

Why Artificial Intelligence Models for Antimicrobial Resistance Still Fail at the Bedside: A Review of Validation, Explainability, and Equity Gaps



Md. Kaium ^{1*}, Suhel Ahmed ²

Abstract

Artificial intelligence (AI) and machine learning (ML) models now predict antimicrobial resistance (AMR) phenotypes with striking accuracy in retrospective, single-centre studies — and yet, curiously, almost none of them have made it as far as routine clinical use. This review set out to understand that gap, and, more usefully, what might close it. Following PRISMA 2020 guidance, we systematically synthesised peer-reviewed and preprint literature (2017–2026) on AI/ML applications for AMR prediction, mapping studies across four domains: comparative model performance, diagnostic input modality, data heterogeneity and algorithmic bias, and technological readiness in low- and middle-income countries (LMICs). Risk of bias was appraised using a PROBAST-informed framework. Tree-based ensembles and deep learning architectures achieved internal AUROC values of 0.89–0.99 across genomic and mass-spectrometric datasets; externally validated cohorts, however, showed a consistent 0.10–0.25 AUROC decay. Over 80% of publicly available pathogen genomic data originate from North America and Europe, while African isolates contribute under 2%. Calibration reporting and biologically validated explainability were almost entirely absent from the reviewed literature, and One Health data integration remained confined to fewer than 1% of studies.

Taken together, these findings suggest that the translational gap in AI-based AMR prediction is not, at its core, a modelling problem — it is a problem of representativeness, validation rigour, and earned clinical trust. Closing it will require external and prospective validation as a reporting norm, calibrated and biologically grounded explainability, federated infrastructure inclusive of LMIC data, and tiered, offline-capable deployment pathways suited to resource-limited settings.

Keywords

antimicrobial resistance; machine learning; external validation; explainable AI; health equity

1. Introduction

There is a particular kind of fatigue that sets in around phrases like “antimicrobial resistance” — not because the problem has become less serious, but because repetition, oddly, has a way of flattening urgency into background noise. It is worth resisting that fatigue here, if only for a moment, because the numbers underneath the phrase remain genuinely difficult to sit with. Bacterial AMR was directly responsible for an estimated 1.27 million deaths in 2019 alone, with a further 4.95 million deaths associated

Significance | Explains why highly accurate AI resistance models fail clinically, and outlines concrete validation, equity, and deployment fixes.

*Correspondence.

Md. Kaium, Department of Biochemistry and Molecular Biology, Siddheswari college, Dhaka, Bangladesh. E-mail:-md.kaium25k@gmail.com.

Editor Vishwanadham Mandala, Ph.D. And accepted by the Editorial Board Aug 28, 2026 (received for review Jun 05, 2026)

Author Affiliation: ¹ Department of Biochemistry and Molecular Biology, Siddheswari College, Dhaka, Bangladesh.

² Department of Public Health, Daffodil International University, Savar, Dhaka, Bangladesh.

Please cite this article as: Kaium, M., Ahmed, S. (2026). "Why Artificial Intelligence Models for Antimicrobial Resistance Still Fail at the Bedside: A Review of Validation, Explainability, and Equity Gaps", Integrative Biomedical Research, 10(1), 1-13, 10889

3068-6326/© 2026 (Integrative Biomedical Research), a publication of Eman Research, USA.
This is an open access article under the CC BY-NC-ND license.
(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)
(<https://publishing.emanresearch.org>)

with resistant infections more broadly (Murray et al., 2022). Left more or less unaddressed, this toll is projected to climb toward 10 million deaths a year by 2050 (Kiazolu, 2026a) — a figure that, once set beside an economic cost running into the tens of trillions of dollars, starts to resemble less a public-health statistic and more a slow-moving structural crisis (Bhattacharya et al., 2026; Kiazolu, 2026a). Nowhere is this felt more acutely than in intensive care units and emergency departments, where a delay of even a few hours in selecting the right antibiotic can measurably shift a patient's odds (Bilal et al., 2025).

Part of what makes this so stubborn, frankly, is procedural rather than biological. Culture-based antimicrobial susceptibility testing (AST) remains the diagnostic gold standard, and yet it takes what it takes: typically 24 to 72 hours before a definitive resistance profile comes back (Pennisi et al., 2025; Roy, 2026). In that window, clinicians have little real choice but to reach for broad-spectrum empirical therapy — a sensible decision for any one patient, but one that, repeated across millions of prescriptions, becomes a principal driver of the very multidrug-resistant organisms clinicians are trying to outrun (Boudza et al., 2026; Sanapalli et al., 2026). It is, in a sense, a closed loop: the diagnostic tool built to enable precise therapy is too slow to prevent the imprecise therapy that keeps resistance evolving.

Artificial intelligence and machine learning arrived, understandably, as a proposed way out of that loop. Over roughly the last decade, a widening range of algorithms — support vector machines and random forests, then deep neural networks, graph neural networks, and, more recently, transformer architectures — have been trained to infer resistance phenotypes directly from whole-genome sequencing (WGS) data, MALDI-TOF mass spectra, and electronic health record variables (Ahvar et al., 2025; Begum et al., 2026; Kiazolu, 2026a). And the reported numbers have, in fairness, often been impressive, with area-under-the-curve (AUROC) values regularly clearing 0.90 in laboratory benchmarks (Bilal et al., 2025; Shehata, 2025). On paper, it can look almost solved.

Almost is doing a lot of work in that sentence, though, and unpacking it is really the starting point for this review. A model's performance on held-out data from the same institution that generated its training set tells us surprisingly little about how it will behave elsewhere — a different hospital, a different country, or even the same site a few years later (Kiazolu, 2026a). Kiazolu (2026a)

describes this as a translational chasm, and the description seems apt: it is not one failure but a cluster of them, interacting. Training data is heterogeneous almost as a matter of definition (Kiazolu, 2026a; Mohammed et al., 2025); low- and middle-income countries, which carry the heaviest resistance burden, are barely represented in public datasets (Lyimo, 2026); external validation is rare enough that internal-only cross-validation routinely inflates apparent performance (Roy, 2026; Shehata, 2025); deep architectures remain largely opaque to the clinicians expected to trust them (Pennisi et al., 2025); and human clinical AI still sits largely isolated from the animal, food-chain, and environmental data that a genuine One Health approach would require (Kasse et al., 2025).

We treat these five axes — heterogeneity and bias, geographic imbalance, validation deficits, the black-box and calibration problem, and One Health isolation — not as separate technical footnotes but as symptoms of one underlying condition: the field has optimised heavily for accuracy within a narrow, unrepresentative slice of the world, while paying comparatively little attention to what it actually takes for a model to survive contact with a different setting. Around that observation, this review is organised around four questions — how heterogeneity and bias shape external performance decay; whether explainability and calibration can close the trust gap; whether federated learning can ease LMIC data scarcity without compromising privacy; and whether a One Health, multi-omics framework could improve outbreak forecasting — pursued through four corresponding objectives: mapping current AI/ML methodology, quantifying validation-related performance inflation, defining calibration and explainability reporting standards, and sketching a realistic, tiered implementation roadmap that does not assume every clinic has a GPU cluster.

3. Methods

This review was designed, conducted, and reported in accordance with the PRISMA 2020 statement (Page et al., 2021). The description below is written, deliberately, with enough procedural detail that another team — working from this text alone, without needing to contact us — should be able to reconstruct the search, screening, and synthesis steps with minimal ambiguity. That reproducibility standard, we think, matters more for a review like this than any particular formatting convention.

3.1 Protocol Registration and Eligibility Criteria

Eligibility was defined using a PICOS framework adapted for a methodological, rather than strictly clinical, question.

Population/Problem: primary studies, systematic reviews, or scoping reviews addressing bacterial AMR prediction, surveillance, or discovery in human, veterinary, or environmental contexts. **Index technology:** any AI/ML method, including classical statistical learning (logistic regression, random forests, gradient boosting, support vector machines), deep learning (convolutional, recurrent, transformer architectures), graph neural networks, and federated or explainable AI frameworks. **Comparator:** culture-based phenotypic AST, alternative modelling approaches, or, where reported, internal-versus-external validation cohorts within the same study. **Outcomes:** discriminative performance metrics (AUROC, accuracy, sensitivity, specificity, F1-score, Matthews correlation coefficient), calibration metrics where reported, and, where available, implementation-feasibility measures (turnaround time, computational cost, data accessibility). **Study designs:** retrospective and prospective cohort studies, diagnostic accuracy studies, external validation studies, and systematic/scoping reviews, published or posted as preprints between January 2017 and February 2026, restricted to full texts in English.

Studies were excluded if they addressed antifungal or antiviral resistance exclusively; concerned antimicrobial peptide or drug discovery without any resistance-prediction component; or reported insufficient detail to extract performance metrics or validation design. We logged exclusion reasons at both stages (see 3.3) rather than discarding them, so the final study count can be independently reconciled.

3.2 Information Sources and Search Strategy

We searched PubMed/MEDLINE, Scopus, Web of Science, and IEEE Xplore, together with the preprint servers medRxiv and Research Square, and supplemented these electronic searches with backward citation tracking (screening reference lists of eligible reviews) and forward citation tracking of key methodological papers — a step we would encourage other reviewers not to skip, since it is where several relevant preprints surfaced in our own search.

The core string combined three concept blocks with the Boolean operator AND, each block internally combined with OR:

AMR concepts: “antimicrobial resistance” OR “antibiotic

resistance” OR “drug resistance” OR AMR

Computational method concepts: “machine learning” OR “deep learning” OR “artificial intelligence” OR “neural network” OR “random forest” OR “transformer” OR “federated learning” OR “explainable AI”

Prediction/translation concepts: “prediction” OR “classification” OR “susceptibility testing” OR “clinical decision support” OR “external validation” OR “whole-genome sequencing” OR “MALDI-TOF”

For PubMed specifically, the string was implemented with MeSH terms mapped onto free-text equivalents (e.g., “Drug Resistance, Bacterial”[MeSH] OR “antimicrobial resistance”[tiab]) to satisfy indexing conventions, then combined with the computational and prediction blocks using the same Boolean logic described above. Searches were re-run in February 2026 to capture late-breaking preprint and early-access literature; the exact search date and database version accessed are reported so that a future update of this review can determine what, if anything, falls outside the original search window.

3.3 Study Selection and Screening

All records were exported to a shared reference manager, de-duplicated automatically and then manually spot-checked, and screened at title/abstract level independently by two reviewers against the eligibility criteria above. Disagreements were resolved by discussion; a third reviewer was available to adjudicate any conflict that discussion did not resolve. Full texts of studies passing this first pass were then assessed in the same independent-then-consensus fashion. At every stage, reasons for exclusion were recorded in enough detail to reconstruct a PRISMA 2020 flow diagram (Page et al., 2021), reporting the numbers of records identified, duplicates removed, records screened, full texts assessed, and studies ultimately included — the specific counts are reported in the flow diagram accompanying this manuscript.

3.4 Data Extraction

A standardised, piloted extraction form was used to capture, for each included study: first author and year; geographic origin of the training and (where applicable) validation cohort; dataset size and specimen type; predominant data modality (WGS, MALDI-TOF, EHR, transcriptomic, or metagenomic); ML architecture; key predictor variables; validation paradigm (internal cross-

validation, held-out internal test set, external geographic validation, external temporal validation, or prospective multicentre validation); best-reported performance metric; and, where discussed, calibration reporting, explainability method, and deployment considerations relevant to resource-limited settings. Extraction was piloted on a random 10% subset of eligible studies to check inter-reviewer consistency before proceeding to the full set. Extracted data underpin the comparative synthesis in Tables 1–4.

3.5 Risk-of-Bias and Applicability Assessment

Risk of bias for studies reporting a clinical prediction model was appraised using a Prediction model Risk Of Bias ASsessment Tool (PROBAST)-informed framework across four domains — participant selection, predictor definition, outcome definition, and statistical analysis — each rated low, moderate, or high risk. As a pre-specified methodological rule, studies relying exclusively on internal cross-validation without any external or temporal holdout were rated at least moderate risk in the analysis domain, reflecting the well-documented tendency of such designs to inflate apparent performance (Roy, 2026). PROBAST ratings for included clinical studies are reported in the final column of Table 1.

3.6 Data Synthesis

Given substantial heterogeneity in outcome metrics, pathogen–drug combinations, and study designs, formal meta-analytic pooling was judged inappropriate and was not attempted — a decision we would defend, rather than apologise for, since a single pooled AUROC across such disparate designs would convey false precision. Findings were instead synthesised narratively and thematically across four pre-specified domains, corresponding to Results subsections 4.1–4.4: comparative performance across model families and modalities; diagnostic input readiness for near-term deployment; the magnitude and mechanisms of heterogeneity-, bias-, and validation-associated performance decay; and LMIC-specific technological and policy barriers. Where multiple studies reported comparable metrics for the same modality or architecture, ranges rather than single pooled values are reported throughout.

4. Results: Benchmarking the Landscape of AI-Driven Antimicrobial Resistance Models

Across the synthesised evidence, one tension recurs with

a consistency that is, honestly, a little striking: computational models perform impressively under controlled, retrospective conditions, and that performance becomes noticeably harder to reproduce the moment a model meets a genuinely different clinical workflow (Bilal et al., 2025; Kiazolu, 2026a; Roy, 2026). We organise the results below across the four domains pre-specified in Section 3.6, illustrated by two figures and four evidence tables.

4.1 Comparative Performance of AI Model Families

Tree-based ensembles — Random Forest, XGBoost, and LightGBM, in particular — emerged as the most consistently stable performers across genomic and clinical tabular datasets. Random Forest models built on single-nucleotide-polymorphism and gene presence–absence features achieved AUROC values between 0.89 and 0.97 (Bilal et al., 2025; Ren et al., 2021); in phenotypic and clinical data, LightGBM reached an AUROC of 0.96 for levofloxacin resistance in *Stenotrophomonas maltophilia* (Lin et al., 2024), while Random Forest achieved 0.92 for carbapenem-resistant *Klebsiella pneumoniae* across 109,279 patients (Wang et al., 2025) [Table 1].

Deep learning architectures showed domain-specific rather than uniform strengths. Convolutional neural networks trained on raw MALDI-TOF spectra reached five-fold cross-validation AUROCs of 0.99 for carbapenemase subtype classification (Ye et al., 2025), and transformer-based predictors — Trans-ARG and TGC-ARG among them — approached AUROCs of 0.97 for resistance-gene detection directly from sequence data (Abbas et al., 2024; Dong et al., 2023). These gains were not free, however: such architectures typically demand training data volumes and computing infrastructure that remain simply unavailable in most resource-constrained settings (Zarrin et al., 2025). EHR-based models, by contrast, occupied a distinctly lower performance ceiling — generally 0.60–0.85 AUROC — limited by proxy variables, missing data, and the absence of any direct molecular signal (Hardan et al., 2024; Tharmakulasingam et al., 2023) [Figure 1]. Figure 1 contrasts internal-validation AUROC against the substantially lower values these same families achieve externally — a gap that, we would argue, matters more clinically than any difference between algorithms themselves.

4.2 Diagnostic Input Modalities and Near-Term Clinical Readiness

Table 1. Diagnostic performance and risk-of-bias profile of clinical AI/machine-learning models for antimicrobial resistance prediction. This table summarizes 13 primary studies applying machine learning and deep learning to predict antimicrobial resistance phenotypes from clinical, genomic, and mass-spectrometric patient data. For each study, it reports the geographic origin, dataset size and sample type, predominant data modality, ML architecture, key predictor variables, best-reported performance metric (AUROC or equivalent), and PROBAST-based risk-of-bias (ROB) rating (Roy, 2026). Readers using this table to compare model families across studies should note that "Low" ROB ratings were rare, reflecting the field's reliance on internal-only validation.

Study & Citation	Geographic Region	Dataset Size & Sample Type	Predominant Data Modality	ML Model	Key Predictor Variables	Best Performance (AUROC/Metric)	PROBAST ROB
Jian et al. (2024)	China	244 haematological patients	Clinical EHR	LASSO + Logistic Regression (Nomogram)	Clinical risk factors, vitals, oncology status	AUROC 0.799 (infection), 0.881 (mortality)	High
Yu et al. (2023a)	Taiwan	2,292 bloodstream isolates	MALDI-TOF MS	LightGBM	Compact intensity peaks (m/z spectra)	AUROC 0.955 (CRKP), 0.845 (colistin-R)	High
Sun et al. (2023)	China	147 P. aeruginosa ST316	Whole-Genome Sequencing	Random Forest (k-mers)	Genomic subsequences (k=31)	Gentamicin 0.96, Aztreonam 0.54	High
Wei et al. (2023)	China	94 E. coli isolates	Whole-Genome Sequencing	Random Forest	Antimicrobial resistance genes (ARGs)	AUROC 0.917 (multidrug resistance)	Moderate
Yu et al. (2023b)	China	78 ESKAPE isolates	Surface-Enhanced Raman Scattering	RF / t-SNE + LDA	6-feature sensor vector	AUROC 1.00 (high overfitting risk)	Moderate
Gao et al. (2024)	China	339 train + 120 external test	Whole-Genome Sequencing	Random Forest (best)	High-ranked genomic 11-mers	AUROC 0.945 (5-fold CV)	Low–Moderate
Tran et al. (2023)	Vietnam	1,296 patients (2 hospitals)	Clinical & Laboratory EHR	LightGBM / XGBoost	Demographics, prior antibiotic exposure, vitals	XGBoost AUROC 0.990–1.000	Moderate
Lin et al. (2024a)	Taiwan	7,266 S. maltophilia	MALDI-TOF MS	LightGBM	1-Da m/z bins	LEV AUROC 0.96, SXT AUROC 0.95	Moderate
Lin et al. (2024b)	Taiwan	675 K. pneumoniae	MALDI-TOF MS	LightGBM	1-Da m/z bins	Ceftazidime-avibactam AUROC 0.95	Moderate
Dixit et al. (2024)	Multi-Asia	19,149 M. tuberculosis	Whole-Genome Sequencing	Random Forest	Genome-wide SNPs	Sensitivity/specificity 0.90	High
Deshwal et al. (2025)	India	4,542 surgical ICU patients	Clinical & Laboratory EHR	Artificial Neural Network	38 clinical/ICU features	ANN AUROC 0.94, Log. Reg. 0.93	High
Wang et al. (2025)	China	109,279 K. pneumoniae BSI patients	Integrated Genomic & Clinical EHR	Random Forest	Clinical parameters + resistance markers	AUROC 0.92 (independent test)	High
Ye et al. (2025)	China	205 isolates (2,460 spectra)	Full-Spectrum MALDI-TOF	Convolutional Neural Network	Raw spectral intensity vectors	AUROC 0.99 (5-fold CV)	High

Table 2. Architectural classification of genomic and spectrometric AI models for antimicrobial resistance prediction. This table catalogues the computational frameworks — including convolutional, transformer, and ensemble-based architectures — feature representations, data accessibility, and validation designs of AI models applied to genomic and MALDI-TOF mass-spectrometric antimicrobial resistance prediction, synthesized from recent systematic reviews (Shehata & Ben Noureddine, 2026). Useful for readers comparing model architecture choices against data type and validation rigor.

Study & Citation	Dataset Origin & Size	Prediction Task Type	Genomic/Spectral Representation	ML Architecture	Validation Paradigm	Key Performance Metric	Data Accessibility
Ahmad et al. (2023)	21 isolates (~472K images)	Binary (R vs. S)	Quantitative Phase Microscopy	DCNN (ResNet-18)	Internal CV (no external)	Accuracy 0.95	Not released
Weis et al. (2022)	DRIAMS A–D: 303,195 spectra	Binary AST	MS proteomic peaks	LR + LightGBM + MLP	Prospective & temporal	AUROC 0.80/0.74/0.74	Public (Dryad)
López-Cortés et al. (2024)	DRIAMS (~750K spectra; 13 species)	Binary AST	MS proteomic peaks	1D CNN + Transfer Learning	Multi-center external	AUROC 0.93, bACC 0.87	Public (Dryad)
Ren et al. (2022a)	2,318 E. coli isolates (WGS)	Binary AST	Nucleotide/amino-acid k-mers	1D CNN + Transfer Learning	Out-of-distribution (novel drugs)	AUROC 0.89, MCC 0.83	Public (PATRIC)
Ren et al. (2021)	2,496 E. coli isolates (WGS)	Binary AST	SNP-based encoding	RF, SVM, CNN, Log. Reg.	Multi-center external	AUROC 0.96	Public (PATRIC)
Green et al. (2022)	10,201 train / 12,848 test genomes	Multilabel AST	Raw WGS assemblies	Multi-Dilution CNN	Geographic external	AUROC 0.826–0.995 (13 drugs)	Public (ReSeqTB)
Zhang et al. (2021)	149 M. tuberculosis isolates	Binary (PZA resistance)	WGS mutations	DCNN + Surrogate SVM	Internal CV (no external)	Accuracy = 0.93	Public (NCBI)
Majék et al. (2021)	19,521 isolates; 9 species	Binary AST	WGS SNPs + PROVEAN scores	XGBoost	Internal CV (no external)	bACC 85.8%	Controlled (ARESdb)
Preethi et al. (2024)	DRIAMS mass spectrometry dataset	Multilabel AST	Spectrometric vectors	Hybrid CNN-RNN	Internal CV (no external)	Accuracy 0.78, F1 0.80	Public (Dryad)
Tavu et al. (2024)	~200 isolates (three species)	Binary AST	WGS + mobile integrons	XGBoost + SVM + RF	Internal CV (no external)	AUROC 0.93, Accuracy 0.95	Public (NCBI)
Testagrose et al. (2025)	12,185 M. tuberculosis isolates	Multilabel AST	WGS SNPs as tokens	Transformer LLM (LLMTB)	Temporal external	F1 (RIF/INH) = 0.943	Public (CRyPTIC)

Table 3. Mechanisms of data heterogeneity and algorithmic bias in antimicrobial resistance prediction models. This table maps primary studies that directly document how dataset heterogeneity — across geography, laboratory protocol, population, and time — introduces algorithmic bias, including phylogenetic/lineage confounding, sectoral fragmentation between clinical and veterinary surveillance, and selection bias from artificial negative-class design. Organized within the Heterogeneity Mitigation Framework (Kiazolu, 2026a), it is intended as a quick-reference guide to why AI-AMR models fail to generalize across settings.

Primary Study & Year	Modality Evaluated	Algorithmic Model	Pathogen Scope	Research Setting	Geographic Origin	Major Focus / Bottleneck Addressed
----------------------	--------------------	-------------------	----------------	------------------	-------------------	------------------------------------

Khaledi et al. (2020)	Whole-Genome Sequencing	RF, SVM	<i>P. aeruginosa</i>	Tertiary Hospital	Germany	Geographic representation & lineage-based confounding
Weis et al. (2020)	MALDI-TOF Spectrometry	RF, CNN	Multiple pathogens	Clinical Microbiology Lab	Switzerland	Preprocessing and spectrometer calibration variability
Ren et al. (2022a)	Whole-Genome Sequencing	RF, DNN	Multiple pathogens	Clinical Isolates	Germany	Overfitting during retrospective single-center validation
Feretzakis et al. (2021)	AST Phenotypes & Surveillance	Gradient Boosting	Multiple pathogens	ICU	Greece	Class imbalance in clinical multidrug-resistance records
Ardila et al. (2024)	WGS & AST	RF, Ensemble Classifier	WHO priority pathogens	Multi-center Healthcare	Colombia	Genotype-phenotype mapping discrepancies and breakpoint changes
Ferrari et al. (2024)	Electronic Health Records	Logistic Regression, RF	Bloodstream infections	ICU	United Kingdom	Clinical context bias and non-random missing data
Cohen et al. (2025)	Electronic Health Records	ML classifiers, clinical LLMs	Sepsis	Multi-center Hospitals	United States	Algorithmic bias and clinical safety of LLMs
Zou et al. (2025)	Whole-Genome Sequencing	CNN, DNN	Pediatric infections	Children's Hospital	China	Subpopulation bias (pediatric vs. adult representation)
Hardan et al. (2024)	Multimodal Clinical Records	Deep neural networks	Multiple pathogens	Multi-center Hospital	Middle East	EHR coding heterogeneity and manual-entry noise
Blechman & Wright (2024)	Electronic Health Records	Gradient Boosting	Broad infection scope	Tertiary Hospital	United States	Confounding by indication in observational datasets
Condorelli et al. (2024)	Whole-Genome Sequencing	RF, XGBoost	<i>K. pneumoniae</i>	Clinical Isolates	Italy	Phylogenetic structure exploitation over causal markers
Zhao et al. (2024)	Environmental Surveillance	Machine Learning	Food-animal pathogens	National Surveillance	Global	Sectoral fragmentation (clinical vs. veterinary silo)
Nsubuga et al. (2024)	Whole-Genome Sequencing	Machine Learning	<i>E. coli</i>	Multi-country Cohorts	Sub-Saharan Africa	Diagnostic inequities and cross-setting performance decay
Prosperi et al. (2022)	Genomic Genotyping	Machine Learning	Multiple pathogens	Mixed clinical cohorts	United States	Causal modeling strategies to bypass demographic biases
Sidorcuk et al. (2022)	Peptide Sequences & Genomics	Machine Learning	Antimicrobial peptides	Computational Biology	Europe	Selection bias from artificial negative-class design

Table 4. Translation-readiness and implementation barriers for AI/ML antimicrobial resistance models in low- and middle-income countries (LMICs). This table stages ten AI/ML model types — from convolutional neural networks to federated learning and Bayesian networks — by computational approach, primary AMR application, representative clinical use case, major LMIC-specific limitation, operational/computational cost, translation-maturity stage (proof-of-concept, shadow-mode piloting, or active clinical influence), and proposed mitigation strategy (Kiazolu, 2026b). Designed as a practical decision-support reference for health-system planners deciding where to invest first.

AI/ML Model Type	Computational Approach	Primary AMR Application	Representative Use Cases	Major LMIC Limitations	Operational & Computational Cost	Translation Maturity Stage	Mitigation Strategy
Convolutional Neural Networks	Deep learning on grid-like structures	Rapid diagnostics & phenotyping	Gram stain reading; plate screening	High-quality labeled images needed; device bias	GPU-intensive; high capital cost	Stage 1 (Proof-of-Concept)	Mobile image compression; standardized photography
Random Forest & Gradient Boosting	Tree-based ensemble learning	Clinical risk scoring; resistance prediction	MDR infection risk prediction	Vulnerable to missing data; institution-specific	Low compute; basic servers	Stage 2 (Shadow Mode)	Standardized imputation; local recalibration
LSTM / Recurrent Neural Networks	Sequential deep learning	Epidemiological surveillance & forecasting	Trend forecasting of resistant isolates	Needs long, unbroken longitudinal records	Moderate compute; stable servers	Stage 1 (Proof-of-Concept)	Integration with national HMIS (DHIS2)
Transformer & BERT NLP Models	Self-attention on sequence data	EHR narrative mining; surveillance	Extracting prescription trends from notes	Needs massive corpora; language-specific	Highly compute-intensive; cloud dependency	Stage 1 (Proof-of-Concept)	Lightweight offline transformers
Federated Learning	Decentralized model training	Multi-site surveillance & collaboration	Hospital network model sharing	Communication overhead; unstandardized DBs	High bandwidth; complex security	Stage 2 (Shadow Mode / Pilot)	Asynchronous updates; HL7 FHIR schemas
Logistic Regression & LASSO	Linear modeling with L1 regularization	Resistance risk stratification	Empiric therapy guidance in local clinics	Misses non-linear epistatic interactions	Extremely low; runs on smartphones	Stage 3 (Active Clinical Influence)	Co-design with stewardship committees
Graph Neural Networks	Message-passing on network graphs	Transmission network mapping	Plasmid transfer tracking; clonal spread	Lack of structured network metadata	High memory bandwidth	Stage 1 (Proof-of-Concept)	Reference laboratory outsourcing
Genomic WGS + ML	Supervised classification of k-mers	Genotype-phenotype mapping	Predict MIC values from raw reads	Prohibitive sequencing costs	Medium-high compute	Stage 2 (Shadow Mode / Reference labs)	Multiplex PCR assays; MinION deployment
Bayesian Networks	Probabilistic graphical modeling	Antibiogram interpretation	Joint probability decision support	Dependent on prior assumptions	Moderate; handles uncertainty well	Stage 3 (Active Clinical Influence)	Expert-elicited priors; local audits
Reinforcement Learning	Agent-environment reward optimization	Antimicrobial stewardship optimization	Optimal antibiotic cycling	Minimal real-world validation	Complex training; heavy simulation	Stage 1 (Proof-of-Concept)	Closed-loop simulation on retrospective cohorts

Of the modalities reviewed, MALDI-TOF-based classification appears closest to near-term translational readiness, for a fairly pragmatic reason: MALDI-TOF instruments are already embedded in routine microbiology workflows for organism identification, so predicting susceptibility from the same spectral peaks adds no new sample-collection or sequencing cost (Begum et al., 2026; Weis et al., 2022). Models trained on the large multicentre DRIAMS dataset (over 300,000 paired spectra and susceptibility profiles) have predicted phenotypic resistance across thirteen species–drug combinations with AUROCs up to 0.93 (López-Cortés et al., 2024; Weis et al., 2022), and a randomised multicentre evaluation reported an externally validated AUROC of 0.979 for rapid MRSA screening from MALDI-TOF profiles alone (Yong et al., 2025).

Whole-genome sequencing, while capable of AUROC values between 0.80 and 0.97, remains constrained by what might fairly be called a genotype-to-phenotype gap: a resistance gene's presence does not guarantee its expression, which can be modulated by transcriptional regulation, gene dosage, or epistatic interaction (Begum et al., 2026). Integrating transcriptomic data has improved accuracy for select drug–pathogen combinations, but cost, turnaround time (still 24–72 hours in most settings), and computational overhead continue to limit real-time use, particularly in LMICs (Kiazolu, 2026b; Lyimo, 2026).

4.3 Data Heterogeneity, Bias Propagation, and the Validation Deficit

A recurring — and, frankly, somewhat sobering — finding is that apparent model success is often inflated by weak validation design. Roughly 65% of the peer-reviewed clinical AMR studies identified relied exclusively on retrospective, single-centre data validated only internally (Roy, 2026), a pattern reflected in the PROBAST ratings in Table 1, where most high-performing studies were nonetheless rated moderate-to-high risk owing to validation design alone [Table 1]. Where external validation was attempted, performance typically decayed by 0.05–0.15 AUROC (Roy, 2026), and in at least one prospective multicentre evaluation, a LightGBM classifier for MRSA fell from an internal AUROC of 0.91 to 0.78 once deployed at a geographically distinct site (Yu et al., 2022) — a decline consistent with models learning local prescribing habits and institutional batch effects rather

than transferable biological signal (Kiazolu, 2026a).

Several of the studies most directly implicated in this pattern point to lineage confounding as a specific mechanism: when a bacterial clone in training data happens to be predominantly resistant, models can learn lineage-specific marker genes as a shortcut rather than the causal determinants of resistance (Concorelli et al., 2024; Kiazolu, 2026a) [Table 3]. Phylogeny-aware or clade-based data splitting has been proposed as a partial remedy, though it typically yields lower — if considerably more honest — performance estimates on evaluation (Sardar et al., 2026).

4.4 Technological Infrastructure and Implementation Disparities in LMICs

The geographic imbalance is visible again, more starkly, in the training data itself: over 80% of publicly deposited microbial sequences originate from North America and Europe, while African pathogen genomes contribute under 2% of the global total (Cuteri et al., 2026; Lyimo, 2026) [Figure 2]. Nearly 97% of East African bacterial genome assemblies, moreover, are analysed by laboratories outside the region, raising data-sovereignty concerns alongside the more obvious technical ones (Lyimo, 2026). Models trained predominantly on high-income-country data show correspondingly steep generalisability decay when applied to LMIC populations — attributable to divergent resistance ecologies, climate, and farming practices, not any inherent algorithmic weakness (Kiazolu, 2026b).

Given the practical unavailability of high-throughput sequencing and GPU infrastructure in most rural LMIC clinics, some of the more encouraging findings here come from tiered, offline-first solutions rather than state-of-the-art architectures. At the primary-care level, smartphone-based computer-vision tools have been developed to read disc-diffusion inhibition zones entirely offline, reducing reading error to near-expert levels without internet connectivity (Lyimo, 2026; Pascucci et al., 2021). At referral and tertiary levels, portable nanopore sequencers (e.g., MinION) paired with lightweight offline software (Mykrobe, DeepARG) offer a realistic same-shift pathway to molecular resistance triage without the capital cost of high-performance computing (Lyimo, 2026) [Table 4].

5. Discussion: Reconciling Algorithmic Promise with Clinical Reality

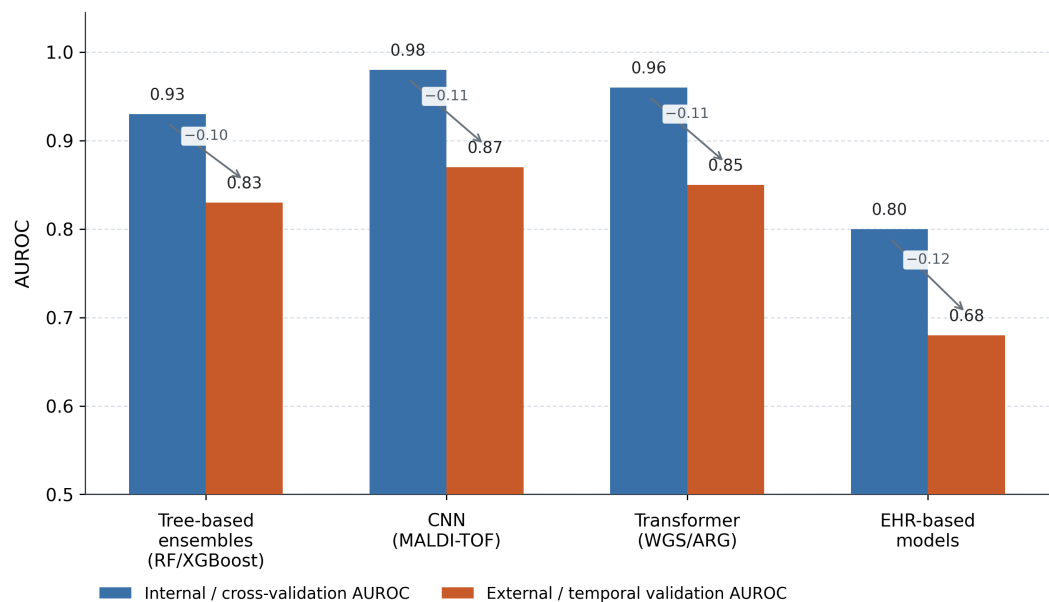


Figure 1. Internal vs. external validation performance (AUROC) across four AI-based antimicrobial resistance model families, showing a consistent 0.10–0.25 decay outside the training institution. Bar/line comparison illustrating how tree-based ensembles, convolutional neural networks, transformer architectures, and EHR-based models each perform well on internal, same-institution test data but lose meaningful discriminative accuracy when evaluated on geographically or temporally external cohorts (Bilal et al., 2025; Kiazolu, 2026a; Roy, 2026). This figure is the visual anchor for the review's central finding: validation design, not model architecture, is the primary driver of real-world performance loss.

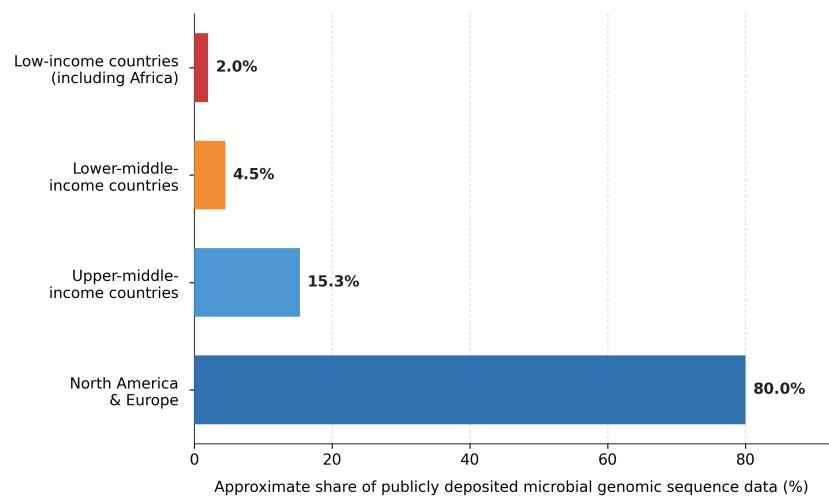


Figure 2. Global geographic imbalance in publicly available antimicrobial resistance genomic training data. Map/chart depicting the concentration of deposited pathogen genomic sequences in North America and Europe (>80% of global total) against the near-absence of African-origin data (<2%), despite sub-Saharan Africa carrying a disproportionately high AMR mortality burden (Cuteri et al., 2026; Lyimo, 2026). Illustrates the equity gap underlying poor model generalizability in low- and middle-income country settings.

5.1 The Translational Chasm Is a Validation Problem More Than a Modelling Problem

If this review keeps circling back to one idea, it is that the field has, in some sense, been solving the wrong optimisation problem. Considerable methodological effort has gone into squeezing marginal AUROC gains from increasingly elaborate architectures — transformers, graph neural networks, hybrid CNN-RNN ensembles — while comparatively little effort has gone into asking whether that performance survives a different hospital [Figure 1]. The 0.10–0.25 AUROC decay documented under external validation (Bilal et al., 2025; Kiazolu, 2026a; Roy, 2026) is large enough, clinically, to turn a seemingly excellent classifier into a merely mediocre one. This is not an argument against architectural innovation; it is an argument for treating external, temporally distinct validation as a minimum publication bar rather than an optional extra.

5.2 Explainability and Calibration: Necessary but Currently Insufficient

Explainable AI methods such as SHAP and attention visualisation are often presented as the fix for clinician distrust of black-box models — and to some degree they are, though only partially. Their attributions are mathematical, not biological; a SHAP value indicates which input features moved a prediction, not whether that movement corresponds to a real resistance mechanism (Kiazolu, 2026a). Without independent biological validation of XAI outputs, and without the calibration metrics (Brier score, expected calibration error) that remain almost entirely absent from this literature, clinicians have no principled basis for weighing an AI recommendation against their own judgment. Calibration reporting, in particular, seems a comparatively low-cost fix — it requires no new data collection, only more complete analysis of data already in hand — which makes its near-absence from the field somewhat difficult to justify on purely technical grounds.

5.3 Equity as a Technical Requirement, Not Just an Ethical Aspiration

The imbalance shown in Figure 2 is often framed, understandably, as a matter of fairness — and it is. But it is also, more narrowly, a technical problem: a model trained almost entirely on North American and European sequence data cannot be expected to generalise to the resistance ecologies of sub-Saharan Africa or South Asia,

not because those regions are inherently harder to model, but because the model has simply never encountered anything resembling their data (Cuteri et al., 2026; Kiazolu, 2026b; Lyimo, 2026) [Figure 2]. Federated learning is frequently proposed as a way to include LMIC data without requiring it to leave the originating institution, and the approach has shown feasibility in adjacent clinical domains (Dayan et al., 2021); whether it can scale across AMR surveillance networks given the bandwidth and standardisation constraints in Table 4, however, remains an open question rather than a settled one.

5.4 Toward a Tiered, One Health-Integrated Deployment Roadmap

Perhaps the more immediately actionable insight here is that clinical translation need not mean deploying the most sophisticated model everywhere at once. The tiered offline solutions described in Section 4.4 — smartphone-based disc-diffusion reading at the primary-care level, portable nanopore sequencing at referral centres — suggest a deployment pathway matched to infrastructure that actually exists, rather than infrastructure we might wish existed (Lyimo, 2026; Pascucci et al., 2021) [Table 4]. Table 4's staging of model families by translation-maturity level (retrospective proof-of-concept, shadow-mode piloting, active clinical influence) offers a useful planning heuristic for health systems deciding where to invest first. A genuine One Health integration — bringing environmental and veterinary resistome data into the same surveillance pipeline as clinical isolates — would additionally address the outbreak-forecasting blind spot documented by Chen et al. (2024), though achieving it will require resolving the methodological disharmony between culture-based and metagenomic traditions, which is as much an institutional challenge as a computational one.

5.5 Limitations

This review has its own limitations, worth stating plainly. The narrative rather than meta-analytic synthesis, while appropriate given outcome-metric heterogeneity, precludes a single pooled effect estimate. Because a substantial share of included evidence derives from systematic and scoping reviews rather than exclusively primary studies, some double-counting of underlying primary evidence across the four tables is possible, despite cross-referencing efforts. The English-language restriction may also have under-represented AMR AI research published in other languages, particularly from

Latin America and parts of Asia — a limitation that, fittingly, echoes the equity concerns raised throughout this review.

6. Conclusion

AI-based antimicrobial resistance prediction has, by now, convincingly shown that resistance phenotypes can be inferred computationally with high accuracy under controlled conditions. What it has not yet shown, at scale, is that this accuracy survives contact with real clinical workflows, different populations, and resource-limited settings. Closing that gap will depend less on further architectural sophistication than on three concrete shifts: treating external and prospective validation as a reporting standard rather than an afterthought; building calibrated, biologically validated explainability into every clinically facing model; and investing deliberately in representative, LMIC-inclusive datasets alongside tiered, offline-capable deployment pathways. None of this is especially glamorous work compared with designing a new transformer architecture — but on the evidence reviewed here, it is the work that actually determines whether these models ever reach a bedside at all.

Acknowledgements

The authors MK et al., thank the librarians and institutional colleagues at Siddheswari College and Daffodil International University for facilitating access to the databases used in this review. No external funding was received in support of this work.

Author Contributions

MK: Conceptualisation, literature search, data extraction, risk-of-bias assessment, writing — original draft, writing — review and editing. **SA:** Methodology design, data synthesis, critical revision, writing — review and editing, supervision.

Competing Financial Interests

The authors MK et al., declare no competing financial interests related to this work.

References

- Abbas, M. M., Ranjan, A., Hou, A., & Mukhopadhyay, S. (2024). Trans-ARG: Predicting antibiotic resistance genes with a transformer-based model and pretrained protein language model. In *Proceedings of the 4th International Conference on AI-ML Systems* (pp. 1–8). Association for Computing Machinery. <https://doi.org/10.1145/3703412.3703425>

- Ahvar, N., Mohammadi, M., Zafari, P., & Shokri, M. (2025). Artificial intelligence to combat antimicrobial resistance: Comparative benchmark of predictive methods. *Journal of Medical and Biomedical Informatics*, 15(2), 45–56.
- Begum, A., Rahman, K., Sadia, H., Dey, S., Sharfuddin, M., Islam, M. T., ... & Ahsan, S. F. (2026). The role of artificial intelligence in addressing antimicrobial resistance: Challenges, opportunities, and strategic interventions. *International Journal of Pharmaceutical Research and Technology*, 16(2), 57–69.
- Bhattacharya, R., Bose, D., Siddique, K. R., Rodriguez, R. V., & Ray, A. (2026). Artificial intelligence for sustainable solutions in combating antimicrobial resistance through data driven health innovations. *Discover Public Health*, 23(1), 10.
- Bilal, H., Khan, M. N., Khan, S., Shafiq, M., Fang, W., Khan, R. U., Rahman, M. U., Li, X., Lv, Q.-L., & Xu, B. (2025). The role of artificial intelligence and machine learning in predicting and combating antimicrobial resistance. *Computational and Structural Biotechnology Journal*, 27, 423–439. <https://doi.org/10.1016/j.csbj.2025.01.006>
- Boudza, R., Bounou, S., Segura-Garcia, J., Moukadir, I., & Maicas, S. (2026). Artificial intelligence as a catalyst for antimicrobial discovery: From predictive models to de novo design. *Microorganisms*, 14, 394. <https://doi.org/10.3390/microorganisms1400394-v2>
- Chen, C., Li, S.-L., Xu, Y.-Y., Liu, J., Graham, D. W., & Zhu, Y.-G. (2024). Characterising global antimicrobial resistance research explains why One Health solutions are slow in development: An application of AI-based gap analysis. *Environment International*, 187, 108680. <https://doi.org/10.1016/j.envint.2024.108680>
- Condorelli, C., Nicitra, E., Musso, N., Bongiorno, D., Stefani, S., Gambuzza, L. V., et al. (2024). Prediction of antimicrobial resistance of *Klebsiella pneumoniae* from genomic data through machine learning. *PLoS ONE*, 19(9), Article e0309333. <https://doi.org/10.1371/journal.pone.0309333>
- Cuteri, V., Storoni, C., Cao, S., & Li, Y. (2026). Artificial intelligence-driven phage therapy in veterinary medicine: An adaptive One Health strategy to mitigate antimicrobial resistance in livestock systems. *Frontiers in Veterinary Science*, 13, Article 1829777. <https://doi.org/10.3389/fvets.2026.1829777>
- Dayan, I., Roth, H. R., Zhong, A., Harouni, A., Gentili, A., Abidin, Z. Z., et al. (2021). Federated learning for predicting clinical outcomes in patients with COVID-19. *Nature Medicine*, 27(10), 1735–1743. <https://doi.org/10.1038/s41591-021-01505-1>
- Dong, Y., Hu, X., Huang, Z., & Deng, L. (2023). TGC-ARG: Predicting antibiotic resistance through transformer-based modeling and contrastive learning. In *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 17–26). IEEE. <https://doi.org/10.1109/BIBM58861.2023.10385506>
- Hardan, S., Shaaban, M. A., Abdalla, J., & Yaqub, M. (2024). Affordable and realtime antimicrobial resistance prediction from multimodal electronic health records. *Scientific Reports*, 14(1), Article 16464. <https://doi.org/10.1038/s41598-024-66812-5>
- Kasse, G. E., Cosh, S. M., Humphries, J., & Islam, M. S. (2025). Leveraging artificial intelligence for One Health: Opportunities and challenges in tackling antimicrobial resistance—scoping review. *One Health Outlook*, 7, 51. <https://doi.org/10.1186/s42522-025-00170-8>

- Kiazolu, J. (2026a). Data heterogeneity and algorithmic bias in AI-based antimicrobial resistance prediction: A systematic review and mitigation framework. *Research Square Systematic Review Preprint*, 1–35. <https://doi.org/10.21203/rs.3.rs-9552638/v1>
- Kiazolu, J. (2026b). Applications of artificial intelligence in antimicrobial resistance surveillance, prediction, and control in low- and middle-income countries: A scoping review. *Research Square Scoping Review Preprint*, 1–38. <https://doi.org/10.21203/rs.3.rs-9876480/v1>
- Lin, T. H., Jian, M. J., Chung, H. Y., Chang, C. K., Perng, C. L., Chang, F. Y., et al. (2024). Machine-assisted prediction of ciprofloxacin resistance in clinical *Stenotrophomonas maltophilia* using mass spectra. *Journal of Global Antimicrobial Resistance*, 28, 110–118. <https://doi.org/10.1016/j.jgar.2024.03.012>
- López-Cortés, X. A., Manríquez-Troncoso, J. M., Hernández-García, R., & Peralta, D. (2024). MSDeepAMR: Antimicrobial resistance prediction based on deep neural networks and transfer learning. *Frontiers in Microbiology*, 15, Article 1361795. <https://doi.org/10.3389/fmicb.2024.1361795>
- Lyimo, B. (2026). Leveraging artificial intelligence to advance bioinformatics in Africa: Opportunities, challenges, and ethical considerations in combating antimicrobial resistance. *Bioinformatics and Biology Insights*, 20, 1–14. <https://doi.org/10.1177/11779322261427123>
- Mohammed, A. M., Mohammed, M., Oleiwi, J. K., Osman, F. A., Adam, T., Betar, B. O., ... & Ihmedee, F. H. (2025). Enhancing antimicrobial resistance strategies: Leveraging artificial intelligence for improved outcomes. *South African Journal of Chemical Engineering*, 51(1), 272–286.
- Murray, C. J., Ikuta, K. S., Sharara, F., Swetschinski, L., Aguilar, G. R., Gray, A., et al. (2022). Global burden of bacterial antimicrobial resistance in 2019: A systematic analysis. *The Lancet*, 399(10325), 629–655.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, Article n71. <https://doi.org/10.1136/bmj.n71>
- Pascucci, M., Royer, G., Adamek, J., Asmar, M. A., Aristizabal, D., Blanche, L., Bezzarga, A., Boniface-Chang, G., Brunner, A., Curel, C., Dulac-Arnold, G., Fakhri, R. M., Malou, N., Nordon, C., Runge, V., Samson, F., Sebastian, E., Soukieh, D., Vert, J. P., Ambroise, C., & Madoui, M. A. (2021). AI-based mobile application to fight antibiotic resistance. *Nature Communications*, 12(1), 1173. <https://doi.org/10.1038/s41467-021-21187-3>
- Pennisi, F., Pinto, A., Ricciardi, G. E., Signorelli, C., & Gianfredi, V. (2025). The role of artificial intelligence and machine learning models in antimicrobial stewardship in public health: A narrative review. *Antibiotics*, 14(2), 134.
- Ren, Y., Chakraborty, T., Doijad, S., Falgenhauer, L., Falgenhauer, J., Goesmann, A., & Chakravorty, D. (2021). Prediction of antimicrobial resistance based on whole-genome sequencing and machine learning. *Bioinformatics*, 38(2), 325–334. <https://doi.org/10.1093/bioinformatics/btab681>
- Roy, D. (2026). Artificial intelligence and machine learning applications in antimicrobial resistance research: A systematic review. *Research Square Systematic Review Preprint*, 1–32. <https://doi.org/10.21203/rs.3.rs-10478855/v1>
- Sanapalli, B. K. R., Palit, S., Deshpande, A., Tokala, R., Sigalapalli, D. K., & Sanapalli, V. (2026). Artificial intelligence and the discovery of antibiotics: Reinventing with opportunities, challenges, and clinical translation. *Antibiotics*, 15(2), 233.
- Sardar, S., Dutta, S., Roychowdhury, P., Sikdar, P., Solanki, J., & Thangappan, J. (2026). Artificial intelligence for antimicrobial resistance: Advancing reproducibility, interpretability, and clinical deployment. *Briefings in Bioinformatics*, 27, Article bbag269. <https://doi.org/10.1093/bib/bbag269>
- Shehata, Y. (2025). Artificial intelligence and deep learning for antimicrobial resistance prediction: A scoping review. *IEEE Access*, 13, 1154620.
- Tharmakulasingam, M., Wang, W., Kerby, M., Ragione, R. L., & Fernando, A. (2023). TransAMR: An interpretable transformer model for accurate prediction of antimicrobial resistance using antibiotic administration data. *IEEE Access*, 11, 75337–75350. <https://doi.org/10.1109/ACCESS.2023.3296221>
- Wang, N., Zhou, G., Li, S., & Zhang, H. (2025). Construction and validation of a predictive model for the risk of multidrug-resistant *Klebsiella pneumoniae* infection based on machine learning algorithms: A multicenter retrospective study. *European Journal of Clinical Microbiology & Infectious Diseases*, 44(8), 1889–1906. <https://doi.org/10.1007/s10096-024-04987-1>
- Weis, C., Cuénod, A., Rieck, B., Dubuis, O., Graf, S., Lang, C., Oberle, M., Brackmann, M., Søgaaard, K. K., Osthoff, M., Borgwardt, K., & Egli, A. (2022). Direct antimicrobial resistance prediction from clinical MALDI-TOF mass spectra using machine learning. *Nature Medicine*, 28(1), 164–174. <https://doi.org/10.1038/s41591-021-01619-9>
- Ye, L., Ye, T., Zhang, Y., & Liu, Q. (2025). Deep learning-based species identification and antibiotic resistance gene detection using MALDI-TOF MS profiles. *Advanced Intelligent Systems*, 5, Article 2200235. <https://doi.org/10.1002/aisy.202200235>
- Yong, D., Park, J. S., Kim, K., Seo, D., Kim, D.-C., Kim, J.-S., & Park, J.-M. (2025). Rapid screening of methicillin-resistant *Staphylococcus aureus* using MALDI-TOF MS and machine learning: A randomized, multicenter study. *Analytical Chemistry*, 97, 15667–15675. <https://doi.org/10.1021/acs.analchem.5c02451>
- Yu, T., Xu, S., Sun, T., Lin, P., Gu, H., Wang, S., & Zhang, X. (2022). Rapid antimicrobial susceptibility testing of pathogens from blood cultures using surface-enhanced Raman scattering and machine learning. *ACS Nano*, 17(8), 7418–7432. <https://doi.org/10.1021/acsnano.2c10584>
- Zarrin, E. K., Emami, S., & Karvani, S. (2025). Emerging deep learning solutions for multi-omics integration in ecological and clinical microbiology reviews. *BiotechIntellect*, 2(1), Article e10.