



OPEN ACCESS

EDITED BY

Margaret Ip,
The Chinese University of Hong Kong, China

REVIEWED BY

Eustachio Cuscianna,
University of Bari Aldo Moro, Italy
Virna Maria Stavros Tsitou,
Medical University Sofia, Bulgaria

*CORRESPONDENCE

Alberto Rizzo
✉ rizzo.alberto@asst-fbf-sacco.it

RECEIVED 29 November 2025

REVISED 31 December 2025

ACCEPTED 13 January 2026

PUBLISHED 29 January 2026

CORRECTED 14 April 2026

CITATION

Scaglione G, Mastroianni N, Rizzo A, Palomba E, Carcione D, Rimoldi SG, Brigante G, Principe L, Colaneri M, Gori A and Borgonovo F (2026) Integrating artificial intelligence with genome sequencing against antimicrobial resistance: a narrative review. *Front. Public Health* 14:1757161. doi: 10.3389/fpubh.2026.1757161

COPYRIGHT

© 2026 Scaglione, Mastroianni, Rizzo, Palomba, Carcione, Rimoldi, Brigante, Principe, Colaneri, Gori and Borgonovo. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Integrating artificial intelligence with genome sequencing against antimicrobial resistance: a narrative review

Giovanni Scaglione¹, Nicolò Mastroianni², Alberto Rizzo^{3*}, Emanuele Palomba^{1,4}, Davide Carcione², Sara Giordana Rimoldi³, Gioconda Brigante², Luigi Principe⁵, Marta Colaneri^{4,6}, Andrea Gori^{1,4,6} and Fabio Borgonovo¹

¹Department of Infectious Diseases, Luigi Sacco Hospital, Milan, Italy, ²Laboratory of Clinical Microbiology and Virology, ASST Valle Olona, Gallarate, Italy, ³Laboratory of Clinical Microbiology, Virology and Bioemergencies, Luigi Sacco Hospital, Milan, Italy, ⁴Centre for Multidisciplinary Research in Health Science (MACH), University of Milan, Milan, Italy, ⁵Clinical Microbiology and Virology Unit, Great Metropolitan Hospital "Bianchi-Melacrino-Morelli", Reggio Calabria, Italy, ⁶Department of Biomedical and Clinical Sciences "L. Sacco", University of Milan, Milan, Italy

Antimicrobial resistance (AMR) represents an escalating global health threat, demanding diagnostic strategies capable of rapid, accurate, and comprehensive pathogen characterization. Genomic sequencing has transformed our ability to elucidate resistance mechanisms and track their evolution, yet its routine clinical adoption remains limited by cost, workflow constraints, and extended turnaround times. This narrative review examines how artificial intelligence (AI) and machine learning (ML) can enhance and operationalize sequencing-based diagnostics across the clinical microbiology continuum. We summarize current AI applications in whole-genome sequencing for AMR prediction, pan-genome feature extraction, and multicenter model generalizability, including emerging approaches such as federated learning. We then explore AI-driven metagenomic analytics for pathogen detection, resistome profiling, outbreak investigation, and prognostic modeling. Complementary non-genomic technologies, Raman spectroscopy and MALDI-TOF MS, are also evaluated for their potential to deliver rapid resistance profiling when integrated with ML. Finally, we discuss practical barriers, including cost, dataset standardization, interpretability, and regulatory challenges, while outlining future directions toward scalable, explainable, and equitable AI-guided diagnostics. Integrating AI with genomic and rapid phenotypic tools offers a pathway to real-time surveillance, optimized antimicrobial stewardship, and strengthened preparedness against emerging infectious threats.

KEYWORDS

antimicrobial resistance, artificial intelligence, diagnosis, genome sequencing, infection, machine learning, surveillance

1 Introduction

Antimicrobial resistance (AMR) is rapidly becoming a major global health threat with the World Bank estimating that AMR could result in US\$ 1 trillion additional healthcare costs by 2050, and US\$ 1–3.4 trillion gross domestic product losses per year by 2030. It has been reported that poor access to quality diagnostics is one of the main contributing factors, among others (1). The advent of genomic sequencing has significantly enhanced our ability to

understand the genetic mechanisms behind resistance, track their evolution over time, and accurately monitor the spread of resistant strains.

However, despite its analytical power, the integration of genome sequencing into routine clinical workflows remains limited. Barriers such as cost, personnel training requirements, and longer turnaround time (TAT) compared to traditional phenotypic testing or polymerase chain reaction (PCR) assays continue to delay its broader adoption (2).

In parallel, artificial intelligence (AI) has been proposed as a revolutionizing tool for laboratory sample processing and data management. Specifically, machine learning (ML) systems, with the ability to learn and improve from training datasets, are emerging with a central role in clinical microbiology (3). Beyond ML, more sophisticated AI frameworks, including convolutional neural networks (CNNs) and large language models (LLMs) are being investigated for their ability to handle complex, high-dimensional datasets and generate clinically relevant outputs (4–6).

Recent studies suggest that AI may play an integrative role across the full continuum of pathogen genomics, from raw data processing to the generation of actionable insights, potentially addressing several obstacles that limit the scalability of sequencing-based diagnostics (7). Here, we examine how AI-driven approaches can enhance genomic surveillance, resistance prediction, and the broader fight against AMR when strategically combined with next-generation sequencing (NGS) and complementary technologies.

2 Approach and scope of the narrative review

This narrative review was conducted to synthesize and critically appraise current evidence on the integration of AI with genomic and

rapid diagnostic technologies for AMR detection and surveillance in hospital and public health settings. Abbreviations, and the corresponding brief definitions, of all AI/ML-related terms used in this manuscript are provided in Table 1. The review was informed by a predefined conceptual PI(C)O framework (Supplementary Table S1), which served as a scoping scaffold to define the population, intervention domains, and outcomes of interest, rather than as a formal systematic or scoping review protocol. The literature search was conducted across three major biomedical databases: MEDLINE (via Ovid), Embase, and the Cochrane Central Register of Controlled Trials (CENTRAL). Searches included publications up to April 2025, with particular emphasis on studies published within the last decade, reflecting the rapid evolution of AI methodologies and sequencing technologies. Search concepts combined terms related to AI and ML (e.g., “AI,” “machine learning,” “deep learning,” “natural language processing”), genomic and rapid diagnostic technologies (e.g., “next-generation sequencing,” “whole-genome sequencing,” “metagenomics,” “nanopore sequencing,” “MALDI-TOF MS,” “Raman spectroscopy”), AMR, and surveillance or infection prevention workflows.

The scope of this review is intentionally restricted to AI applications integrated with genomic and rapid diagnostic technologies for AMR detection, characterization, and surveillance. It does not aim to provide a comprehensive review of AI tools focused primarily on patient-level MDRO carrier prediction, hospital logistics, or infection control management, which fall outside the focus of the present manuscript.

The review prioritized studies relevant to human health, particularly those addressing hospital or inpatient settings, population-level surveillance, infection prevention and control (IPC), and antimicrobial stewardship. One Health or environmental studies were included when they directly informed surveillance frameworks, transmission dynamics, or public health

TABLE 1 Abbreviations and brief definitions of AI/ML-related terms used in the manuscript.

Abbreviation	Extended name	Brief definition
AI	Artificial Intelligence	Umbrella term for computational systems designed to perform tasks requiring human-like intelligence, including learning, reasoning, and decision-making.
ML	Machine Learning	Subfield of AI in which algorithms learn patterns from data to make predictions or decisions without explicit rule-based programming.
DL	Deep Learning	Subset of ML using multi-layered neural networks to model complex, high-dimensional relationships in data.
CNN	Convolutional Neural Network	DL architecture optimized for grid-like data (e.g., spectra, signals, images), using convolutional filters to learn local patterns.
ANN	Artificial Neural Network	ML model inspired by biological neurons, consisting of interconnected layers that transform inputs into predictive outputs.
DT	Decision Tree	Supervised ML model that recursively splits data using decision rules to predict outcomes.
RF	Random Forest	Ensemble ML method combining multiple decision trees trained on random feature subsets to improve predictive accuracy and robustness.
SVM	Support Vector Machine	Supervised ML algorithm that finds optimal separating hyperplanes to classify or regress high-dimensional data.
LR	Logistic Regression	Supervised statistical-ML model estimating the probability of a binary outcome using a logistic function.
XAI	Explainable Artificial Intelligence	Methods that provide transparent, interpretable explanations of AI model predictions to support trust and auditability.
k-mer	k-length nucleotide substring	Fixed-length sequence fragments used as alignment-free genomic features in ML models.

implementation. We explicitly excluded studies limited to: (i) AI applications focused solely on antimicrobial prescribing optimization or laboratory automation without infection control relevance, and (ii) theoretical models, simulations, or proof-of-concept frameworks lacking a clear pathway to real-world application.

Priority was given to multicenter studies, externally validated models, real-world or near-real-world implementations, and consensus documents or guidelines where available. Single-center and early proof-of-concept studies were included selectively to illustrate emerging approaches, but were interpreted cautiously with respect to generalizability, validation, and readiness for implementation. Up-to-date narrative reviews were included to enhance the completeness of the literature coverage. During manuscript revision, nine relevant articles published after April 2025 were identified and incorporated as substantive updates in the field, and four studies initially excluded were subsequently included following re-evaluation of their relevance.

This review is intentionally narrative and non-exhaustive. No formal systematic review methodology, risk-of-bias assessment, or quantitative synthesis was undertaken. The objective was to provide a structured and critical overview of key AI-enabled genomic and rapid diagnostic applications for AMR surveillance, highlighting methodological limitations, operational constraints, and implications for public health practice.

3 Artificial intelligence and genome sequencing

The critical threat of AMR is prompting the need for advanced computational methods capable of rapid and precise resistance prediction. Recent developments in ML leverage genomic data, pan-genome analyses, and ensemble techniques to enhance detection accuracy. Nevertheless, challenges persist in feature selection, model generalizability, and validation across heterogeneous datasets.

A foundational step in AMR prediction is identifying informative genomic features. Traditional models often rely on known resistance genes or mutations, which may overlook novel or combinatorial signals. Her and Wu demonstrated that pan-genome-based feature extraction improved prediction accuracy in *Escherichia coli* by capturing accessory genome elements linked to strain-specific resistance mechanisms (8). Notably, non-core genes, present only in subsets of strains, outperformed core-genome features, even though not all directly conferred resistance. Many of these genes, in fact, often carried on mobile elements like plasmids, contributed to broader multidrug resistance signals. The proposed AI-based genetic algorithm approach based on non-core genes identified informative clusters that discriminated resistant from susceptible *E. coli*, achieving area under the curve (AUCs) above 0.9 for ampicillin, gentamicin, trimethoprim-sulfamethoxazole, and ciprofloxacin, surpassing conventionally examined clusters. Notably, the model flagged the *pmrC/pmrE* cluster (classically associated with polymyxin resistance) as predictive not only of polymyxin resistance but also of ampicillin and trimethoprim-sulfamethoxazole resistance, underscoring the capacity of ML to reveal latent resistance patterns.

Similarly, Wang et al. (9) showed that combining ML with whole-genome sequencing (WGS) predicted *Staphylococcus aureus* minimum inhibitory concentrations with >95% accuracy, particularly excelling

in detecting cefoxitin resistance and reliably identifying methicillin-resistant *S. aureus* (MRSA) strains.

In this context Hyun et al. (10) developed a pan-genome framework for *S. aureus*, *P. aeruginosa*, and *Escherichia coli*, creating a binary genome-by-feature matrix to train multiple support vector machines (SVM) models. Their ensemble approach, based on randomly sampled strain-feature subsets, achieved accuracies between 79.3 and 99.5%, outperforming classical statistical and ranking methods. Impressively, the model recovered 45 known resistance genes and suggested 25 novel candidates (10). By training every sub model on a randomly selected part of the dataset, they reduced overfitting and improved model generalizability.

Álvarez et al. (11), for instance, applied SVM to 5,299 *A. baumannii* genomes, including 512 global clone 1 (GC1) and 4,787 non-GC1, and identified a previously unknown 367 base pair (bp) fragment of *moaCB*, termed U1, uniquely defining the high-risk GC1. The model achieved notable performance (sensitivity 1.00, specificity 1.00, accuracy 1.00; 0 false positives and 0 false negatives), illustrating how ML can reveal clinically actionable biomarkers and directly support rapid diagnostic development in outbreak scenarios (11).

Most studies validate on single datasets, risking inflated performance estimates. When this is not the case, model performance is not consistent across different settings, and often fails to replicate across hospitals, likely due to differences in sequencing platforms, population structure, and metadata quality. On the other hand, while multicenter datasets enhance robustness, they also introduce new biases from heterogeneous lab protocols. Federated learning presents a promising alternative, allowing multi-institutional training without sharing raw genomic data. This approach can balance generalizability with privacy, laying the foundation for regional ML-based AMR early-warning systems (12).

Future models may benefit from integrating multi-omics layers, such as transcriptomics or proteomics, to uncover complex interactions or cryptic features that lead to AMR.

4 Artificial intelligence and metagenomics

Thanks to rapid advancements in sequencing technologies, metagenomics has the potential to provide a detailed view of the entire microbial ecosystem within clinical and environmental samples. Metagenomics provides a panoramic, high-resolution snapshot of the entire microbial genomic landscape, representing a substantial advance in microbiological diagnostics by moving beyond the targeted detection of individual pathogens and capturing comprehensive genomic information even in the absence of culture growth.

For example, Fida et al. (13) observed that 16S rRNA gene metagenomics was able to identify a potentially pathogenic organism in 47% of septic patients compared to 32% with blood cultures. As sequencing technologies continue to advance, this approach may potentially bypass traditional methods in the future, beginning with direct nucleic acid extraction from the collected specimen.

These technologies, like shotgun metagenomic sequencing, generate massive volumes of raw genetic data. Therefore, scaling this approach requires sequencing and data analysis technologies capable of managing the complexity and volume of information generated (14).

Hence, AI-based data management could be of paramount importance since the task of assembling millions of short reads into a coherent biological narrative is expected to be very time consuming without advanced computational pipelines. AI algorithms are used not only to align reads with reference databases, but more importantly, to perform deeper reconstruction and inference tasks. ML algorithms, specifically deep learning (DL) methods like CNN, could address several aspects of metagenomics sequencing, including microbial population analysis, novel pathogen detection, sequence classification, patient stratification and disease prediction (15).

Additionally, they have been successfully applied to functional annotation, generating comprehensive taxonomic profiles, classifying whether a given sequence is an antibiotic resistance gene, identifying all resistance determinants that together compose the sample's resistome, and detecting virulence genes or metabolic biomarkers (16, 17).

This broader, systems-level, insight can redefine how antimicrobial resistance is understood and monitored. While conventional methods profile resistance at the level of individual isolates, metagenomics unveils the resistome embedded within a sample's microbiota, including silent or subdominant reservoirs of resistance that may be mobilized under antibiotic pressure (18). This capacity is especially valuable for colonization screening, where characterizing the full spectrum of resistance genes enables more informed infection control and outbreak prevention strategies.

In a 2025 study, Mikolas et al. (19), through shotgun sequencing on Illumina platform, applied high-throughput metagenomic analysis to 96 intensive care unit (ICU) patient, staff, and environmental samples, aiming to identify early microbial biomarkers for prognostic assessment and uncovered extensive AMR reservoirs. Their computational models identified early oropharyngeal resistome signatures that discriminated early- from late-mortality patients with strong prognostic performance (AUC up to 0.90), illustrating how advanced metagenomic analytics can extract clinically actionable signals from complex datasets to support patient-centered clinical management strategies in settings where rapid clinical deterioration is frequent (19).

A promising application of AI to metagenomics is represented by the ability to differentiate between viable and dead microorganisms, addressing a common challenge in modern microbiology where molecular positivity cannot distinguish live from dead organisms. In standardized laboratory conditions, Ürel et al. (20) developed a model capable of detecting the viability of bacteria with an accuracy of 0.83 and an AUC value of 0.90 through the analysis of shotgun nanopore sequencing data.

Yet, the promise of this panoramic vision must be tempered by current limitations. The perceived high costs of sequencing, coupled with the need for robust bioinformatic infrastructure and expertise, currently restrict metagenomics to high-stakes clinical scenarios, such as culture-negative infections (e.g., endocarditis), intraoperative diagnostics, or precision outbreak tracking.

However, it should be noted that in a 2021 cost-effectiveness simulation model of a 32-month carbapenem-resistant *A. baumannii* (CRAB) outbreak, early use of WGS and shotgun metagenomics was predicted to avert 18 additional infections, yield 74 extra quality-adjusted life years (QALYs), and reduce overall hospital expenditures by AU\$93,822 compared with the real-life outbreak management scenario, highlighting the potential clinical and economic value of

sequencing-guided infection control (21). Then, the integration into routine, high-throughput clinical microbiology should move from aspirational to practical.

5 AI-integrated sequencing for real-time genomic surveillance

During and after SARS-CoV-2-caused COVID-19 pandemic, pathogen WGS shifted from reduced capacity to surveillance backbone, enabling rapid lineage tracking, transmission mapping, and detection of emerging resistance across clinical and public-health settings. Consensus guidance now frames genomic surveillance as essential for infection control and AMR monitoring across One-Health domains (22).

The term “real-time” should not be interpreted as “instantaneous.” It should refer to a data flow that makes information timely for effective infection control decisions and public health policies. Real-time genomic surveillance should refer to the continuous process of monitoring and analyzing the genetic material of pathogens and AMR-associated distribution, as they circulate in a population, with the goal of providing actionable insights to control outbreaks as soon as possible as they occur.

Unlike traditional surveillance, which often relies on retrospective data, real-time surveillance should aim for a fast “sample-to-result” TAT, typically within days or weeks, to allow public health officials to intervene during an active transmission event rather than simply documenting it after the event.

Real-time genomic surveillance has been proposed to anticipate the emergence of resistant strains, track their spread, link human, animal, and environmental data, and support rapid response.

As partially anticipated, currently, two complementary sequencing paradigms are fundamental for surveillance: (i) isolate WGS, which delivers high-resolution phylogenetics and precise AMR determinant calling, and (ii) metagenomic NGS (mNGS), which profiles mixed specimens (e.g., blood, cerebrospinal fluid, bronchoalveolar lavage, wastewater) without culture.

A 7-year clinical mNGS evaluation highlighted stable diagnostic yield but also the operational drivers of TAT: 8.2–11.4 days from sample collection to result and 3.6–3.8 days from start of processing to result. It emphasizes the need for workflow engineering and automation that include the pre-analytical phase (decision to order testing, sample shipping, batch testing) to achieve true “real-time” (23). AI-assisted triage could be a practical target for improvement of the pre-analytics. Examples of AI-based systems applications include auto-eligibility of cases that might benefit from NGS (e.g., culture-negative samples), automated creation of samples shipping orders, and efficient batching of a sufficient number of specimens for NGS (24).

In 2021, in a study on new workflow for real-time prediction of pathogenic microorganisms, Bartoszewicz et al. (25) built DL models specialized in inference from incomplete short- and long-read sequencing data. The models showed an accuracy higher than 78.5% in inferring whether an Illumina or a Nanopore sequence belongs to pathogenic bacteria or viruses, for reads as short as 50 bp obtained during the Illumina sequencing run or after the first 250 bp (0.5 s of sequencing time) of Nanopore reads (25). If properly validated in the clinical context, such AI-based systems, performed in real-time during the sequencing run, could potentially save hours (depending

on the platform used) in the post-analytical phase of pathogens identification.

However, Liu et al. (26) showed that adaptive nanopore sequencing, even without AI capabilities, already enabled reliable pathogen identification using the reads obtained within the first hour of sequencing, with a categorical agreement (CA) of 89.3%. Overall, it achieved 90.6% CA for AMR prediction with very major error rates of only 1–3% in key pathogens, and detected 23 polymicrobial infections, including 7 missed by standard diagnostics, substantially improving timely stewardship and outbreak control (26).

Beyond diagnostics, long read-based technologies like nanopore resolve complex AMR loci and structural variants that short-read pipelines may miss, strengthening surveillance signals about transmission of resistance cassettes (2).

Multiple evaluations now show that embedding WGS into infection-prevention programs improves timeliness and precision of outbreak recognition, with measurable patient and economic benefits when results are integrated into routine workflows. Optimized pipelines can move from sample to action in a clinically relevant window, particularly when coupled to automated analytics and visualization for infection prevention and control teams (27). This combination of AI and WGS has been successfully evaluated for outbreak investigation in hospital settings. Sundermann et al. (28) developed an AI algorithm that helped identify likely transmission routes after genomic clustering was established. The model increased the identification of epidemiological links by 3.8% as compared to traditional review methods of infection prevention and control, with an overall sensitivity of 91.7% (28).

For population-scale surveillance studies, including resistome surveillance, metagenomic analysis of sewage has been proposed as a powerful tool, capturing antibiotic resistance gene (ARG) diversity and trends independent of testing behavior. Large international datasets demonstrated global heterogeneity of ARGs and their association with socio-economic and environmental factors, while newer studies use serial wastewater metagenomes to relate resistome shifts to antimicrobial use and interventions, positioning wastewater-based epidemiology (WBE) as a sensitive barometer for public-health AMR policies (29, 30). WBE might be integrated with AI-based tools to improve surveillance capabilities. Liu et al. (31) demonstrate that ML models trained on WGS data using gene presence and absence profiles, SNP variants, and k-mer features can reliably predict antimicrobial resistance in *S. aureus*. In a dataset of 1,764 isolates tested against ten antibiotics, their best performing models achieved AUC values ranging from 0.83 to 0.99, with six antibiotics reaching an AUC above 0.99 and nine exceeding 0.96. These findings show that rich genomic feature sets enable ML classifiers to capture resistance determinants with consistent accuracy across various antimicrobial agents (31).

Emerging studies clarify that “real-time” also involves an efficient integration of sample logistics, run time, computation, expert review, and feedback into decisions. Reported clinical mNGS experiences indicate that shaving days off logistics and report generation often dwarfs gains from faster chemistry alone, areas where AI-driven automation (eligibility triage, templated narratives, anomaly alerts) can reduce wall-clock time (32).

6 Integration of artificial intelligence with other technologies

While this review centers on genome-based AI for AMR prediction, emerging non-genomic tools like Raman spectroscopy and matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS) can complement these approaches. When paired with ML, they offer rapid, label-free resistance profiling, especially useful when sequencing is unavailable or delayed.

6.1 Raman spectroscopy and machine learning

While culture-based methods remain the diagnostic standard, their prolonged TAT might be inappropriate for the urgency of severe infections. The integration of Raman spectroscopy with ML emerges as a powerful challenger, potentially poised to redefine rapid microbiological identification, before culture results are available (33, 34). Specimens from nasopharyngeal, rectal, or wound swabs are analyzed with minimal preparation, significantly shortening diagnostic timelines. The derived spectral patterns can then be fed to supervised ML classifiers, such as SVMs, Random Forests, and deep neural networks, which promptly identify not only microbial species, but also related AMR phenotypes (35). The single-cell resolution offered by Raman spectroscopy with ML, could represent a valuable tool for the diagnostics of pathogenic bacteria and AMR. The study by Lu et al. (36) showed an average accuracy in the identification of 12 species of common pathogenic bacteria (e.g., *E. coli*, *K. pneumoniae*, *S. aureus*) equal to $90.73 \pm 9.72\%$. While the model built for the study was unable to discern specific antibiotic resistance, it was able to discriminate with $99.92 \pm 0.06\%$ accuracy drug-resistant *A. baumannii* isolates from susceptible clones. However, Raman spectroscopy has not yet been applied to pathogenic identification and antibiotic susceptibility detection in clinical practice, limiting the generalizability of these observations (36). Practical barriers include the need for standardized spectral databases, variability in sample preparation, model generalizability across diverse clinical settings, and limited interpretability of spectral features (37). While the approach shows clear promise, its adoption will depend on addressing technical, operational, and regulatory hurdles.

6.2 Mass spectrometry and machine learning

While MALDI-TOF MS has already revolutionized microbial identification, its integration with ML is now extending its capabilities into strain-level epidemiology and predictive resistance profiling. This synergy is particularly impactful for the direct analysis of clinical swabs, moving beyond mere identification to deliver actionable insights at a pace that could directly influence patient outcomes. The performance of any downstream algorithm is tightly coupled to the quality and diversity of the reference spectral datasets. The added value of the integration of ML with MALDI-TOF becomes evident precisely where traditional methods reach their limits. For example, complex taxonomic groups like the *Enterobacter cloacae* complex often confuse conventional diagnostics, whereas MALDI-TOF

combined with ML can accurately carry out species-level discrimination tasks (38). Perhaps more significantly, it can be scaled to outbreak investigations, differentiating epidemiologically relevant strains (e.g., *P. aeruginosa* ST175) when WGS is not available, promptly informing infection control teams (39). For common pathogens like *Klebsiella* spp. or *E. coli*, the challenge is no longer identification, but rapid resistance profiling since predicting antibiotic resistance within hours rather than days allows a swift transition from empirical to targeted therapy. Ultimately, ML does not merely automate MALDI-TOF interpretation; but transforms a gold-standard identification tool into a dynamic source of high-fidelity intelligence for clinical decision-making as shown in several recent studies. In one example, a 2024 investigation on *Staphylococcus epidermidis*, through the analysis of best-annotated protein, reported 94% accuracy in predicting methicillin resistance based on subtle spectral variations, circumventing the 24-h delay associated with traditional testing (40). Comparably, in Gram-negative rods a neural network trained to identify carbapenem-resistant *Klebsiella pneumoniae* showed over 90% accuracy (41). The combination of ML and MALDI-TOF marks a pivotal shift in AMR diagnostics. Although the technological potential is undeniable, bringing these innovations into daily clinical use reveals a familiar tension: progress in research outpacing its clinical integration. In fact, high performance often depends on narrowly selected datasets, while real-world clinical samples are often polymicrobial infections and with poor-quality spectra. Additionally, many ML-derived spectral features lack clear biological annotation. It is one thing to identify a peak statistically associated with resistance; it is another to link it mechanistically without confirmatory genomic evidence, a step frequently skipped. Regulatory concerns further complicated the matter. Still, the momentum behind this technology should not be lost. Addressing these issues is not just desirable but needed, if we aspire to move from proof-of-concept to practical, trustworthy tools in patient care.

7 State of the art and current challenges

7.1 State of the art

The integration of AI into clinical microbiology might represent a valuable transition from culture-dependent diagnostics to proactive, data-driven decision-making. The convergence of genome sequencing and other technologies is delivering significant improvements in speed, accuracy, and diagnostic resolution (33, 38). One of the most significant benefits would lie in the ability to potentially reduce the number of hours imposed by the standard diagnostic timeline, allowing clinicians to act earlier and with greater precision. This emerging diagnostic landscape, summarized in Table 2, also promises to alleviate the heavy manual workload in clinical microbiology laboratories, automating time-consuming tasks, and allowing skilled personnel to focus on higher-order analytical and interpretive challenges. The field of AI-assisted microbiology against AMR is growing rapidly, with a diverse array of models and modalities emerging to address different clinical and public health needs. While this variety reflects innovation and potential, it can be overwhelming for clinicians and decision-makers. Many published models are trained in tightly scoped settings, optimized for specific data types or

institutions, and reported using performance metrics that may not generalize beyond their original context. The result is a proliferation of tools whose applicability, reproducibility, and true clinical impact are not always clear. Publication bias further complicates interpretation, as models are often internally tuned and validated on the same data until optimal results are achieved, potentially giving a misleading impression of robustness. Looking at common failure modes that may limit model performance or generalizability in real-life scenarios, we found (1) the risk of dataset shifts, for both geographical and methodological reasons, to be relevant as many studies were analyzing data coming from publicly available microbiology datasets collecting isolates' data from all over the world, and often the geographic origin of the isolate, or methodological differences in isolates' analyses, were not taken into account (8, 10, 11, 25, 30), other than in the studies by Munk et al. (29) and Candela et al. (38). Moreover, (2) only two studies mentioned training the model on one dataset and then testing on other datasets pursuing multicenter external validation (38, 40), whereas many experiments relied on hold-out validation or internal train/test splits (9, 20, 25, 35, 41); or pursued k-fold cross validation (10, 11, 32, 36). Thirdly, (3) Only a minority explicitly addressed the risk of phylogenetic confounding (10, 20, 40), for example by analyzing DNA degradation signatures, despite evidence that some models may be capturing and learning lineage-related features (e.g., k-mer counts or taxonomic features), rather than detecting causal resistance determinants mechanisms (9, 11, 25, 31, 35, 36, 41). Furthermore, (4) clinical interpretability of findings remains uneven among studies reporting on antimicrobial susceptibility as only one study framed errors in terms of major and very major rates (9), while the vast majority reported only global metrics such as accuracy or AUROC and precision-recall curves, sensitivity, and specificity, without extensively addressing misclassification risk, which obscure risks relevant to frontline use in different scenarios (10, 11, 20, 31, 35, 36, 40, 41). Finally, (5) transparency also varied but there are encouraging signals worth mentioning: several included studies shared public access to the reference code on platforms such as GitHub, or released the model under an open license (9, 11, 20, 25). Moreover, 6 papers are directly referencing publicly available microbiology datasets (e.g., NCBI Sequence Read Archive, PATRIC ID, DRIAMS) (8, 10, 19, 29, 38); only one study guaranteed both full microbiological data and code availability (40). This shows an effort towards reproducibility is being made, which is a fundamental step towards broader adoption and future multicenter validation of these models. Taken together, these findings underscore the importance of rigorous validation and full-spectrum risk reporting before deployment of AI tools into clinical microbiology and surveillance workflows.

7.2 Current challenges

Several critical obstacles, some of which have already been mentioned above, must be addressed before these technologies can be adopted widely and safely. First, the cost of acquiring and implementing AI-integrated platforms, especially those requiring advanced sequencing or high-resolution imaging, is a major barrier for many institutions. Second, the performance and fairness of AI systems are directly tied to the quality and diversity of their training data. As Peiffer-Smadja et al. (34) highlight, the lack of large,

TABLE 2 Summary of included studies describing AI task and models, input data, examples among included studies, and public-health applications.

Core model task	Main clinical/lab application	Input data types	AI approaches	Example in paper	Public health use case
Pathogen or taxon identification and characterization	Rapid species identification	MALDI-TOF MS spectra	RF; SVM; PLS-DA; PCA	<i>E. cloacae complex</i> species discrimination vs. reference databases (38)	Outbreak-relevant taxonomy
	Rapid pathogen identification	Raw NGS reads (Illumina/Nanopore)	CNN; LSTM	Real-time DL classification of pathogenic reads during sequencing (25)	Biosecurity surveillance
	Rapid pathogen identification	Single-cell Raman spectra	CNN compared with RF; SVM; AdaBoost; RF + PCA	Single-cell Raman → RF pathogen identification and AST (35, 36)	Early AST-guided public health response
	Viability classification	Raw nanopore signal time-series	CNN; XAI (Class activation maps)	DL-based viability inference with explainable “signal-drop” patterns in metagenomics (20)	Explainable AI application in fast microbiology; Antimicrobial and diagnostic stewardship
Outbreak detection and clone characterization	Rapid high-risk clone detection	MALDI-TOF MS; FTIR-S spectra; WGS data	RF; SVM; NCA-kNN; PLS-DA; PCA	Identification of a high-risk clone (GC1- vs. non-GC1) via genome sequencing (11) Threshold vs. biomarker peak-matrix discrimination of a high-risk clone via MALDI-TOF (39)	Outbreak strain surveillance
	Hospital outbreak investigation support + transmission-route attribution	WGS data + EHR spatiotemporal data	Logistic regression; likelihood-ratio ML; importance sampling	EDS-HAT: WGS-EHR integration identifies hidden outbreaks and routes (27, 28)	Improved IPC containment; population data mining for AMR surveillance
Discovery of Resistance determinants and resistome structure	Resistome profiling	Shotgun metagenomics; antibiotic resistance genes profiles	DL; Kraken2; PCA; PERMANOVA; networks	Plasmid detection + mapping of global resistome (29); framing and interpretation of resistome structure and antibiotic usage associations (30)	Global AMR surveillance (One Health perspective); early-warning systems
	Resistance determinant discovery	Public genomes (PATRIC); WGS assemblies; pan-genome gene/allele matrices	RF; SVM; NB; AdaBoost; GA feature selection	Pan-genome SVM ranking identifies genes potentially involved in new mechanisms of resistance (8, 10)	Advancement in AMR genotype knowledge
	ICU resistome analysis for clinical risk stratification purposes	Clinical variables; shotgun metagenomics (from patient and environmental sampling); distance matrices	ML-adjacent methods (PCA, network inference)	ROC/AUC-based regression plus multivariate ordination and network analysis (19)	Population-level pattern discovery

(Continued)

TABLE 2 (Continued)

Core model task	Main clinical/lab application	Input data types	AI approaches	Example in paper	Public health use case
Antimicrobial susceptibility prediction	Binary AMR classification	WGS assemblies; pan-genomic library (gene presence/absence)	RF; SVM; NB; AdaBoost; GA feature selection	Pan-genome workflow identifying known AMR determinants (8, 10)	Improved IPC containment Enhanced antimicrobial stewardship
	Binary AMR classification	MALDI-TOF MS spectra; +/- AST labels	RF; SVM; LR; NB; LightGBM; MLP; SHAP; ANN + feature selection	RF-mediated interpretable AMR predictions (40) Full-spectrum ANN CRKP screening before AST (41)	Improved IPC containment; Enhanced antimicrobial stewardship; Explainable AI-based AMR surveillance
	Accelerated identification and AST	Single-cell Raman spectra	CNN compared with RF; SVM; AdaBoost; RF + PCA	Culture free (Raman spectroscopy + RF) pathogen identification and AST (35, 36)	Early AST-guided public health response
	MIC prediction → S/I/R classification	WGS reads; k-mer matrices	RF; SVM; XGBoost;	k-mer-based MIC prediction capturing <i>S. aureus</i> resistance mechanism (SCCmec) and phenotypic resistance prediction (9)	Improved IPC containment; Enhanced antimicrobial stewardship

AI, Artificial Intelligence; ML, Machine Learning; AMR, Antimicrobial Resistance; AST, Antimicrobial Susceptibility Testing; IPC, Infection Prevention and Control; ROC, Receiver Operator Characteristic curve; AUC, Area Under the Curve; SVM, Support Vector Machine; DL, Deep Learning; kNN, k-Nearest Neighbour; CNN, Convolutional Neural Network; ANN, Artificial Neural Network; DT, Decision Trees; RF, Random Forest; NB, Naïve Bayes; AdaBoost, Adaptive Boosting; XGBoost, Extreme Gradient Boosting; LightGBM, Light Gradient Boosting Machine; PCA, Principal Component Analysis; SHAP, SHapley Additive exPlanations; FTIR-S, Fourier-transform Infrared Spectroscopy; MALDI-TOF MS, Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry; LR, Logistic Regression; GA, Genetic Algorithm; XAI, eXplainable AI; PERMANOVA, Permutational Multivariate Analysis of Variance; PLS-DA, Partial Least Squares Discriminant Analysis; CRKP, Carbapenem-resistant *Klebsiella pneumoniae*; CSKP, Carbapenem-susceptible *Klebsiella pneumoniae*; LSTM, Long Short-Term Memory; NGS, Next Generation Sequencing; EDS-HAT, Enhanced Detection System for Healthcare-Associated Transmission; EHR, Electronic Health Record; MLP, Multilevel Perceptron.

standardized, and globally representative datasets, particularly for spectral and genomic data, limits model generalizability and raises the risk of diagnostic bias across different populations and healthcare settings. Third, the so-called “black box” nature of many DL models continues to undermine clinical trust. Clinicians are understandably reluctant to rely on predictions that lack transparent reasoning, especially in high-stakes decisions. Fourth, the lack of reproducibility requirements and external validation that result in lack of generalizability. Finally, such AI-based systems should be integrated, as appropriate, with specific resources like databases, that collect and organize reference information on AMR, and international guidelines to interpret AMR (42, 43).

To advance, the field must commit to addressing these challenges through open, interdisciplinary collaboration. Key actionable priorities could include: (a) developing expansive, curated, and ethically sourced multimodal databases that include diverse patient populations and microbial clones, (b) improving explainable AI methodologies that make model outputs interpretable and auditable (e.g., enabling clinicians to visualize which specific spectral peaks or resistance genes underpin a given prediction), (c) designing hybrid, tiered diagnostic systems that combine rapid techniques (e.g., Raman spectroscopy) with high-resolution methods (e.g., metagenomic sequencing) to create scalable workflows tailored to clinical urgency and case complexity, and (d) ensuring the development and availability of cost-effective and accessible technologies across diverse healthcare settings to mitigate disparities in AI adoption and prevent technology-driven discrimination, (e) fostering prospective, multicentric

validation of AI tools to explore real-world usability and understand how AI tools behave under heterogeneous epidemiology, laboratory practices, patient case-mix, and data quality to support safer deployment and more appropriate scope definition before widespread adoption, and (f) encouraging the use of AI agent-based coordination approaches to integrate outputs from multiple task-specific tools, with the aim of reducing information fragmentation and supporting coherent, prioritized clinical decision-making in complex clinical workflows.

By addressing these limitations head-on, the fusion of AI and microbial diagnostics can evolve from isolated innovations into a robust, intelligent, and equitable diagnostic ecosystem, one that accelerates results and enhances precision, adaptability, and resilience in clinical microbiology.

8 Relevance for public health and future perspectives

The methodologies reviewed in this paper, specifically the application of AI to genomic strategies, offer transformative opportunities for pathogen identification and AMR detection. These AI-driven tools hold the potential to substantially reduce diagnostic TAT, enabling clinicians to initiate targeted therapies with greater promptness and appropriateness. As a result, the implications of AI-aided sequencing and spectroscopy for public health are profound: at the bedside, ML models applied to MALDI-TOF or Raman spectra

can deliver rapid AMR predictions, offering valuable support for early treatment decisions and antimicrobial stewardship. Further upstream, WGS and mNGS provide higher-resolution insights, enabling the detection of transmission clusters, characterization of resistance mechanisms, and tracking of high-risk clones across healthcare facilities and regions. These genomic approaches may also underpin sentinel surveillance programs and broader population-level initiatives, such as national strain monitoring and prevalence mapping, though they require robust infrastructure, standardized workflows, and centralized coordination. Wastewater-based epidemiology is emerging as a complementary strategy, leveraging mNGS to detect shifts in community-level resistomes and early signs of emerging threats, though analytical complexity remains a barrier to widespread implementation. In low-resource settings, simplified and portable solutions, such as real-time nanopore sequencing or MALDI-based species identification, can provide actionable insights with lower technical requirements, provided there is investment in training and system integration.

The near future likely lies in a tiered diagnostic framework, which may help balance speed, cost, and resolution: for instance, deploying MALDI-TOF or Raman-based ML for rapid, priority triage, while reserving WGS or mNGS for high-risk outbreak scenarios, such as MDRO outbreaks in large ICUs or oncohematological facilities, that require deeper genomic resolution and contextual interpretation. Moving forward, the integration of population-based metrics other than hospital systems' EHRs may warrant future integration of AI in public health interventions against AMR such as improving vaccine uptake in fragile populations or assist and monitor the prescription of antimicrobials in primary care (44, 45).

Such a stratified approach balances diagnostic depth with operational feasibility, ensuring that high-resolution insight is applied precisely where it is needed most.

9 Conclusion

Taken together, the evidence reviewed indicates that AI has the potential to substantially enhance clinical microbiology and AMR surveillance, but its impact will depend on disciplined translation rather than technological novelty alone. Progress will require robust validation, transparent reporting of clinically meaningful errors, and seamless integration of AI outputs into resource-aware and coherent diagnostic workflows. Only through such a rigorous and pragmatic path can AI move from isolated proofs of concept to reliable components of routine microbiological practice.

Author contributions

GS: Conceptualization, Writing – original draft, Writing – review & editing, Supervision. NM: Writing – original draft, Writing – review & editing. AR: Supervision, Writing – review & editing. EP: Supervision, Writing – review & editing. DC: Writing – review & editing. SR: Resources, Writing – review & editing. GB: Writing – review & editing. LP: Writing – review & editing. MC: Supervision,

Writing – review & editing. AG: Funding acquisition, Resources, Writing – review & editing. FB: Conceptualization, Writing – original draft, Writing – review & editing, Supervision.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This research was supported by the EU funding within the Next Generation EU-MUR PNRR Extended Partnership initiative on Emerging Infectious Diseases (Project no. PE00000007, INF-ACT).

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author AG declared that they were an editorial board member of *Frontiers* at the time of submission. This had no impact on the peer review process and the final decision.

Correction note

A correction has been made to this article. Details can be found at: [10.3389/fpubh.2026.1824508](https://doi.org/10.3389/fpubh.2026.1824508).

Generative AI statement

The author(s) declared that Generative AI was used in the creation of this manuscript. Generative AI was used for copyediting.

Any alternative text (alt text) provided alongside figures in this article has been generated by *Frontiers* with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2026.1757161/full#supplementary-material>

References

- World Health Organization. Available online at: <https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance> (2023) (Accessed February 10, 2025).
- Sherry NL, Lee JYH, Giulieri SG, Connor CH, Horan K, Lacey JA, et al. Genomics for antimicrobial resistance—progress and future directions. *Antimicrob Agents Chemother.* (2025) 69:e0108224. doi: 10.1128/aac.01082-24
- Vandenberg O, Durand G, Hallin M, Diefenbach A, Gant V, Murray P, et al. Consolidation of clinical microbiology laboratories and introduction of transformative technologies. *Clin Microbiol Rev.* (2020) 33:33. doi: 10.1128/CMR.00057-19
- Zheng D, Pang G, Liu B, Chen L, Yang J. Learning transferable deep convolutional neural networks for the classification of bacterial virulence factors. *Bioinformatics.* (2020) 36:3693–702. doi: 10.1093/bioinformatics/btaa230
- Rizzo A, Mensa E, Giacomelli A. The future of large language models in fighting emerging outbreaks: lights and shadows. *Lancet Microbe.* (2024) 5:100954. doi: 10.1016/j.lanmic.2024.07.017
- Rizzo A, De Vito A, Lucis R, Mileto D, Giacomelli A. Chain of thought reasoning in enhancing infectious disease diagnosis and microbiological analysis. *Am J Infect Control.* (2025) 53:4111–2. doi: 10.1016/j.ajic.2024.11.009
- Suster CJE, Pham D, Kok J, Sintchenko V. Emerging applications of artificial intelligence in pathogen genomics. *Front Bacteriol.* (2024) 3:1326958. doi: 10.3389/fbct.2024.1326958
- Her H-L, Wu Y-W. A pan-genome-based machine learning approach for predicting antimicrobial resistance activities of the *Escherichia coli* strains. *Bioinformatics.* (2018) 34:i89–95. doi: 10.1093/bioinformatics/bty276
- Wang S, Zhao C, Yin Y, Chen F, Chen H, Wang H. A practical approach for predicting antimicrobial phenotype resistance in *Staphylococcus aureus* through machine learning analysis of genome data. *Front Microbiol.* (2022) 13:13. doi: 10.3389/fmicb.2022.841289
- Hyun JC, Kavvas ES, Monk JM, Palsson BO. Machine learning with random subspace ensembles identifies antimicrobial resistance determinants from pan-genomes of three pathogens. *PLoS Comput Biol.* (2020) 16:e1007608. doi: 10.1371/journal.pcbi.1007608
- Álvarez VE, Quiroga MP, Centrón D. Identification of a specific biomarker of *Acinetobacter baumannii* global clone 1 by machine learning and PCR related to metabolic fitness of ESKAPE pathogens. *mSystems.* (2023) 8:8. doi: 10.1128/mSystems.00734-22
- Zwiers LC, Grobbee DE, Uijl A, Ong DSY. Federated learning as a smart tool for research on infectious diseases. *BMC Infect Dis.* (2024) 24:1327. doi: 10.1186/s12879-024-10230-5
- Fida M, Wolf MJ, Hamdi A, Vijayvargiya P, Esquer Garrigos Z, Khalil S, et al. Detection of pathogenic bacteria from septic patients using 16S ribosomal RNA gene-targeted metagenomic sequencing. *Clin Infect Dis.* (2021) 73:1165–72. doi: 10.1093/cid/ciab349
- Batool M, Galloway-Peña J. Clinical metagenomics—challenges and future prospects. *Front Microbiol.* (2023) 14:14. doi: 10.3389/fmicb.2023.1186424
- Roy G, Prifti E, Belda E, Zucker J-D. Deep learning methods in metagenomics: a review. *Microb Genom.* (2024) 10:10. doi: 10.1099/mgen.0.001231
- Pillay S, Calderón-Franco D, Urhan A, Abeel T. Metagenomic-based surveillance systems for antibiotic resistance in non-clinical settings. *Front Microbiol.* (2022) 13:13. doi: 10.3389/fmicb.2022.1066995
- Lal Gupta C, Kumar Tiwari R, Cytryn E. Platforms for elucidating antibiotic resistance in single genomes and complex metagenomes. *Environ Int.* (2020) 138:105667. doi: 10.1016/j.envint.2020.105667
- de Abreu VAC, Perdigão J, Almeida S. Metagenomic approaches to analyze antimicrobial resistance: an overview. *Front Genet.* (2021) 11:575592. doi: 10.3389/fgene.2020.575592
- Mikolas M, Fauszt P, Petrilla A, Nemeth P, David P, Szilagyi-Tolnai E, et al. Analysis of ICU resistome dynamics in patients, staff and environment for the identification of predictive biomarkers of sepsis and early mortality. *Sci Rep.* (2025) 15:25080. doi: 10.1038/s41598-025-10848-8
- Ürel H, Benassou S, Marti H, Reska T, Sauerborn E, Pinheiro Alves De Souza Y, et al. Nanopore- and AI-empowered microbial viability inference. *Gigascience.* (2025) 14:giaf100. doi: 10.1093/gigascience/giaf100
- Elliott TM, Harris PN, Roberts LW, Doidge M, Hurst T, Hajkovic K, et al. Cost-effectiveness analysis of whole-genome sequencing during an outbreak of carbapenem-resistant *Acinetobacter baumannii*. *Antimicrob Steward Healthc Epidemiol.* (2021) 1:e62. doi: 10.1017/ash.2021.233
- Struelens MJ, Ludden C, Werner G, Sintchenko V, Jokelainen P, Ip M. Real-time genomic surveillance for enhanced control of infectious diseases and antimicrobial resistance. *Front Sci.* (2024) 2. doi: 10.3389/fsci.2024.1298248
- Benoit P, Brazer N, de Lorenzi-Tognon M, Kelly E, Servellita V, Oseguera M, et al. Seven-year performance of a clinical metagenomic next-generation sequencing test for diagnosis of central nervous system infections. *Nat Med.* (2024) 30:3522–33. doi: 10.1038/s41591-024-03275-1
- Athanasopoulou K, Michalopoulou V-I, Scorilas A, Adamopoulos PG. Integrating artificial intelligence in next-generation sequencing: advances, challenges, and future directions. *Curr Issues Mol Biol.* (2025) 47:470. doi: 10.3390/cimb47060470
- Bartoszewicz JM, Genske U, Renard BY. Deep learning-based real-time detection of novel pathogens during sequencing. *Brief Bioinform.* (2021) 22:22. doi: 10.1093/bib/bbab269
- Liu P-Y, Wu H-C, Li Y-L, Cheng H-W, Liou C-H, Chen F-J, et al. Comprehensive pathogen identification and antimicrobial resistance prediction from positive blood cultures using nanopore sequencing technology. *Genome Med.* (2024) 16:141. doi: 10.1186/s13073-024-01416-2
- Sundermann AJ, Kumar P, Griffith MP, Waggle KD, Rangachar Srinivasa V, Raabe N, et al. Real-time genomic surveillance for enhanced healthcare outbreak detection and control: clinical and economic impact. *Clin Infect Dis.* (2025) [Epub ahead of Print]. doi: 10.1093/cid/ciaf216
- Sundermann A, Chen J, Saul M, Shutt K, Griffith M, Snyder G, et al. Artificial intelligence enhances genomic surveillance in healthcare outbreak investigations. *Infect Control Hosp Epidemiol.* (2025):1–5. doi: 10.1017/ice.2025.10355
- Munk P, Brinch C, Møller FD, Petersen TN, Hendriksen RS, Seyfarth AM, et al. Author correction: genomic analysis of sewage from 101 countries reveals global landscape of antimicrobial resistance. *Nat Commun.* (2023) 14:178. doi: 10.1038/s41467-023-35890-w
- Maile-Moskowitz A, Brown CL, Rumi MA, Mouri NA, Majeed H, Finkielstein CV, et al. Relating antimicrobial use to wastewater resistance gene patterns via metagenomic analysis of two neighboring treatment plants circa the COVID-19 pandemic. *NPJ Antimicrob Resist.* (2025) 3:82. doi: 10.1038/s44259-025-00153-9
- Liu Y, Wang X, Wu L, Shi Z, Zhu M, Ye L, et al. Prediction of antimicrobial resistance in *Staphylococcus aureus* with a machine learning classifier based on WGS data. *Microbiol Spectr.* (2025) 13:e0006525. doi: 10.1128/spectrum.00065-25
- Ardila CM, Yadalam PK, González-Arroyave D. Integrating whole genome sequencing and machine learning for predicting antimicrobial resistance in critical pathogens: a systematic review of antimicrobial susceptibility tests. *PeerJ.* (2024) 12:e18213. doi: 10.7717/peerj.18213
- Alsulimani A, Akhter N, Jameela F, Ashgar RI, Jawed A, Hassani MA, et al. The impact of artificial intelligence on microbial diagnosis. *Microorganisms.* (2024) 12:1051. doi: 10.3390/microorganisms12061051
- Peiffer-Smadja N, Dellière S, Rodriguez C, Birgand G, Lescure F-X, Fourati S, et al. Machine learning in the clinical microbiology laboratory: has the time come for routine practice? *Clin Microbiol Infect.* (2020) 26:1300–9. doi: 10.1016/j.cmi.2020.02.006
- Kang H, Wang Z, Sun J, Song S, Cheng L, Sun Y, et al. Rapid identification of bloodstream infection pathogens and drug resistance using Raman spectroscopy enhanced by convolutional neural networks. *Front Microbiol.* (2024) 15:15. doi: 10.3389/fmicb.2024.1428304
- Lu W, Li H, Qiu H, Wang L, Feng J, Fu YV. Identification of pathogens and detection of antibiotic susceptibility at single-cell resolution by Raman spectroscopy combined with machine learning. *Front Microbiol.* (2023) 13. doi: 10.3389/fmicb.2022.1076965
- Bilal H, Khan MN, Khan S, Shafiq M, Fang W, Khan RU, et al. The role of artificial intelligence and machine learning in predicting and combating antimicrobial resistance. *Comput Struct Biotechnol J.* (2025) 27:423–39. doi: 10.1016/j.csbj.2025.01.006
- Candela A, Guerrero-López A, Mateos M, Gómez-Asenjo A, Arroyo MJ, Hernandez-García M, et al. Automatic discrimination of species within the *Enterobacter cloacae* complex using matrix-assisted laser desorption/ionization-time of flight mass spectrometry and supervised algorithms. *J Clin Microbiol.* (2023) 61:e0104922. doi: 10.1128/jcm.01049-22
- Candela A, Arroyo MJ, Sánchez-Cueto M, Marín M, Cercenado E, Méndez G, et al. Rapid discrimination of *Pseudomonas aeruginosa* ST175 isolates involved in a nosocomial outbreak using MALDI-TOF mass spectrometry and FTIR spectroscopy coupled with machine learning. *Transbound Emerg Dis.* (2023) 2023:1–11. doi: 10.1155/2023/8649429
- Ren M, Chen Q, Zhang J. Repurposing MALDI-TOF MS for effective antibiotic resistance screening in *Staphylococcus epidermidis* using machine learning. *Sci Rep.* (2024) 14:24139. doi: 10.1038/s41598-024-75044-6
- Zhang Y-M, Tsao M-F, Chang C-Y, Lin K-T, Keller JJ, Lin H-C. Rapid identification of carbapenem-resistant *Klebsiella pneumoniae* based on matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and an artificial neural network model. *J Biomed Sci.* (2023) 30:25. doi: 10.1186/s12929-023-00918-2
- Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya AR, Wlodarski MA, et al. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the comprehensive antibiotic resistance database. *Nucleic Acids Res.* (2023) 51:D690–9. doi: 10.1093/nar/gkac920

43. Kahlmeter G, Brown DFJ, Goldstein FW, MacGowan AP, Mouton JW, Odenholt I, et al. European committee on antimicrobial susceptibility testing (EUCAST) technical notes on antimicrobial susceptibility testing. *Clin Microbiol Infect.* (2006) 12:501–3. doi: 10.1111/j.1469-0691.2006.01454.x

44. Bianchi FP, Stefanizzi P, Di Lorenzo A, Cuscianna E, Tafuri S, Germinario CA. Vaccine coverage for recommended vaccines among splenectomised patients in Apulia,

South Italy: a retrospective cohort study. *BMJ Open.* (2023) 13:e069316. doi: 10.1136/bmjopen-2022-069316

45. Simão BLP, Pereira CM, Jácome M, Oliveira C, Ferreira LM, Paiva JM, et al. Artificial intelligence in the prescription of acute medical treatments in primary healthcare - comparison of the performance of family physicians and ChatGPT. *BMC Prim Care.* (2025) 26:284. doi: 10.1186/s12875-025-02963-2