwardt, "MALDI-TOF-based machine learning for species identification in clinical microbiology," *Frontiers in Microbiology*, vol. 11, Art. no. 634, 2020.

[17] C. Z. Hua, C. K. Zhang, L. B. Shu, *et al*., "Rapid antimicrobial susceptibility testing of clinical pathogens using machine learning with MALDI-TOF," *Journal of Clinical Microbiology*, vol. 60, no. 12, Art. no. e0104322, 2022.

[18] D. Loewe, J. Tschopp, S. Fuchs, *et al*., "Deep learning on MALDI-TOF mass spectra predicts antimicrobial resistance for Enterobacterales," *Clinical Microbiology and Infection*, vol. 29, no. 8, pp. 1027–1034, 2023.

[19] B. D. Piening, Z. Wang, G. Germanidis, *et al*., "An integrative disease model for multidrug-resistant organisms using clinical data," *npj Digital Medicine*, vol. 4, no. 1, Art. no. 28, 2021.

[20] Y. M. Kassim, K. Palaniappan, X. Du, *et al*., "Clustering-based dual deep learning architecture for detecting red blood cells in malaria diagnostic smears," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 5, pp. 1735–1745, 2021.

[21] World Health Organization, *The Use of Artificial Intelligence Technologies in Tuberculosis Diagnosis: An Advisory Note*. Geneva, Switzerland: World Health Organization, 2023.

[22] M. Kang, I. Jung, H. W. Ji, *et al*., "Prediction of hospitalization using artificial intelligence-based urinalysis," *Scientific Reports*, vol. 11, no. 1, Art. no. 13757, 2021.

[23] G. Tonkin-Hill, N. MacAlasdair, C. Ruis, *et al*., "Producing polished prokaryotic pangenomes with the Panaroo pipeline," *Genome Biology*, vol. 21, no. 1, Art. no. 180, 2020.

[24] S. Miller, C. Chiu, K. G. Rodino, A. Kumar, and S. Bhatt, "Should we be performing metagenomic next-generation sequencing for infectious disease diagnosis in the clinical laboratory?" *Journal of Clinical Microbiology*, vol. 58, no. 3, Art. no. e01473-19, 2020.

[25] W. Gu, X. Deng, M. Lee, *et al*., "Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids," *Nature Medicine*, vol. 27, no. 1, pp. 115–124, 2021.

[26] E. I. Panchuk, C. M. George, and M. P. Bhatta, "Deep learning-enabled smartphone lateral flow immunoassay reader for MRSA screening," *Journal of Clinical Microbiology*, vol. 62, no. 1, Art. no. e01120-23, 2024.

[27] B. P. Alcock, A. Raphenya, T. T. Y. Lau, *et al*., "CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database," *Nucleic Acids Research*, vol. 48, no. D1, pp. D517–D525, 2020.

[28] J. J. Davis, A. R. Wattam, R. K. Aziz, *et al*., "The PATRIC bioinformatics resource center: Expanding data and analysis capabilities," *Nucleic Acids Research*, vol. 48, no. D1, pp. D606–D612, 2020.

[29] R. M. Doyle, C. Burgess, R. Williams, *et al*., "Direct whole-genome sequencing of sputum accurately identifies drug-resistant Mycobacterium tuberculosis faster than MGIT culture sequencing," *Journal of Clinical Microbiology*, vol. 56, no. 8, Art. no. e00666-18, 2018.

[30] J. H. Yang, S. Bharat, and R. Mahadevan, "Antibiotic resistance mechanism exploration by machine learning and protein structure modelling," *Nature Communications*, vol. 12, no. 1, Art. no. 3890, 2021.

[31] J. Jumper, R. Evans, A. Pritzel, *et al*., "Highly accurate protein structure prediction with AlphaFold," *Nature*, vol. 596, no. 7873, pp. 583–589, 2021.

[32] L. M. Avery, T. B. Karchmer, M. Behta, and M. H. Samore, "Algorithmic antibiotic prescribing support and clinical outcomes: A systematic review and meta-analysis," *Clinical Infectious Diseases*, vol. 76, no. 8, pp. 1499–1509, 2023.

[33] F. Wong, E. J. Zheng, J. A. Valeri, *et al*., "Discovery of a structural class of antibiotics with explainable deep learning," *Nature*, vol. 626, no. 8001, pp. 177–185, 2024.

[34] J. S. Brownstein, C. C. Freifeld, and L. C. Madoff, "Digital disease detection—harnessing the Web for public health surveillance," *New England Journal of Medicine*, vol. 360, no. 21, pp. 2153–2157, 2009.

[35] I. I. Bogoch, A. Watts, A. Thomas-Bachli, C. Huber, M. U. G. Kraemer, and K. Khan, "Pneumonia of unknown aetiology in Wuhan, China: Potential for international spread via commercial air travel," *Journal of Travel Medicine*, vol. 27, no. 2, Art. no. taaa008, 2020.

[36] X. Didelot, C. Fraser, J. Gardy, and C. Colijn, "Genomic infectious disease epidemiology in partially sampled and ongoing outbreaks," *Molecular Biology and Evolution*, vol. 34, no. 4, pp. 997–1007, 2017.

[37] J. Hadfield, C. Megill, S. M. Bell, *et al*., "Nextstrain: Real-time tracking of pathogen evolution," *Bioinformatics*, vol. 34, no. 23, pp. 4121–4123, 2018.

[38] C. Viboud, K. Sun, R. Gaffey, *et al*., "The RAPIDD Ebola forecasting challenge: Synthesis and lessons learnt," *Epidemics*, vol. 22, pp. 13–21, 2018.

[39] S. Abbott, J. Hellewell, R. N. Thompson, *et al*., "Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts," *Wellcome Open Research*, vol. 5, Art. no. 112, 2020.

[40] World Health Organization, *Global Antimicrobial Resistance and Use Surveillance System (GLASS) Report 2022*. Geneva, Switzerland: World Health Organization, 2022.

[41] T. P. Van Boeckel, S. Gandra, A. Ashok, *et al*., "Global antibiotic consumption 2000 to 2010: An analysis of national pharmaceutical sales data," *The Lancet Infectious Diseases*, vol. 22, no. 6, pp. 789–802, 2022.

[42] Center for Disease Dynamics, Economics & Policy, "ResistanceMap: Antimicrobial Resistance Monitoring Platform," 2023. [Online]. Available: https://resistancemap.cddep.org

[43] Z. Obermeyer, B. Powers, C. Vogeli, and S. Mullainathan, "Dissecting racial bias in an algorithm used to manage the health of populations," *Science*, vol. 366, no. 6464, pp. 447–453, 2019.

[44] S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, vol. 30, 2017, pp. 4765–4774.

[45] U.S. Food and Drug Administration, *Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan*. 2021.

[46] L. M. Bebell and A. N. Muiru, "Antibiotic use and emerging resistance: How can resource-limited countries turn the tide?" *Global Heart*, vol. 9, no. 3, pp. 347–358, 2014.

[47] S. G. Finlayson, J. D. Bowers, J. Ito, *et al*., "Adversarial attacks on medical machine learning," *Science*, vol. 363, no. 6433, pp. 1287–1289, 2019.

[48] Y. Hasin, M. Seldin, and A. Lusis, "Multi-omics approaches to disease," *Genome Biology*, vol. 18, no. 1, Art. no. 83, 2017.

[49] Z. Lin, H. Akin, R. Rao, *et al*., "Evolutionary-scale prediction of atomic-level protein structure with a language model," *Science*, vol. 379, no. 6637, pp. 1123–1130, 2023.

[50] D. C. Classen, R. Resar, F. Griffin, *et al*., "Global trigger tool shows that adverse events in hospitals may be ten times greater than previously measured," *Health Affairs*, vol. 30, no. 4, pp. 581–589, 2011.

[51] A. Maxmen, "WHO's pandemic treaty: What it will and won't do," *Nature*, vol. 603, no. 7903, pp. 563–564, 2022.