



AI-Assisted Prediction of Antimicrobial Resistance Evolution in Emerging Bacterial Pathogens

Zehra Fatima^{1*}, Shehr Hayat², Maham Rafiq³, Aiman Nishat⁴ and Sadia Amjad⁵

¹Department of Zoology, University of Sargodha, Pakistan.

²Department of Life Sciences, University of Management and Technology, Lahore, Pakistan.

³Central Park Medical College, Lahore, Pakistan.

⁴Institute of Molecular Biology and Biotechnology (IMBB), University of Lahore, Pakistan.

⁵Jambi University, Indonesia.

Article Info

Keywords:

Antimicrobial resistance;
Artificial intelligence;
Resistance evolution

Date:

Received: 6 April 2026

Accepted: 18 May 2026

Published: 15 June 2026

Volume: 04 (01), 2026



Copyright: © 2026 by the authors. This article is distributed under the Creative Commons Attribution (CC BY) license, allowing for open access and sharing under specified terms and conditions. (<https://creativecommons.org/licenses/by/4.0/>).

Abstract

*Antimicrobial resistance (AMR) is a dynamic evolutionary process in which bacteria acquire and disseminate resistance through mutation, recombination, and horizontal gene transfer. This review synthesizes machine learning (ML) and deep learning (DL) applications in AMR diagnostics, genomic and metagenomic surveillance, and antibiotic discovery, emphasizing prediction of resistance evolution rather than classification of its current state. We map genomic, spectral, imaging, metagenomic, and clinical data to predictive architectures, including tree ensembles, convolutional and recurrent networks, graph neural networks, transformer-based genomic encoders, and generative models. A 57-study systematic review of *Klebsiella pneumoniae* found that 74.3% of AUROC-primary studies reported AUROC ≥ 0.90 , yet validation remained predominantly retrospective: internal cross-validation was used in 96.5% of studies, compared with <30% external validation, 15.8% temporal validation, and only one prospective evaluation. High risk of bias affected 78.9% of studies, while approximately half relied on automated susceptibility testing rather than reference-grade broth microdilution. Fewer than 10% of AI-microbiology studies originated from low- and middle-income countries. These findings expose the gap between cross-sectional classification and evolution-aware prediction. We propose a six-component roadmap integrating longitudinal genomic surveillance, phylogeny-aware representation, temporal and relational architectures, forecast-oriented validation, federate learning, and explainable AI with prospective laboratory validation to anticipate resistance and accelerate countermeasure discovery.*

1. Introduction

Antimicrobial resistance was associated with an estimated 4.95 million deaths globally in 2019 (Murray *et al.*, 2022), and forecasts based on the 2021 Global Burden of Disease study project that AMR-related mortality could rank among the leading causes of death worldwide by 2050 (Naghavi *et al.*, 2024), with some earlier projections placing the annual toll as high as ten million deaths by mid-century (de Kraker *et al.*, 2016). Unlike many other global health threats, AMR

is fundamentally an evolutionary phenomenon. Bacterial populations under antibiotic selection pressure acquire resistance through point mutation, gene amplification, plasmid-mediated horizontal gene transfer, and recombination, and these mechanisms interact with clonal expansion and geographic spread to produce resistance trajectories that differ markedly between species, clinical settings, and time periods. The co-occurrence of metal and antibiotic resistance genes on shared mobile genetic

Corresponding Author: Zehra Fatima

Email: zeea25143@gmail.com

elements, for example, has been shown to amplify the dissemination of resistance beyond what either pressure would produce alone (Bombaywala *et al.*, 2021), illustrating that resistance emergence is rarely reducible to a single gene or a single antibiotic exposure.

Conventional antimicrobial susceptibility testing (AST) - broth microdilution, disc diffusion, and automated platforms such as VITEK 2 or BD Phoenix -reports the resistance phenotype of an isolate as it exists at the moment of testing, typically 36 to 72 hours after sample collection (Aggarwal *et al.*, 2026). These methods were not designed to anticipate how a lineage's resistance profile will change over subsequent transmission cycles, nor to flag which currently susceptible strains are genomically primed to acquire resistance next. Artificial intelligence has been proposed repeatedly as a tool for accelerating and improving AMR detection, and a substantial and rapidly growing literature now documents machine learning and deep learning models that classify resistance phenotypes from whole-genome sequencing (WGS) data, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) spectra, vibrational spectroscopy, and microscopy images, frequently reporting areas under the receiver operating characteristic curve (AUROC) above 0.90 (Li *et al.*, 2024; Hassan *et al.*, 2026).

The great majority of this work, however, treats resistance prediction as a cross-sectional classification problem: given a snapshot of an isolate's genotype or phenotype, predict its current susceptibility category. Considerably less attention has been paid to the complementary and arguably more consequential problem of forecasting resistance evolution - using genomic, phylogenetic, and epidemiological data collected over time to anticipate which resistance determinants are likely to emerge, spread, or combine next within a given pathogen population (Arnold *et al.*, 2025; Khangarot *et al.*, 2026). This gap is most consequential for emerging or newly prioritized bacterial pathogens, whose surveillance history is comparatively short, whose genomic reference base is still being assembled, and whose resistance landscape may be actively shifting as the organism adapts to new hosts, environments, or treatment regimens. This review synthesizes recent syntheses and primary studies of AI applications in AMR to build a focused account of what is known, and what remains unresolved, about using AI to predict the evolutionary trajectory of resistance in emerging bacterial pathogens.

2. Methods

To guide and help inform this review, we drew on five recent reviews of artificial intelligence applications in antimicrobial resistance: a review that adopted a narrative synthesis approach (Li *et al.*, 2024) to cover a mechanistic and methodological account of AI/ML models for predicting pathogen resistance, a PRISMA-guided systematic review of 57 studies that used AI for AMR detection and prediction in *Klebsiella pneumoniae* from the period 2018 to 2025 (Aggarwal *et al.*, 2026), a four-theme narrative review of the use of AI in bacterial diagnostics, susceptibility testing, surveillance, and antibiotic development (Hassan *et al.*, 2026), a review of AI applications across bacterial infection diagnostics, AMR surveillance, and antibiotic discovery with particular emphasis on model architecture (Arnold *et al.*, 2025), and a review of the applications of AI in microbiology spanning microbial taxonomy, AMR prediction, metagenomics, microbial ecology, and drug discovery (Khangarot *et al.*, 2026). Several attempts have been made to trace primary studies discussed in these reviews back to their original publication, when this was possible, to help address the evolutionary-forecasting focus of this synthesis.

Primary studies were chosen for discussion if they explicitly model genomic or phenotypic data that have a temporal or longitudinal or phylogenetic structure; or if they concern an organism with known instances of lineage replacement or rapid emergence of resistance; or if they used a model architecture capable, in principle, of representing sequential or relational dependencies (recurrent networks, graph neural networks, transformer-based genomic encoders, or generative models); and/or if the authors used a form of validation that was relevant to forecasting but not just classification, such as external, temporal, or prospective validation. Findings were summarized narratively rather than quantitatively, reflecting the considerable diversity in the definitions of the outcome over the literature used, and in the model design, modality of input, and design.

3. Results

3.1 Data Representations Underlying AI-Based AMR Prediction

Feature representation of whole genome or metagenomic sequence data is usually performed prior to model training, relying on one of a few representations (Table 1). Nucleotide/amino-acid k-mer counts are counts of short subsequences of DNA/protein and can easily be extended

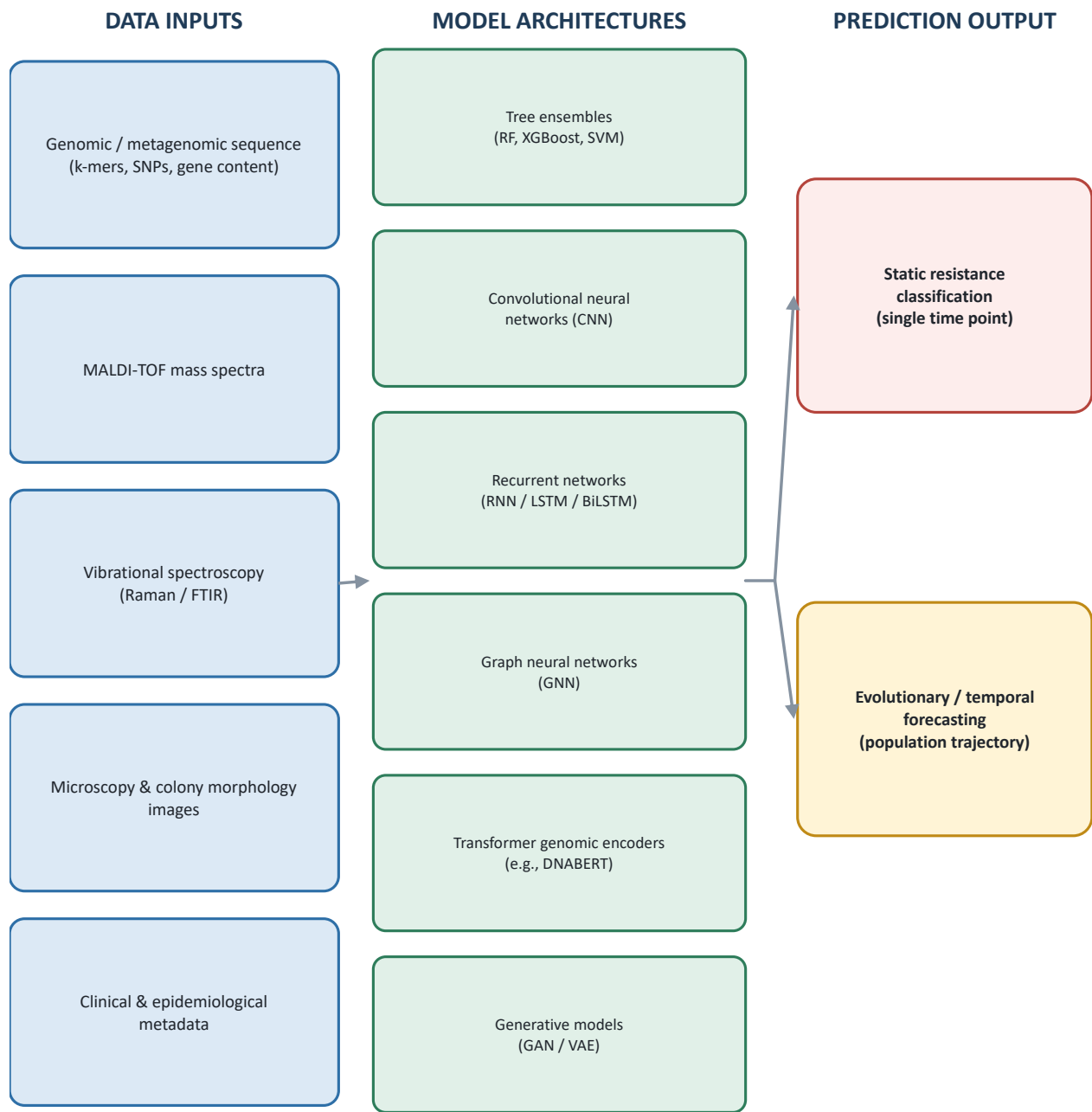


Figure 1. Conceptual framework linking AMR data modalities, AI/ML architectures, and prediction outputs. Most published models terminate in static, single-time-point classification (top output); evolutionary forecasting of resistance trajectories (bottom output) remains comparatively underdeveloped.

for novel sequences or sequences divergent from the reference, while high dimensionality of very long k-mers lowers the interpretability (Li *et al.*, 2024). Gene-content representations represent the detection—logical yes or no—

of the set of resistance associated gene clusters, usually done by clustering tools such as MMseqs2 or by aligning against curated databases like the Comprehensive Antibiotic Resistance Database (CARD) (Alcock *et al.*, 2023),

Table 1. Representations and modalities leveraged for the prediction of AMRs in AI.

Data type	Representation / format	Example tools or databases	Reference
Nucleotide / amino-acid k-mers	Sequence subsequences of fixed length k, counted without alignment to a reference	KMC3, MerCat	Li et al., 2024
Gene content	Presence/absence of clustered resistance-gene families	MMseqs2, CARD, ResFinder, ARDB, BLDB	Li et al., 2024; Alcock et al., 2023
SNP / mutation matrices	Presence, absence, or type of variant at each polymorphic site relative to a reference genome	Snippy; one-hot, label, or chaos-game encoding	Li et al., 2024
MALDI-TOF mass spectra	Spectrum binned into fixed mass-to-charge intervals (e.g., 3 Da bins, 2000–20,000 Da)	DRIAMS database	Weis et al., 2022; Aggarwal et al., 2026
Vibrational spectroscopy	Raman or Fourier-transform infrared spectral signatures	Site-specific spectral libraries	Hassan et al., 2026
Microscopy / colony imaging	Colony morphology, Gram-stain, or single-cell images	AGAR dataset; site-specific image sets	Hassan et al., 2026; Smith and Kirby, 2018; Zieliński et al., 2017
Metagenomic reads	Short-read sequence data from mixed microbial communities, unassembled or assembled	MetaPhlAn, CARD-based alignment	Baker et al., 2023; Hendriksen et al., 2019
Clinical / epidemiological metadata	Patient demographics, infection site, antibiotic exposure history, hospital data	Electronic health records	Arnold et al., 2025; Khangarot et al., 2026

Table 2. Key genomic, spectral, and surveillance resources supporting AI-based AMR prediction and evolutionary forecasting.

Resource	Type	Content	Relevance to evolutionary forecasting
CARD	Curated resistance-gene database	Annotated antibiotic resistance genes with a standardized resistance ontology	Reference baseline for known determinants; divergence from CARD entries can flag candidate novel resistance genes
ResFinder	Alignment-based annotation tool and database	Identification of acquired resistance genes via BLAST-based sequence similarity	Fast, high-confidence annotation of known genes; not designed to detect novel or divergent determinants
ARDB / BLDB	Legacy and class-specific resistance databases	Antibiotic Resistance Genes Database; beta-lactamase-specific sequence collection	Historical reference sets, largely superseded by CARD but still used for cross-validation
DRIAMS	MALDI-TOF spectral repository	Over 300,000 mass spectra linked to roughly 768,000 resistance labels across 803 species	Large enough to support repeated, site-linked sampling that could be restructured as a temporal training set
PATRIC / BV-BRC	Pathogen genomic resource	Whole-genome sequences and pan-genome data across thousands of bacterial isolates	Source of pan-genome feature sets used in phylogeny-informed resistance prediction
NARMS	National surveillance program	Non-typhoidal Salmonella and related enteric pathogen susceptibility data linked to WGS	Longitudinal, multi-year sampling structure suited to temporal MIC-forecasting models
WHO GLASS	Global surveillance system	Antimicrobial consumption and resistance data aggregated across participating countries	Population-level trend data supporting One Health and cross-border evolutionary monitoring

ResFinder, and Antibiotic Resistance Genes Database (ARDB) (Li *et al.*, 2024), or the beta-lactamase database (BLDB) (Li *et al.*, 2024). Mutation or single-nucleotide-polymorphism (SNP)-based representations align mapping

reads to a reference genome and represent whether the sequence exists at the polymorphic positions (or its type in the case of a SNP) with a one-hot encoding, label encoding, or chaos game representation (Li *et al.*, 2024).

Effective use of the non-genomic modalities is also vital in practice. Unlike other mass spectra, binned (or fixed) mass-to-charge intervals have been one of the most popular input formats for resistance-phenotype classifiers as they can be obtained when the pathogen is routinely identified and do not require an extra wet-laboratory step (Weis *et al.*, 2022; Aggarwal *et al.*, 2026). Vibrational spectroscopy (Raman and Fourier-transform infrared) and microscopy or colony-morphology imaging can deliver complementary phenotypic signals which are independent from sequencing, with wide applicability (Hassan *et al.*, 2026). Resistance determinants are acquired and transition from their ancestors via mutation, and this is not accounted for by static gene-presence or single-timepoint mutation matrices for evolutionary forecasting specifically. Genetic variants have been ranked prior to model training to locate known and unknown resistance associated variants in *Mycobacterium tuberculosis* (Yurtseven *et al.*, 2023) and genetic-interaction analysis along with the structural representation of each mutation has been used to trace the sequential evolution of multi-drug resistance to 13 antibiotics in *Mycobacterium tuberculosis* (Kavvas *et al.*, 2018).

The reliability of any prediction, static or evolutionary, is bounded by the reference resource used to generate labels and features. Table 2 summarizes the principal genomic, spectral, and surveillance resources that recur across the reviewed literature, spanning curated resistance-gene databases, large spectral repositories, and population-level surveillance programs, several of which are the natural substrate for the longitudinal, phylogeny-linked datasets that evolutionary forecasting requires.

3.2 Model Architectures and Their Suitability for Temporal Prediction

Classical machine learning methods -support vector machines, logistic regression, random forests, and gradient-boosted tree ensembles such as XGBoost and LightGBM - remain the most widely benchmarked approaches for AMR phenotype prediction, and frequently outperform more complex architectures on structured, moderate-sized datasets such as MALDI-TOF spectra or curated gene-presence matrices, in part because tree-based ensembles resist overfitting when sample sizes are limited (Aggarwal *et al.*, 2026). Random forests are reported as both the most frequently used and the most often best-performing model family across a systematic review of 57 *Klebsiella pneumoniae* AMR-prediction studies, though their

prominence partly reflects their common use as a benchmark comparator (Aggarwal *et al.*, 2026).

Deep learning architecture extends this toolkit in ways directly relevant to evolutionary modeling (Table 3). Convolutional neural networks (CNNs) extract local patterns from structured or grid-like data -mass spectra, Raman spectra, or micrographs -and have been used for tasks from carbapenemase-producing *K. pneumoniae* detection to automated Gram-stain interpretation (Smith and Kirby, 2018; Zieliński *et al.*, 2017) and single-cell bacterial species identification from three-dimensional quantitative phase imaging (Kim *et al.*, 2022). Recurrent architectures, particularly long short-term memory (LSTM) and bidirectional LSTM (BiLSTM) networks, are explicitly designed for sequential and temporal data and have been applied to predict sepsis risk from time-ordered electronic health record measurements (Zhang *et al.*, 2021; Goh *et al.*, 2021; Shashikumar *et al.*, 2021); the same architectural logic -learning how a system's state at one time step constrains its state at the next -is directly transferable to modeling how a lineage's resistance genotype changes across sequential surveillance samples. Graph neural networks (GNNs) represent genomic or metagenomic data as nodes and edges denoting sequence similarity, co-abundance, or genomic context, and because horizontal gene transfer and plasmid exchange are inherently relational events, GNNs are a natural architecture for modeling the propagation of a resistance determinant across a bacterial population or between species (Khangarot *et al.*, 2026). Transformer-based genomic encoders such as DNABERT model long-range contextual dependencies within a sequence more effectively than k-mer-frequency-based random forests (Ji *et al.*, 2021; Khangarot *et al.*, 2026), and generative architectures -variational autoencoders (VAEs) and generative adversarial networks (GANs) -learn a latent representation of resistance-associated genomic variation that can augment scarce training data for rare or emerging resistance phenotypes (Jin *et al.*, 2018; Khangarot *et al.*, 2026).

3.3 From Static Classification to Evolutionary Forecasting

In the *K. pneumoniae* systematic review, internal validation using k-fold cross-validation or hold-out splitting was reported in 96.5% of studies, while external validation on independent datasets appeared in fewer than 30%, and only a single study evaluated a model prospectively on newly

Table 3. Machine learning and deep learning architectures applied to AMR prediction, contrasted by relevance to static classification versus evolutionary forecasting.

Architecture	Typical input	Current predominant use	Relevance to evolutionary forecasting	Reference
Tree ensembles (RF, XGBoost, SVM)	Gene-presence matrices, MALDI-TOF spectra	Static phenotype classification; most frequently used and best-performing family in current literature	Limited — single-timepoint by design; can be repeated cross-sectionally but does not model trajectories	Aggarwal et al., 2026
Convolutional neural networks (CNN)	Mass spectra, Raman spectra, microscopy images	Spectral/image-based resistance and species classification	Low-moderate — captures local structure only; requires pairing with a temporal layer for forecasting	Smith and Kirby, 2018; Zieliński et al., 2017
Recurrent networks (RNN / LSTM / BiLSTM)	Time-ordered clinical or genomic measurements	Sepsis risk prediction from sequential EHR data	High — designed to model sequential dependencies across time steps	Zhang et al., 2021; Goh et al., 2021; Shashikumar et al., 2021
Graph neural networks (GNN)	Sequence-similarity, co-abundance, or genomic-context graphs	Strain-level classification in metagenomic datasets	High — naturally represents relational events such as horizontal gene transfer	Khangarot et al., 2026
Transformer genomic encoders (e.g., DNABERT)	Raw or tokenized nucleotide sequence	Taxonomic classification, sequence representation learning	Moderate-high — captures long-range context; temporal extensions still emerging	Ji et al., 2021; Khangarot et al., 2026
Generative models (GAN / VAE)	Latent representation of resistance-associated variation	Data augmentation for rare/emerging phenotypes; drug and peptide design	Moderate — can simulate plausible future genotypes but largely unvalidated for forecasting	Jin et al., 2018; Stokes et al., 2020; Swanson et al., 2024

collected samples (Aggarwal *et al.*, 2026). Temporal validation -testing a model on data collected after the training period, the minimum requirement for assessing whether a model captures genuinely predictive, evolution-relevant signal -was reported in only about 16% of the reviewed studies, most often through simple chronological data splits rather than explicit forecasting of future resistance states (Aggarwal *et al.*, 2026). Table 6 collects these figures alongside the review's risk-of-bias and reference-standard findings, since together they quantify precisely how far the current evidence base sits from a validated forecasting capability, rather than only illustrating the point qualitatively. Figure 2 contrasts this conventional single-timepoint paradigm with a longitudinal, evolution-aware alternative in which sequential surveillance samples are linked through a temporally or relationally structured model. Anticipatory framing of this kind is not entirely new: more than a decade before the current wave of genomic AI models, an artificial neural network trained on clinical risk factors was used to predict the likelihood of future methicillin-resistant *Staphylococcus aureus* carriage rather than to classify an isolate already in hand (Hsu *et al.*, 2008),

an early demonstration that AI could be pointed at a forward-looking question when the underlying data were structured to support one.

Large-scale genomic and metagenomic surveillance programs provide raw material for this kind of modeling. Analyses of *Salmonella* whole-genome sequences collected through national resistance monitoring have been used to train models predicting minimum inhibitory concentrations (MICs) within a two-fold dilution range at roughly 95% accuracy (Nguyen *et al.*, 2019), and pan-genome feature-selection strategies applied to gradient-boosted models have outperformed both whole-genome and previously known AMR-gene feature sets for predicting resistance in other species (Yang and Wu, 2022). A random forest trained on whole-genome sequencing data from five geographically diverse countries to predict ciprofloxacin MIC in *Escherichia coli* illustrates both the promise and the limits of this approach: performance was strong overall but markedly worse for MIC values underrepresented in the training data, underscoring that evolutionary forecasting models inherit the sampling biases of their surveillance inputs (Pataki *et al.*, 2020). Comparable work applying whole-genome

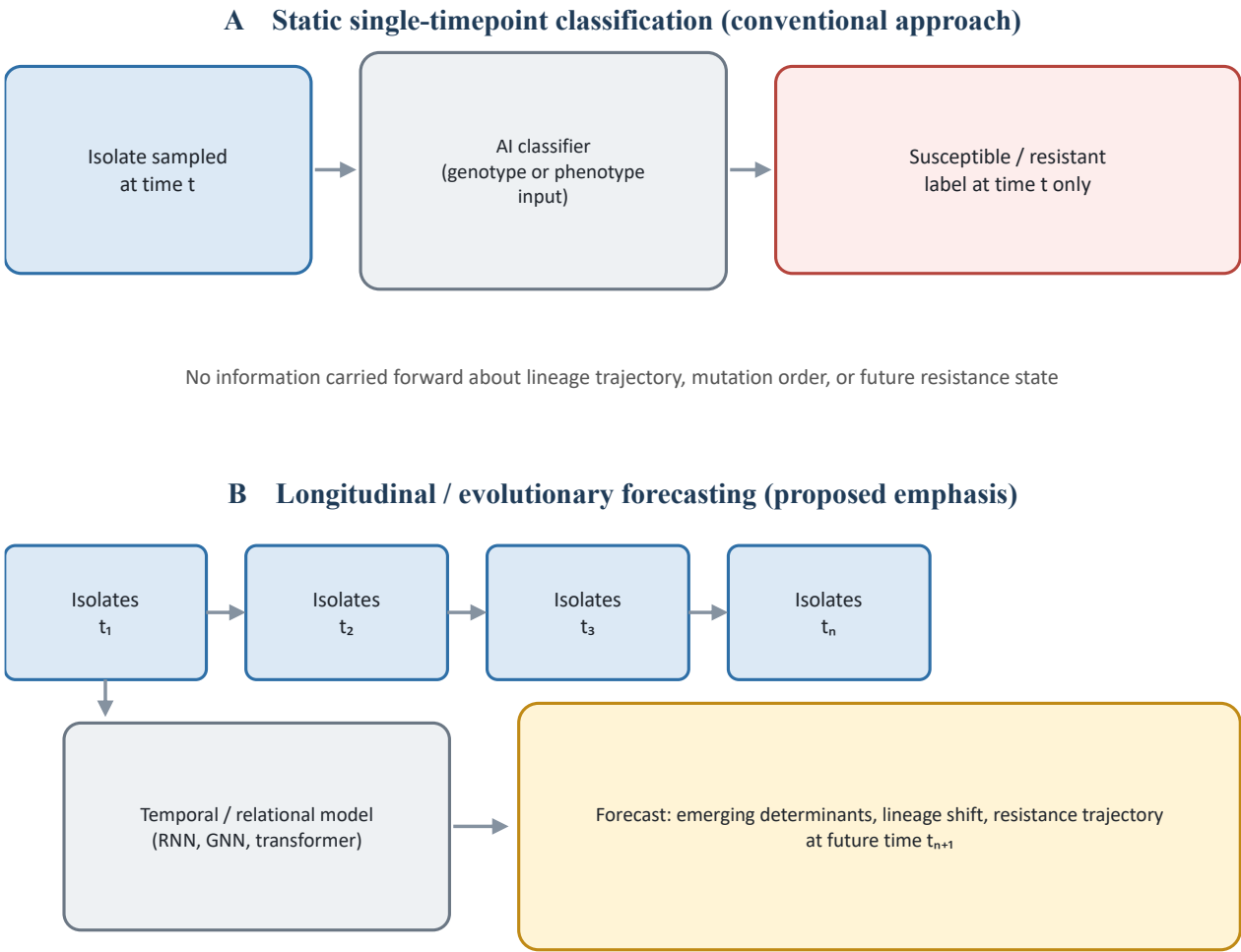


Figure 2. Static, single-timepoint classification (A) versus longitudinal, evolution-aware forecasting (B). Panel B links sequential surveillance samples through a temporally or relationally structured model to anticipate future resistance states rather than only label a present one.

evolutionary forecasting requires is itself constrained by patient-privacy regulation and by the practical difficulty of pooling raw sequencing data across institutional and national boundaries. Federated learning addresses this by training local models at each participating site and sharing only model parameters or gradients with a coordinating server (Zwiers *et al.*, 2024; Khangarot *et al.*, 2026). Federated convolutional models trained across multiple hospitals have shown improved cross-site generalizability for resistance-gene classification relative to single-institution models (Khangarot *et al.*, 2026), though privacy-preserving mechanisms such as homomorphic encryption and differential privacy introduce a trade-off between privacy strength and predictive performance, and

vibrational spectroscopy (Aggarwal *et al.*, 2026). Reported discrimination is frequently strong -roughly three-quarters of AUROC-primary studies reported best-model performance of 0.90 or higher -yet dataset origin was heavily concentrated in China and Taiwan, with comparatively little representation from Europe, North America, or high-AMR regions such as sub-Saharan Africa and Latin America (Aggarwal *et al.*, 2026). A synthesis of comparable work in *Staphylococcus aureus* applying gradient-boosted multi-label learning to large-scale MALDI-TOF data further illustrates how routinely generated spectra can be repurposed for rapid, serial resistance prediction across several antibiotics at once (Zhang *et al.*, 2022).

Table 4. Selected studies illustrating genomic, phylogenetic, cross-host, or temporal approaches relevant to AMR evolution prediction.

Pathogen	Data type	Model	Key finding	Reference
Mycobacterium tuberculosis	WGS, pan-genome allele presence/absence	Linear SVM with phylogeny-informed feature selection	Identified known and previously unrecognized resistance-associated variants across multiple drug classes	Kavvas et al., 2018; Yurtseven et al., 2023
Mycobacterium tuberculosis	SNVs across 16-country isolate collection	Deep denoising autoencoder (DeepAMR)	Predicted co-occurrent multidrug resistance with sensitivity up to 96.3%	Yang et al., 2019
Escherichia coli	WGS-derived SNPs and gene content, 5-country cohort	Random forest, leave-one-country-out validation	Predicted ciprofloxacin MIC; error concentrated in underrepresented MIC classes	Pataki et al., 2020
Escherichia coli	Pan-genome WGS data	Random forest / gradient boosting	AUROC 0.95 for ciprofloxacin resistance on external validation set	Ren et al., 2022; Moradigaravand et al., 2018
Nontyphoidal Salmonella	WGS and NARMS susceptibility data	Machine learning MIC predictor	Predicted MIC within ± 1 two-fold dilution at ~95% overall accuracy	Nguyen et al., 2019
Nontyphoidal Salmonella (agricultural)	WGS	Machine learning classifiers	Identified primary genetic drivers of resistance consistent with known mechanisms	Maguire et al., 2019
Escherichia coli (poultry farms, China)	Metagenomic reads, multi-source samples	XGBoost, SVM, random forest	Identified 10 mobile ARGs as cross-sample predictors; revealed co-localization of resistance genes	Baker et al., 2023
Acinetobacter baumannii	WGS, human and non-human isolates	Alignment to resistance-gene databases	Demonstrated cross-host circulation of beta-lactamase and non-beta-lactamase determinants	Wareth et al., 2021
Escherichia coli (livestock/companion animals)	WGS	Interpretable ML classifier	Genotype-phenotype relationships varied by antibiotic and host species	Chung et al., 2023
Klebsiella pneumoniae (57-study synthesis)	MALDI-TOF, WGS, vibrational spectroscopy	Random forest, gradient boosting, CNN	High internal accuracy but marked geographic/lineage concentration limiting generalizability	Aggarwal et al., 2026

This geographic concentration matters specifically for evolutionary forecasting because *K. pneumoniae*'s high-risk clones differ by region -sequence type ST11 predominates across much of Asia, while ST258 is more prominent in Europe and North America -and MALDI-TOF spectra can carry clade-specific peak patterns tied to these lineages (Aggarwal *et al.*, 2026). A model trained predominantly on an ST11-dominant dataset may therefore learn resistance-associated signal that is partly a proxy for lineage identity rather than a generalizable resistance mechanism and would be expected to generalize poorly if that lineage's dominance shifts or if the model is deployed where a different clone predominates. Because lineage replacement is itself one of the principal ways resistances in a pathogen population evolves over time, this is a direct illustration of why evolution-aware model validation must explicitly test

performance across shifting lineage composition rather than only across held-out isolates drawn from the same underlying population.

Emerging resistance frequently moves between hosts and environments before it is recognized clinically. Comparative genomic analyses of *Acinetobacter baumannii* isolates from human and non-human sources have used whole-genome sequencing and curated resistance-gene databases to demonstrate that beta-lactamase and non-beta-lactamase resistance determinants circulate across host boundaries (Wareth *et al.*, 2021), and machine learning models trained to predict *Escherichia coli* resistance phenotypes in livestock and companion animals have shown that the relationship between genotype and phenotype varies not only with the antibiotic in question but with the host species from which the isolate was recovered (Chung *et al.*, 2023).

These findings imply that a model intended to anticipate resistance evolution in an emerging pathogen cannot safely assume that genotype-phenotype relationships learned in one host context will transfer cleanly to agricultural or environmental reservoirs where the same organism may be evolving under different selective pressures, even though those reservoirs frequently precede the clinical emergence of a newly resistant lineage.

A core requirement for anticipating resistance evolution in an emerging pathogen is the ability to recognize mechanisms that have not yet been catalogued in reference databases, since alignment-based annotation tools can only detect sufficiently similar to already-known resistance genes. Deep learning models trained on sequence features rather than explicit homology are specifically positioned to close this gap. DeepARG uses convolutional filtering over sequence-derived features to classify resistance genes across roughly thirty categories and reports precision above 0.97 on benchmark data, though performance varies considerably by class, with sparsely represented categories such as glycopeptides and polymyxins classified less reliably (Arango-Argoty *et al.*, 2018). HMD-ARG improves on this with a hierarchical multi-label architecture that better separates closely related resistance-gene subfamilies and improves recall for underrepresented classes (Li *et al.*, 2021). Because these models learn continuous sequence representations rather than relying on exact database matches, they can, in principle, flag divergent or entirely novel resistance-gene candidates for experimental follow-up—an essential capability for surveillance of pathogens whose resistance repertoire is still being characterized, though this flexibility comes at the cost of increased false-positive risk and reduced interpretability relative to alignment-based annotation.

3.5 Technical and Translational Challenges

A substantial share of AMR prediction models are trained against automated commercial susceptibility systems such as VITEK 2 and BD Phoenix rather than the reference-grade broth microdilution method; in the *K. pneumoniae* systematic review, just over half of studies used automated commercial systems as the ground truth, compared with roughly 16% using broth microdilution (Aggarwal *et al.*, 2026). Automated systems carry recognized limitations, including historically reduced sensitivity for detecting KPC-producing organisms and unreliable performance for colistin susceptibility testing. When a predictive model is trained to

reproduce the output of an imperfect reference standard, it risks a form of diagnostic mimicry in which the model reproduces the reference method's blind spots rather than the underlying resistance biology, a problem that compounds over successive surveillance cycles (Aggarwal *et al.*, 2026). Direct comparison of rapid automated identification and susceptibility platforms against panel-based reference methods illustrates the magnitude of this concordance problem in practice (Sze *et al.*, 2021).

Deep learning architectures are frequently described as black boxes because the basis for an individual prediction is not readily traceable to specific input features, which is problematic where an erroneous susceptibility prediction could contribute directly to treatment failure. Explainable AI (XAI) techniques are increasingly applied to address this: SHAP (Shapley Additive Explanations) values quantify the contribution of individual genes, mutations, or spectral features to a given prediction (Lundberg and Lee, 2017), and gradient-based attribution and attention-weight visualization in transformer or graph-neural-network architectures indicate which sequence positions or genomic neighborhoods most influenced a classification (Khangarot *et al.*, 2026). These tools have already been used to show that some high-performing MALDI-TOF resistance classifiers are partly discriminating based on phylogenetic lineage rather than resistance mechanism per se, directly illustrating the generalizability risk described in Section 3.4 (Khangarot *et al.*, 2026). Current XAI methods nonetheless remain largely generic machine-learning tools adapted to microbiology rather than frameworks purpose-built for microbiological reasoning, and do not yet routinely account for domain-specific considerations such as plasmid context, strain-level phylogenetic uncertainty, or contamination from co-occurring organisms in metagenomic samples (Khangarot *et al.*, 2026).

AMR datasets vary substantially in sequencing platform, sample preparation, antimicrobial susceptibility testing protocol, and metadata schema across institutions, introducing batch effects that degrade model performance when a model trained at one site is deployed at another (Khangarot *et al.*, 2026). Genomic training data are also markedly skewed toward high-income regions and East Asia in particular, with published estimates suggesting that fewer than one in ten AI-microbiology studies originate from low- and middle-income countries, despite these regions frequently bearing the greatest clinical burden of

Table 5. Key translational challenges for evolution-aware AI-AMR surveillance and proposed mitigations.

Challenge	Description	Proposed mitigation
Reference-standard imperfection	Many models are trained against automated commercial AST systems rather than reference-grade broth microdilution, risking reproduction of the comparator's diagnostic blind spots	Prioritize broth-microdilution-referenced datasets; explicitly report reference-standard provenance
Limited external / prospective validation	Internal cross-validation is near-universal; external and prospective (temporal) validation remain rare across the reviewed literature	Mandate temporal hold-out and lineage-shift testing as a condition of reporting forecasting performance
Interpretability	Deep learning architectures act as black boxes; predictions may reflect lineage or batch confounders rather than resistance mechanism	Apply microbiology-adapted explainable AI (SHAP, attention visualization) to audit whether models track resistance mechanism versus phylogenetic proxy
Geographic / lineage bias	Training data concentrated in high-income regions and specific dominant clones; poor transfer to underrepresented lineages or regions	Prioritize geographically diverse, multi-clone datasets; test explicitly across lineage-replacement scenarios
Cross-institutional data privacy	Longitudinal, multi-site data pooling is constrained by patient-privacy regulation and cross-border data-sharing restrictions	Federated learning with calibrated homomorphic encryption or differential privacy
Data scarcity for emerging pathogens	Newly prioritized or rare pathogens have limited historical surveillance data, constraining model training	Generative data augmentation (GAN/VAE) combined with prospective wet-laboratory validation of flagged candidates

antimicrobial resistance (Khangarot *et al.*, 2026). Because resistomes and circulating clones differ geographically, models trained predominantly on data from one region generalize poorly to genomically distinct populations elsewhere -a direct obstacle to building an evolutionary forecasting system intended to provide early warning on a global rather than a purely local scale.

Across the reviewed literature, internal validation is near-universal, while external and especially prospective validation on newly collected samples remains rare. In the *K. pneumoniae* systematic review, no included study met criteria for an overall low risk of bias under either the PROBAST or QUADAS-2 assessment frameworks, and high risk of bias was identified in the majority of studies, driven principally by weaknesses in the analysis domain and in participant or patient selection (Aggarwal *et al.*, 2026). This validation gap is especially consequential for evolutionary forecasting, since a model's ability to predict resistance in isolates drawn from the same time period and population as its training data provides essentially no evidence about its ability to anticipate resistance states that have not yet been observed.

3.6 AI-Assisted Countermeasure Design as a Complementary Arm of Evolutionary Surveillance

Anticipating how resistance will evolve is only half of a genuinely evolution-aware response; the complementary task is accelerating discovery of the countermeasures

needed to stay ahead of that evolution. Deep learning models trained on bacterial growth-inhibition data have been used to screen chemical libraries far larger than any feasible wet-laboratory campaign, contributing to the identification of structurally novel antibacterial compounds such as halicin (Stokes *et al.*, 2020), and generative approaches that assemble molecules from synthetically tractable building blocks have extended this to designing antibiotic candidates with activity against *Acinetobacter baumannii* (Swanson *et al.*, 2024). Antimicrobial peptides represent a parallel avenue: machine learning models trained to estimate the contribution of each amino acid position to antimicrobial activity have been used to design peptides such as DP7, with demonstrated in vivo efficacy against both drug-sensitive and drug-resistant *Staphylococcus aureus* (Wu *et al.*, 2014). Genome-mining approaches such as DeepBGC apply recurrent architectures to identify biosynthetic gene clusters likely to encode novel secondary metabolites, offering a route to previously uncharacterized natural-product antibiotics (Hannigan *et al.*, 2019).

Advances in protein structure prediction have further accelerated structure-based design of these countermeasures. AlphaFold2 predicts the three-dimensional structure of proteins from amino acid sequence with near-experimental accuracy (Jumper *et al.*, 2021), its successor AlphaFold3 extends this capability to complexes of proteins, nucleic acids, and small molecules (Abramson

Table 6. Quantitative validation landscape in AI-based AMR prediction, evidence from a 57-study systematic review of *Klebsiella pneumoniae* (Aggarwal et al., 2026).

Metric	Reported value	Interpretation
Studies included	57 (2018–2025)	Nearly half (47.4%) published in 2024–2025 alone, reflecting rapid recent growth
Best-model AUROC $\geq$ 0.90	26/35 (74.3%) of AUROC-primary studies	Strong cross-sectional discrimination is now common
Best-model AUROC $\geq$ 0.95	21/35 (60.0%)	A majority clear an even higher discrimination threshold
Internal validation only	55/57 (96.5%)	Near-universal reliance on k-fold or hold-out splitting from the same population
External validation (independent dataset)	17/57 (29.8%)	Fewer than one in three studies test generalizability beyond their training source
Temporal validation (chronological split)	9/57 (15.8%)	The minimum evidence needed for a forecasting claim is rarely reported
Prospective, real-world validation	1/57 (1.8%)	Only a single study evaluated a model on newly collected samples in real time
Overall high risk of bias (PROBAST/QUADAS-2)	45/57 (78.9%)	No included study met criteria for an overall low risk of bias
Reference standard: automated commercial AST	29/57 (50.9%)	Models risk reproducing the reference system's own diagnostic blind spots
Reference standard: gold-standard broth microdilution	9/57 (15.8%)	Reference-grade phenotypic ground truth remains the minority case
Dataset origin: China and Taiwan combined	26/57 (45.6%)	Marked geographic concentration limits transferability to other resistomes and lineages

Table 7. Roadmap components, rationale, and expected impact for an evolution-aware AI-AMR surveillance framework.

Component	Rationale	Expected impact
1. Longitudinal genomic & phenotypic surveillance	Static, single-timepoint snapshots cannot reveal lineage replacement or the order of mutation accumulation	Produces training data with genuine temporal structure rather than incidental time-stamping
2. Phylogeny- and time-aware feature representation	Unordered gene-presence or mutation matrices discard the evolutionary context in which resistance arose	Improves detection of stepwise, combinatorial resistance accumulation rather than isolated markers
3. Temporal / relational model architectures	RNN, GNN, and transformer architectures are designed for sequential or relational dependencies that tree ensembles cannot represent	Captures lineage dynamics and transmission events such as horizontal gene transfer directly
4. Forecast-oriented validation	Internal cross-validation on same-population data cannot demonstrate forecasting ability, only recall of familiar patterns	Establishes genuine evidence of predictive, rather than purely descriptive, performance
5. Federated, privacy-preserving learning	Cross-institutional and cross-border data pooling is constrained by patient-privacy regulation	Expands geographic and lineage diversity in training data without centralizing sensitive records
6. Explainable AI & prospective validation	Black-box predictions risk reflecting lineage or batch confounders rather than true resistance mechanisms	Builds clinical and regulatory trust and confirms flagged candidate mechanisms experimentally

et al., 2024), and open replications such as BOLTZ-1 have made comparable capability more widely accessible for docking-based antibiotic design (Wohlwend *et al.*, 2024). Integrating this countermeasure-design capability with the evolutionary forecasting framework proposed in Section 4 - so that a flagged emerging resistance mechanism can trigger a targeted discovery campaign rather than only a surveillance alert -represents a natural, though largely unrealized, extension of both literatures.

4. Discussion: Toward an Evolution-Aware AI Surveillance Framework

Bringing together the technical capabilities reviewed above suggests a set of design principles for AI systems specifically intended to anticipate resistance evolution in emerging bacterial pathogens, rather than only to classify current resistance status (Figure 3). First, training data should be structured longitudinally wherever possible, with genomic and phenotypic samples linked by collection date

Proposed roadmap for an evolution-aware AI-AMR surveillance framework

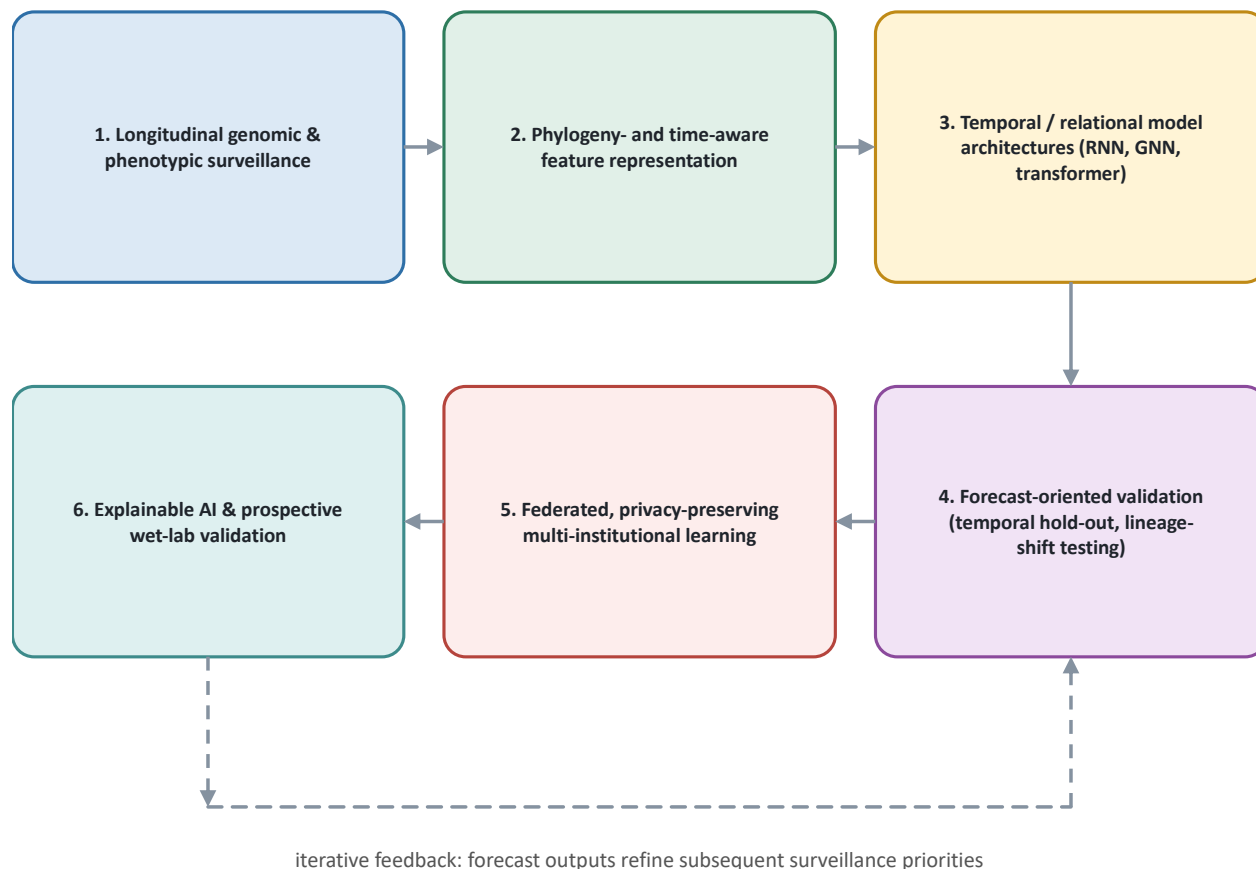


Figure 3. Six-component roadmap moving AI-based AMR research from cross-sectional classification toward validated, evolution-aware surveillance of emerging bacterial pathogens.

and, ideally, by phylogenetic relationship, so that temporal and lineage-replacement dynamics are learnable rather than incidental. Second, model architecture should be matched to this temporal structure: recurrent and graph-based models that explicitly represent sequence-of-events or network-propagation dynamics are better suited to this task than architecture designed purely for single-timepoint classification. Third, validation protocols should be redesigned around forecasting rather than classification, explicitly testing whether a model trained on data up to a given date can anticipate resistance patterns observed after that date, and whether performance is maintained when the dominant circulating lineage shifts.

Fourth, federated learning architectures, supported by appropriately calibrated privacy-preserving mechanisms,

offer a practical route to assembling the geographically and temporally diverse datasets this kind of modeling requires without centralizing sensitive genomic or clinical data. Fifth, explainable AI methods purpose-built for microbiological reasoning -capable of distinguishing a genuine resistance mechanism from a lineage-associated confounder, and of flagging phylogenetic or contamination-related uncertainty -should be treated as a prerequisite for clinical and public health deployment rather than an optional add-on. Finally, because emerging pathogens by definition have limited historical surveillance data, generative and transfer-learning approaches that can extrapolate cautiously from related, better-characterized organisms, combined with prospective wet-laboratory validation of model-flagged candidate resistance

mechanisms, will likely be necessary to compensate for genuinely novel or data-sparse threats. Table 7 summarizes the rational and expected impact of each of these six components alongside the visual roadmap in Figure 3.

None of these components is entirely absent from the current literature -phylogeny-aware feature selection, recurrent and graph-based temporal models, federated AMR prediction, and microbiology-adapted explainable AI have each been demonstrated individually -but they have rarely been integrated into a single validated pipeline explicitly evaluated on its ability to forecast future resistance states. Closing that integration gap, rather than further incremental improvement of single-timepoint classification accuracy, is arguably the most consequential open problem for AI-assisted AMR research over the next several years.

Conclusion

Artificial intelligence has already demonstrated clear value in accelerating and improving the accuracy of antimicrobial resistance detection, from MALDI-TOF-based phenotype classification to genomic identification of known and novel resistance genes. The great majority of this progress, however, addresses a static classification problem, while the more consequential public health need -anticipating how resistance in emerging bacterial pathogens will evolve, spread, and combine before it becomes clinically entrenched -remains comparatively underdeveloped. Realizing this evolutionary forecasting capability will require deliberately temporal and phylogenetically structured training data, model architectures suited to sequential and relational dynamics, validation protocols that explicitly test forecasting rather than classification performance, privacy-preserving federated infrastructure capable of pooling globally diverse surveillance data, and explainable AI methods grounded in microbiological rather than generic machine-learning reasoning. Achieving this will depend on sustained collaboration between microbiologists, computational scientists, clinicians, and public health authorities, and on a shift in how the field measures success -from whether a model classifies an isolate accurately today, to whether it can be trusted to warn us about resistance before it arrives.

Ethical Approval

Not applicable

Conflict of Interest

The authors declare no conflict of interest.

Funding

Not applicable

Author Contributions

Zehra Fatima contributed to the conceptualization, literature search, data collection, analysis, and drafting of the manuscript. Shehr Hayat contributed to the literature review, data interpretation, and manuscript writing. Maham Rafiq contributed to data collection, literature analysis, and critical revision of the manuscript. Aiman Nishat contributed to literature screening, data synthesis, and manuscript editing. Sadia Amjad supervised the overall work, contributed to the conceptualization and critical evaluation of the manuscript, and provided final approval of the version to be published. All authors reviewed and approved the final manuscript.

Acknowledgment

The authors have no acknowledgments to declare.

References

- Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., & Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630(8016), 493-500.
- Aggarwal, R. V., Shah, N., Wong, J. J. E., & Dajani, Z. (2026). Artificial Intelligence for Antimicrobial Resistance Detection and Prediction in *Klebsiella pneumoniae*: A Systematic Review of Clinical Microbiology Applications. *Infection and Drug Resistance*, 614240.
- Alcock, B. P., Huynh, W., Chalil, R., Smith, K. W., Raphenya, A. R., Wlodarski, M. A., & McArthur, A. G. (2023). CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic acids research*, 51(D1), D690-D699.
- Arango-Argoty, G., Garner, E., Pruden, A., Heath, L. S., Vikesland, P., & Zhang, L. (2018). DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Microbiome*, 6(1), 23.
- Arnold, A., McLellan, S., & Stokes, J. M. (2025). How AI can help us beat AMR. *npj Antimicrobials and Resistance*, 3(1), 18.
- Baker, M., Zhang, X., Maciel-Guerra, A., Dong, Y., Wang, W., Hu, Y., & Dottorini, T. (2023). Machine learning and metagenomics reveal shared antimicrobial resistance profiles across multiple chicken farms and abattoirs in China. *Nature Food*, 4(8), 707-720.
- Bombaywala, S., Purohit, H. J., & Dafale, N. A. (2021). Mobility of antibiotic resistance and its co-occurrence with metal resistance in pathogens under

- oxidative stress. *Journal of Environmental Management*, 297, 113315.
- Chung, H. C., Foxx, C. L., Hicks, J. A., Stuber, T. P., Friedberg, I., Dorman, K. S., & Harris, B. (2023). An accurate and interpretable model for antimicrobial resistance in pathogenic *Escherichia coli* from livestock and companion animal species. *PLoS One*, 18(8), e0290473.
- de Kraker, M. E., Stewardson, A. J., & Harbarth, S. (2016). Will 10 million people die a year due to antimicrobial resistance by 2050?. *PLoS medicine*, 13(11), e1002184.
- Gerada, A., Harper, N., Howard, A., Reza, N., & Hope, W. (2024). Determination of minimum inhibitory concentrations using machine-learning-assisted agar dilution. *Microbiology spectrum*, 12(5), e04209-23.
- Goh, K. H., Wang, L., Yeow, A. Y. K., Poh, H., Li, K., Yeow, J. J. L., & Tan, G. Y. H. (2021). Artificial intelligence in sepsis early prediction and diagnosis using unstructured data in healthcare. *Nature Communications*, 12(1), 711.
- Hannigan, G. D., Prihoda, D., Palicka, A., Soukup, J., Klempir, O., Rampula, L., & Bitton, D. A. (2019). A deep learning genome-mining strategy for biosynthetic gene cluster prediction. *Nucleic Acids research*, 47(18), e110-e110.
- Hassan, M., Abu El-Ruz, R., & Zughaier, S. M. (2026). Role of artificial intelligence in bacterial diagnostics and surveillance of anti-microbial resistance. *Digital Health*, 12, 20552076261443074.
- Hendriksen, R. S., Munk, P., Njage, P., Van Bunnik, B., McNally, L., Lukjancenko, O., & Aarestrup, F. M. (2019). Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nature Communications*, 10(1), 1124.
- Hsu, C. C., Lin, Y. E., Chen, Y. S., Liu, Y. C., & Muder, R. R. (2008). Validation study of artificial neural network models for prediction of methicillin-resistant *Staphylococcus aureus* carriage. *Infection Control & Hospital Epidemiology*, 29(7), 607-614.
- Jacot, D., Gizha, S., Orny, C., Fernandes, M., Tricoli, C., Marcelpoil, R., & Croxatto, A. (2024). Development and evaluation of an artificial intelligence for bacterial growth monitoring in clinical bacteriology. *Journal of clinical microbiology*, 62(5), e01651-23.
- Ji, Y., Zhou, Z., Liu, H., Davuluri, R. V. (2021) DNABERT: contextual DNA language model for genomics enabling k-mer residue/token contribution mapping in AMR classifiers. *Nature Machine Intelligence*, 3, 123-132.
- Jin, W., Barzilay, R., & Jaakkola, T. (2018, July). Junction tree variational autoencoder for molecular graph generation. In *International conference on machine learning* (pp. 2323-2332). PMLR.
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *nature*, 596(7873), 583-589.
- Kavvas, E. S., Catoi, E., Mih, N., Yurkovich, J. T., Seif, Y., Dillon, N., & Palsson, B. O. (2018). Machine learning and structural analysis of *Mycobacterium tuberculosis* pan-genome identifies genetic signatures of antibiotic resistance. *Nature communications*, 9(1), 4306.
- Khangarot, R., Kumari, V., Mishra, R., & Singh, A. (2026). Artificial intelligence in microbiology: implications for metagenomics, diagnostics, and AMR surveillance. *BioMedical Engineering Online*.
- Kim, G., Ahn, C.S., Jo, Y. (2022) Rapid species identification of pathogenic bacteria from a minute quantity exploiting three-dimensional quantitative phase imaging and artificial neural network. *Light: Science & Applications*, 11, 190.
- Li, Y., Xu, Z., Han, W., Cao, H., Umarov, R., Yan, A., & Gao, X. (2021). HMD-ARG: hierarchical multi-task deep learning for annotating antibiotic resistance genes. *Microbiome*, 9(1), 40.
- Li, Y., Cui, X., Yang, X., Liu, G., & Zhang, J. (2024). Artificial intelligence in predicting pathogenic microorganisms' antimicrobial resistance: challenges, progress, and prospects. *Frontiers in Cellular and Infection Microbiology*, 14, 1482186.
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30.
- Maguire, F., Rehman, M. A., Carrillo, C., Diarra, M. S., & Beiko, R. G. (2019). Identification of primary antimicrobial resistance drivers in agricultural nontyphoidal *Salmonella enterica* serovars by using machine learning. *Msystems*, 4(4), 10-1128.
- Moradigaravand, D., Palm, M., Farewell, A., Mustonen, V., Warringer, J., & Parts, L. (2018). Prediction of antibiotic resistance in *Escherichia coli* from large-scale pan-genome data. *PLoS Computational Biology*, 14(12), e1006258.
- Murray, C. J., Ikuta, K. S., Sharara, F., Swetschinski, L., Aguilar, G. R., Gray, A., & Tasak, N. (2022). Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The lancet*, 399(10325), 629-655.
- Naghavi, M., Vollset, S. E., Ikuta, K. S., Swetschinski, L. R., Gray, A. P., Wool, E. E., & Dekker, D. M. (2024). Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with

- forecasts to 2050. *The Lancet*, 404(10459), 1199-1226.
- Nguyen, M., Long, S. W., McDermott, P. F., Olsen, R. J., Olson, R., Stevens, R. L., & Davis, J. J. (2019). Using machine learning to predict antimicrobial MICs and associated genomic features for nontyphoidal *Salmonella*. *Journal of Clinical Microbiology*, 57(2), 10-1128.
- Pataki, B. Á., Matamoros, S., van der Putten, B. C., Remondini, D., Giampieri, E., Aytan-Aktug, D., & Schultsz, C. (2020). Understanding and predicting ciprofloxacin minimum inhibitory concentration in *Escherichia coli* with machine learning. *Scientific Reports*, 10(1), 15026.
- Ren, Y., Chakraborty, T., Doijad, S., Falgenhauer, L., Falgenhauer, J., Goesmann, A., & Heider, D. (2022). Prediction of antimicrobial resistance based on whole-genome sequencing and machine learning. *Bioinformatics*, 38(2), 325-334.
- Shashikumar, S. P., Wardi, G., Malhotra, A., & Nemati, S. (2021). Artificial intelligence sepsis prediction algorithm learns to say “I don’t know”. *NPJ Digital Medicine*, 4(1), 134.
- Smith, K. P., Kang, A. D., & Kirby, J. E. (2018). Automated interpretation of blood culture gram stains by use of a deep convolutional neural network. *Journal of Clinical Microbiology*, 56(3), 10-1128.
- Stokes, J. M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N. M., & Collins, J. J. (2020). A deep learning approach to antibiotic discovery. *Cell*, 180(4), 688-702.
- Swanson, K., Liu, G., Catacutan, D. B., Arnold, A., Zou, J., & Stokes, J. M. (2024). Generative AI for designing and validating easily synthesizable and structurally novel antibiotics. *Nature Machine Intelligence*, 6(3), 338-353.
- Sze, D. T., Lau, C. C., Chan, T. M., Ma, E. S., & Tang, B. S. (2021). Comparison of novel rapid diagnostic of blood culture identification and antimicrobial susceptibility testing by Accelerate Pheno system and BioFire FilmArray Blood Culture Identification and BioFire FilmArray Blood Culture Identification 2 panels. *BMC Microbiology*, 21(1), 350.
- Wareth, G., Brandt, C., Sprague, L. D., Neubauer, H., & Pletz, M. W. (2021). WGS based analysis of acquired antimicrobial resistance in human and non-human *Acinetobacter baumannii* isolates from a German perspective. *BMC Microbiology*, 21(1), 210.
- Weis, C., Cuénod, A., Rieck, B., Dubuis, O., Graf, S., Lang, C., & Egli, A. (2022). Direct antimicrobial resistance prediction from clinical MALDI-TOF mass spectra using machine learning. *Nature Medicine*, 28(1), 164-174.
- Wohlwend, J., Corso, G., Passaro, S., Getz, N., Reveiz, M., Leidal, K., & Barzilay, R. (2025). Boltz-1 democratizing biomolecular interaction modeling. *BioRxiv*, 2024-11.
- Wu, X., Wang, Z., Li, X., Fan, Y., He, G., Wan, Y., & Yang, L. (2014). In vitro and in vivo activities of antimicrobial peptides developed using an amino acid-based activity prediction method. *Antimicrobial Agents and Chemotherapy*, 58(9), 5342-5349.
- Yang, Y., Walker, T. M., Walker, A. S., Wilson, D. J., Peto, T. E., Crook, D. W., & Clifton, D. A. (2019). DeepAMR for predicting co-occurrent resistance of *Mycobacterium tuberculosis*. *Bioinformatics*, 35(18), 3240-3249.
- Yang, M. R., & Wu, Y. W. (2022). Enhancing predictions of antimicrobial resistance of pathogens by expanding the potential resistance gene repertoire using a pan-genome-based feature selection approach. *Bmc Bioinformatics*, 23(Suppl 4), 131.
- Yurtseven, A., Buyanova, S., Agrawal, A. A., Bochkareva, O. O., & Kalinina, O. V. (2023). Machine learning and phylogenetic analysis allow for predicting antibiotic resistance in *M. tuberculosis*. *BMC Microbiology*, 23(1), 404.
- Zhang, D., Yin, C., Hunold, K. M., Jiang, X., Caterino, J. M., & Zhang, P. (2021). An interpretable deep-learning model for early prediction of sepsis in the emergency department. *Patterns*, 2(2).
- Zhang, J., Wang, Z., Wang, H. Y., Chung, C. R., Horng, J. T., Lu, J. J., & Lee, T. Y. (2022). Rapid antibiotic resistance serial prediction in *Staphylococcus aureus* based on large-scale MALDI-TOF data by applying XGBoost in multi-label learning. *Frontiers in Microbiology*, 13, 853775.
- Zieliński, B., Plichta, A., Misztal, K., Spurek, P., Brzywczy-Włoch, M., & Ochońska, D. (2017). Deep learning approach to bacterial colony classification. *PloS one*, 12(9), e0184554.
- Zwiers, L. C., Grobbee, D. E., Uijl, A., & Ong, D. S. (2024). Federated learning as a smart tool for research on infectious diseases. *BMC Infectious Diseases*, 24(1), 1327.

Disclaimer/Publisher's Note: The views, opinions, and data presented in all publications are exclusively those of the respective author(s) and contributor(s) and do not reflect the stance of Pagepal or its editorial team. Pagepal and the editor(s) disclaim any responsibility for harm to individuals or damage to property resulting from the use of any ideas, techniques, instructions, or products mentioned in the content.