Review Articles

Research on the Use of Artificial Intelligence in Controlling Drug-Resistant Bacteria

Brady Wilson, Mia Carter *

Department of Public Health, College of Public Health, University of South Florida, Tampa, FL, United States

ARTICLE INFO

Article history:
Received 19 Sep. 2026
Accepted 23 Sep. 2026

Keywords:
Antimicrobial resistance
Artificial intelligence
Machine learning
Antibiotic discovery
Resistance prediction
Clinical decision support

ABSTRACT

Antimicrobial resistance (AMR) has become one of the most urgent threats to global public health, with an estimated 4.95 million deaths associated with bacterial resistance in 2019 alone. Conventional approaches to antibiotic development, resistance detection, and treatment selection are increasingly unable to keep pace with the evolution of resistant pathogens. In recent years, artificial intelligence (AI) has emerged as a powerful complement across the entire AMR control pipeline, from the discovery of novel antibiotics and antimicrobial peptides to the prediction of resistance phenotypes, the interpretation of genomic and mass-spectrometry data, and the personalization of antibiotic treatment. This review synthesizes the current state of research on AI-based approaches to controlling drug-resistant bacteria. It examines deep learning-driven antibiotic discovery, machine learning models for resistance gene surveillance and susceptibility prediction, and AI-assisted clinical decision support for antimicrobial stewardship. It then critically evaluates persistent challenges, including data heterogeneity and label noise, confounding by bacterial population structure, limited external validation, interpretability, and regulatory hurdles. Finally, it outlines promising directions, including foundation models, population-structure-aware learning, prospective clinical trials, and the integration of AI into One Health surveillance systems. The review concludes that AI is poised to become a central component of a data-driven response to AMR, but that its clinical impact will depend on rigorous validation, transparent reporting, and close collaboration between computational scientists, clinicians, and public health authorities.

1. Introduction

The discovery of antibiotics in the twentieth century transformed medicine, turning once-fatal bacterial infections into routinely treatable conditions. Decades of widespread and often indiscriminate use, however, have selected for bacteria that no longer respond to first-line therapies, and the global reservoir of drug-resistant pathogens continues to expand. Resistance has been documented for every class of antibiotic in clinical use, often within a short time of introduction, while the development of genuinely novel agents has slowed dramatically. The convergence of rising resistance and a depleted discovery pipeline defines the contemporary AMR crisis, and it has prompted a search for tools that can accelerate every step of the response, from finding new drugs to using existing ones more wisely.

Artificial intelligence has attracted intense interest as one such tool. Machine learning (ML) and deep learning (DL) methods are particularly well suited to problems in which the relevant signals are high-dimensional, nonlinear, and distributed across heterogeneous data sources, which is an accurate description of most problems in microbiology. In the context of AMR, these methods have been applied to tasks as diverse as predicting whether a small molecule will inhibit bacterial growth, identifying resistance determinants in whole-genome and metagenomic sequencing data, inferring susceptibility from routine mass spectra, and recommending antibiotics tailored to an individual patient's clinical history. What unites these applications is a shift from reliance on human-curated rules and single experimental assays toward models that learn directly from large collections of chemical, genomic, phenotypic, and clinical data. In several of these tasks, models now match or exceed the accuracy of expert-curated approaches, and in some cases they have produced discoveries that would have been difficult to reach by conventional screening alone.

* *Corresponding author: Mia Carter. mia.carter@gmail.com*
ISSN: 3071-4265/ This work is licensed under a Creative Commons Attribution 4.0 International License.
DOI: 10.64904/20260908

This review examines the use of AI in controlling drug-resistant bacteria, with three objectives. First, it summarizes the scale of the AMR burden and the biological mechanisms that make resistance a moving target. Second, it synthesizes the principal application areas of AI against AMR: the discovery of new antibiotics and antimicrobial peptides, the prediction and surveillance of resistance at genomic and clinical levels, and the optimization of treatment decisions. Third, it appraises the obstacles that currently separate promising models from dependable clinical and public health tools, and identifies the directions most likely to translate AI-based approaches into measurable reductions in resistant infections.

Literature for this narrative review was identified through structured searches of PubMed, Web of Science, Scopus, and the Crossref index, supplemented by snowballing from the reference lists of retrieved articles. Search terms combined domain terms (artificial intelligence, machine learning, deep learning, neural network) with AMR-specific terms (antimicrobial resistance, antibiotic discovery, antimicrobial peptide, resistance prediction, antimicrobial stewardship, hospital-acquired infection), and the search was restricted to English-language publications from 2015 to 2026. Priority was given to original research and clinical trials published in peer-reviewed journals, together with high-quality systematic reviews and the reference reports of international health organizations. Because this is a narrative review rather than a systematic review, no formal screening protocol or risk-of-bias assessment was applied; instead, landmark studies that introduced widely replicated approaches were privileged, alongside recent analyses that interrogate the validity of earlier results. The discussion is organized thematically rather than chronologically, claims about clinical utility are distinguished from demonstrations of technical feasibility, and, throughout, the emphasis is on what the evidence actually supports.

2. The Global Burden of Antimicrobial Resistance

Quantifying the human cost of drug-resistant infection has become an essential foundation for policy and research. In the most comprehensive global assessment to date, Murray et al. [1] estimated that in 2019 bacterial AMR was associated with 4.95 million deaths worldwide, of which 1.27 million were directly attributable to resistance. On this measure, AMR claimed more lives in that year than either HIV/AIDS or malaria, and the burden was concentrated in low- and middle-income regions, with the highest age-standardized mortality rates in sub-Saharan Africa and South Asia. The same analysis identified a small set of pathogens, led by Escherichia coli, Staphylococcus aureus, Klebsiella pneumoniae, Streptococcus pneumoniae, Acinetobacter baumannii, and Pseudomonas aeruginosa, as responsible for the large majority of attributable deaths. These figures, produced from a systematic synthesis of hospital data, surveillance systems, and published literature, have become the standard reference point for the scale of the problem, and they underpin international calls for coordinated action on antibiotic stewardship and new drug development.

Several biological features make this burden difficult to reverse. Bacteria acquire resistance through multiple, often simultaneous routes: mutations in drug targets, enzymatic inactivation or modification of antibiotics, active efflux, reduced outer-membrane permeability, and the horizontal transfer of resistance genes carried on plasmids and other mobile genetic elements. Horizontal gene transfer is particularly consequential because it allows resistance determinants to spread across species boundaries and across ecological niches, linking clinical settings, livestock, and the environment in a single interconnected reservoir. The result is that resistance phenotypes are shaped by both vertical evolution within lineages and the lateral movement of genetic material between lineages, a duality that complicates prediction and surveillance: a resistance mechanism may be rare in one region and dominant in another, may emerge abruptly, and may be encoded by genes that are highly diverse in sequence while identical in function.

The clinical consequences extend beyond treatment failure. When first-line therapy is ineffective, patients require broader-spectrum drugs, longer hospital stays, and more invasive procedures, each of which increases the risk of adverse outcomes and of further resistance selection. Empirical prescribing, the dominant mode of antibiotic use in most settings, is inherently probabilistic, and the probability of choosing a drug to which a given infection is susceptible has been declining as resistance rises. The economic burden, although more difficult to quantify, includes the costs of prolonged hospitalization, intensive care, lost productivity, and the development of replacement drugs. Against this backdrop, the case for investing in new methods of detection, prediction, and therapeutic decision-making is compelling, and it is precisely at these points that AI-based approaches are being developed.

3. Artificial Intelligence in Antibiotic Discovery

3.1. Deep Learning Screening of Chemical Libraries

The conventional route to a new antibiotic requires screening large chemical libraries against target pathogens, followed by years of medicinal chemistry and clinical development. Machine learning offers a way to prioritize which compounds are worth testing, effectively learning the implicit rules that distinguish antibiotic activity from inactivity from large collections of labeled molecules. Instead of testing every compound in a library against every pathogen, a model can score candidates by predicted activity, allowing researchers to concentrate experimental effort where the probability of success is highest and to explore chemical space far beyond the collections that a laboratory could physically screen.

The potential of this approach was demonstrated decisively by Stokes et al. [2], who trained a message-passing neural network on more than 2,300 molecules with known activity against E. coli and used the model to screen the Drug Repurposing Hub, a library of more than 6,000 compounds with established safety profiles. The model identified halicin, a compound originally developed for other indications and structurally divergent from known antibiotics, as a potent inhibitor of E. coli. Halicin displayed broad-spectrum activity against a panel of resistant pathogens, including carbapenem-resistant Enterobacteriaceae and A. baumannii, and was effective in a mouse model of A. baumannii infection. Because halicin was drawn from a repurposing library, its existing safety data accelerated the path toward potential clinical evaluation, and subsequent mechanistic work suggested that it acts by disrupting the proton motive force across the bacterial membrane. The study also identified several additional hits, illustrating that deep learning can enrich the output of library screens far beyond what random testing would produce. The importance of the result lies not only in the molecule itself but in the demonstration that a neural network trained on modest amounts of labeled data could generalize to find active compounds in chemical space far from its training set.

Subsequent work has refined both the targets and the generality of this strategy. Liu et al.[3] applied deep learning to the discovery of antibiotics specifically active against A. baumannii, a Gram-negative pathogen that international priority lists rank among the most critical targets for new drug development. Screening a focused compound library, the model identified abaucin, a compound with narrow-spectrum activity against A. baumannii,

including carbapenem-resistant clinical isolates. The narrow-spectrum profile is noteworthy because it may reduce collateral damage to the commensal microbiota, a frequent concern with broad-spectrum agents, and because resistance selection may be slower when the drug does not pressure the entire microbial community. Mechanistic studies revealed that abaucin disrupts the Lol lipoprotein trafficking system, a transport pathway that is essential in A. baumannii, and the compound showed efficacy in a mouse wound infection model. The identification of a species-selective mechanism by an in silico screen illustrates how deep learning can move beyond generic antibiotic activity toward precision targeting of particularly problematic pathogens, a direction that aligns with the growing recognition that narrow-spectrum therapy is often preferable from both a patient and a stewardship perspective.

A further step toward practical deployment was taken by Wong et al. [4], who addressed the interpretability gap that had limited earlier models. Using graph neural networks trained on molecular structures, the authors generated predictions accompanied by explainable rationales: chemical substructures that the model associated with antibacterial activity. This design allowed the model to screen more than 12 million commercially available compounds while flagging the structural basis of each prediction, and it led to the discovery of a structural class of antibiotics with potent activity against methicillin-resistant S. aureus (MRSA). The lead compounds were active in mouse models of MRSA infection and exhibited low toxicity, and their mechanism of action was distinct from that of existing anti-staphylococcal drugs. By making the model's decisions legible to chemists, the work demonstrated a template for human-AI collaboration in which computational suggestions can be inspected, prioritized, and optimized by domain experts rather than accepted as black boxes.

Taken together, these studies establish deep learning as a credible engine for antibiotic discovery. The field is nevertheless at an early stage: most candidates have been validated in vitro and in small-animal models, and none has yet completed the clinical trials required for regulatory approval. Still, the speed, cost, and chemical novelty advantages of the approach, combined with the ability to target specific resistant species, make it one of the most actively pursued applications of AI in AMR control, and the pace of follow-up work suggests that the first clinical candidates are a realistic medium-term prospect.

3.2. Artificial Intelligence for Antimicrobial Peptide Discovery

In parallel with small-molecule discovery, AI has been used to identify antimicrobial peptides (AMPs), short amino-acid sequences that kill bacteria through mechanisms often involving membrane disruption. AMPs are attractive candidates because they tend to act rapidly, target multiple sites, and exhibit lower propensity for resistance selection than conventional antibiotics, although their clinical development has been slowed by stability, toxicity, and delivery challenges. The same screening problem that plagues small-molecule discovery applies with even greater force to peptides: the space of possible sequences is astronomically large, and the number of experimentally validated AMPs is tiny by comparison.

A landmark contribution came from Ma et al. [5], who developed a deep learning framework, combining meta-learning with integrated-gradient attribution, to mine the human gut microbiome for peptide sequences with antimicrobial potential. The approach was designed to overcome the scarcity of experimentally validated AMPs relative to the enormous sequence space encoded in metagenomes. By transferring knowledge across related prediction tasks and using attribution analysis to identify the sequence features driving predictions, the model prioritized candidate peptides for experimental testing. Of the candidate peptides synthesized and assayed, 181 displayed antimicrobial activity, corresponding to a positive rate of approximately 83%, and many were active against Gram-negative pathogens, including multidrug-resistant isolates. The results demonstrated that the gut microbiome, a reservoir of largely unexplored functional peptides, can be systematically mined with machine learning, and they provided one of the clearest demonstrations to date that metagenomic sequence data can be converted into experimentally validated therapeutic leads.

The significance of this work extends beyond the specific peptides identified. It showed that deep learning models trained on sparse experimental labels can be made sufficiently reliable to guide wet-lab testing at scale, converting a needle-in-a-haystack search into a tractable pipeline in which the model does the prioritization and the laboratory does the confirmation. Subsequent efforts in the field have extended the same logic to the global microbiome, scaling sequence mining across thousands of metagenomes and broadening the chemical diversity of predicted peptides. The practical bottleneck has accordingly shifted from discovery to development: translating peptide hits into drugs requires optimization of stability, immunogenicity, and delivery, stages at which AI-guided design is only beginning to make inroads, and where close coupling between generative models and experimental iteration will likely be required.

4. Artificial Intelligence in Resistance Prediction and Detection

4.1. Genome-Based Prediction of Resistance

Resistance control depends on knowing which resistance determinants are present in clinical isolates and in the wider microbial community, ideally before they cause treatment failure. Whole-genome sequencing (WGS) has made it possible to detect resistance-associated genes and mutations directly from DNA, and the declining cost of sequencing has made WGS increasingly attractive as a routine diagnostic and surveillance tool. Machine learning adds a layer of interpretation that conventional annotation cannot provide: rather than matching sequences to a fixed database of known determinants, models can learn the complex, often nonlinear relationships between genomic variation and phenotypic resistance, capturing interactions between genes, mutations, and genetic background that are missed by rule-based pipelines.

The scale at which such models must operate is considerable. Yu et al. [6] assembled and analyzed more than 24,000 whole-genome sequences of E. coli and K. pneumoniae with associated resistance phenotypes, demonstrating both the feasibility of training resistance predictors on population-scale genomic collections and the analytical challenges that arise at this scale. Their work highlights a fundamental tension in genome-based prediction: bacterial populations are structured into lineages, resistance determinants are unevenly distributed across those lineages, and a model can exploit this structure to achieve impressive-looking performance without learning the biology of resistance at all. The lesson is that genomic prediction must be evaluated with methods that respect population structure, and that the promise of WGS-based resistance prediction, which is substantial, can only be realized if the models are trained and validated in ways that separate lineage effects from genuine genotype-phenotype associations. In practice, this means combining large, well-characterized training collections with evaluation designs that withhold entire lineages or phylogenetic partitions, so that reported

accuracy reflects the model's ability to generalize to unseen bacterial populations.

When these conditions are met, genome-based prediction offers clear advantages. It can detect resistance mechanisms that are not probed by conventional phenotypic panels, identify the genetic basis of resistance for epidemiological tracking, and support the rapid tailoring of therapy when sequencing can be completed within a clinically useful timeframe. Metagenomic applications extend the same logic to complex communities, allowing resistance genes to be surveyed across hospitals, farms, and environmental samples without culture, which is essential for One Health surveillance. The main limitations are shared with sequencing-based diagnostics more broadly: cost and turnaround time, the imperfect correlation between genotype and phenotype, which can be modulated by gene expression and host factors, and the need for continuously updated training data as new resistance mechanisms emerge. Machine learning does not eliminate these limitations, but it can make the interpretation of genomic data substantially more powerful, provided that the methodological discipline described above is maintained.

4.2. Machine Learning in Clinical Diagnostics

Even where genomic data are unavailable, ML can infer susceptibility from data that are already collected during routine care. Two recent studies illustrate the breadth of this opportunity, operating at different points in the diagnostic workflow, and together they show that the raw material for ML-based diagnostics already exists in most hospitals.

Weis et al. [7] demonstrated that matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectra, which are routinely acquired for species identification in clinical microbiology laboratories, contain sufficient information to predict antimicrobial resistance directly. Using a dataset of approximately 300,000 clinical mass spectra paired with more than 750,000 susceptibility test results from two hospital networks, the authors trained models that predicted resistance with an average area under the receiver operating characteristic curve (AUROC) of about 0.80 across the species-antibiotic combinations examined, while simultaneously maintaining high accuracy in species identification. Because MALDI-TOF analysis is performed within minutes of colony growth, this approach yields same-day resistance predictions, potentially days faster than conventional phenotypic susceptibility testing, and requires no additional instrumentation beyond what most laboratories already possess. The study is notable not only for its performance but for its scale, which reflects the kind of routine data reuse that makes ML diagnostics practical in real clinical settings, and for the demonstration that a single ubiquitous assay can serve a dual diagnostic purpose.

Complementing laboratory-based approaches, Yelin et al. [8] showed that ML can predict antibiotic resistance from clinical records alone, without any microbiological data for the current episode. Trained on more than 700,000 community-acquired urinary tract infections, the models predicted resistance to each of eleven antibiotics with an AUROC of up to about 0.79, and the single most informative input was the patient's personal clinical history, in particular prior infections and previously documented antibiotic resistance. This finding has a straightforward clinical corollary: when a patient has previously harbored a resistant organism, the probability that the current infection shares that resistance is substantially elevated, and an antibiotic that the patient's own history suggests is likely to fail can be avoided. The ability to generate resistance risk estimates from electronic records, before culture results are available, positions ML as a tool for empirical prescribing, the point at which most antibiotic decisions are actually made, and it suggests that even settings without advanced laboratory capacity can benefit from resistance prediction built on administrative data.

These complementary studies highlight a broader principle: the most clinically useful AI models may not require new assays but rather the systematic reuse of data that hospitals already generate. Susceptibility results, mass spectra, electronic medical records, and prescribing histories collectively constitute a rich resource that, when harnessed at scale, can reduce the delays and uncertainties that currently drive empirical broad-spectrum therapy. They also illustrate an important division of labor: some models operate at the point of care using only patient history, while others operate in the laboratory using instrument data, and the two can be combined, with the former guiding initial therapy and the latter refining it as laboratory information becomes available.

A further application of ML in infection control is the prediction and early detection of hospital-acquired infections, which are frequently caused by drug-resistant organisms. Models trained on electronic health record data, including vital signs, laboratory results, and antibiotic exposure, can estimate the risk that a patient will acquire a resistant infection during hospitalization, allowing infection prevention teams to direct barrier precautions and surveillance toward high-risk patients. The challenges inherent in such models are well illustrated by Neuman et al. [9], who developed an XGBoost model to predict hospital-acquired Acinetobacter baumannii infection in intensive care unit patients using the MIMIC-III database. The model achieved an area under the ROC curve of approximately 0.76 internally, but its discrimination fell to about 0.62 when adapted and evaluated on an external cohort, underscoring how much performance can degrade when models cross institutional boundaries. Even at moderate accuracy, such tools can support outbreak detection by flagging unusual clusters of resistant isolates, and their integration with automated surveillance may shorten the interval between the emergence of a resistant clone and an effective response.

4.3. Personalized Treatment Optimization

The ultimate purpose of resistance prediction is better treatment decisions. Stracy et al. [10] moved from prediction to prescription by building machine learning models that recommend specific antibiotics for individual patients on the basis of their microbiological history. The models were developed and validated on large retrospective cohorts of Enterobacteriaceae infections, including both urinary tract and bloodstream infections, and were evaluated against the treatments that patients actually received and against standard empirical guidelines.

The central result was that machine-learning-personalized recommendations were associated with a lower risk of resistance and reinfection than either the treatments actually prescribed in practice or guideline-based empirical choices. The improvement reflects the models' ability to exploit patient-specific information, such as prior isolates and their susceptibility profiles, that generic guidelines cannot incorporate, and it quantifies, in real clinical cohorts, the cost of treating patients as statistical averages rather than as individuals with a documented microbiological history. The work also addressed an important practical question, namely whether such models can be deployed with minimal data infrastructure: the authors showed that useful predictions could be generated from a small set of structured variables, easing integration into hospital information systems and lowering the barrier to adoption in resource-limited settings.

The clinical logic of this approach is compelling. Empirical therapy is inherently probabilistic, and its accuracy determines both patient outcomes

and the selective pressure that drives resistance. When a model can shift prescribing toward antibiotics that the individual patient's bacteria are more likely to be susceptible to, it simultaneously improves the probability of cure and reduces the use of unnecessarily broad agents, the classic goals of antimicrobial stewardship. The evidence so far is retrospective, and prospective interventional trials will be required to establish whether the predicted benefits materialize in practice; nevertheless, the framework represents one of the clearest routes by which AI can exert a direct, measurable influence on AMR control, because it acts at the moment of prescribing, when resistance selection and patient outcomes are most directly shaped.

Evidence that such decision support can change prescribing at scale is now emerging from randomized trials. The INSPIRE program, conducted across 59 US community hospitals, evaluated computerized provider order entry prompts that provided real-time, patient-specific estimates of the risk of multidrug-resistant organism infection. In the urinary tract infection trial, Gohil et al. [11] found that prompting clinicians to select standard-spectrum antibiotics for patients at low estimated risk reduced empiric extended-spectrum antibiotic days of therapy by 17.4% (rate ratio, 0.83; 95% CI, 0.77–0.89) across more than 127,000 admissions, without significant changes in hospital length of stay or days to intensive care unit transfer. The trial demonstrates that algorithmic risk stratification, embedded in the prescribing workflow and coupled with feedback and education, can reduce unnecessary broad-spectrum antibiotic use without measurable harm to patients, providing some of the first randomized evidence that AI-based stewardship translates into real-world improvements in prescribing. Table 1 summarizes the representative applications discussed above, showing the model approach and principal reported performance for each.

Table 1. Representative applications of artificial intelligence in controlling drug-resistant bacteria.

Application area	Representative study	Model approach	Principal result
Antibiotic discovery	Stokes et al. [2]	Message-passing neural network	Identified halicin; broad-spectrum activity against resistant pathogens
Antibiotic discovery	Liu et al. [3]	Deep learning library screening	Identified abaucin; narrow-spectrum activity against A. baumannii
Antibiotic discovery	Wong et al. [4]	Explainable graph neural network	Identified a new structural class active against MRSA
Antimicrobial peptide discovery	Ma et al. [5]	Meta-learning with integrated gradients	181 active peptides; approximately 83% positive rate
Resistance prediction (MALDI-TOF)	Weis et al. [7]	Machine learning on mass spectra	Average AUROC approximately 0.80
Resistance prediction (clinical records)	Yelin et al. [8]	Machine learning on EHR data	AUROC up to approximately 0.79
Hospital-acquired infection prediction	Neuman et al. [9]	XGBoost on ICU data	Internal AUROC ~0.76; external AUROC ~0.62
Personalized treatment	Stracy et al. [10]	Machine learning	Lower risk of resistance and reinfection than empirical therapy
Stewardship decision support	Gohil et al. [11]	Risk-based CPOE prompts	17.4% reduction in empiric extended-spectrum antibiotic use

5. Challenges and Limitations

Despite the enthusiasm surrounding AI-based approaches, a growing body of work cautions that model performance in research settings may overstate what can be achieved in practice. The most important recent contribution in this vein is the analysis by Yu et al. [6], who examined more than 24,000 whole-genome sequences of E. coli and K. pneumoniae with associated resistance phenotypes. The authors demonstrated that machine learning models trained to predict resistance from genomic data frequently learn population structure rather than causal resistance determinants: because susceptible and resistant isolates cluster in distinct lineages, models can achieve high cross-validation accuracy by recognizing which lineage a genome belongs to, without learning anything generalizable about the genetic basis of resistance. When the models were evaluated in a population-structure-aware manner, performance dropped markedly, exposing the inflated estimates produced by naive cross-validation. The magnitude of the effect was such that models which appeared highly accurate under standard evaluation were barely better than chance under structure-aware evaluation, a striking illustration of how evaluation design, rather than model quality, can determine reported performance.

This finding has sobering implications for the field. It suggests that some published performance figures for genomic resistance prediction are optimistic, and that models developed in one geographic or clinical setting may fail when applied to populations with different lineage compositions. It also implies that the interpretability of models is not merely a convenience but a scientific necessity: only by inspecting the features that drive predictions can researchers determine whether a model has learned biology or artifact. The appropriate response is methodological rather than pessimistic. Population-structure-aware evaluation, phylogenetic partitioning, and explicit modeling of lineage effects should become standard practice, and reported performance should distinguish within-lineage predictive power from between-lineage generalization. Journals and reviewers have a role to play here, by insisting that genomic prediction studies report both standard and structure-aware performance estimates.

Beyond population structure, several obstacles are common across the applications discussed in this review. Data heterogeneity is pervasive: resistance phenotypes are generated by different testing methods, breakpoints, and clinical contexts, and labels are often noisy, especially when derived retrospectively from routine records. Small sample sizes, class imbalance, and the rarity of resistance to certain drugs further complicate training and can produce models that are confidently wrong for the very cases that matter most. External validation, the gold standard for assessing generalizability, remains rare relative to the number of published models, and few studies have been evaluated prospectively in the settings where they would be deployed; models trained on data from one hospital or one country may perform substantially worse elsewhere, and the

magnitude of this degradation is rarely quantified. Interpretability, although advanced by explainability methods such as those used in antibiotic discovery, remains incomplete for many deep learning models, which can impede both clinical trust and regulatory review. Finally, the regulatory pathway for AI-based diagnostics and clinical decision support is still evolving, and questions of liability, data privacy, and equity, particularly the risk that models trained on unrepresentative populations will widen existing disparities, have yet to be resolved. Each of these issues is tractable, but none will be solved by algorithmic innovation alone.

The equity implications of AI-based AMR control deserve particular emphasis. As documented by Murray et al. [1], the burden of resistant infection is heaviest in low- and middle-income countries, yet the datasets on which most models are trained come predominantly from high-income settings with advanced laboratory and informatics infrastructure. Models developed in one context may perform substantially worse in another, and the very infrastructure on which they depend- whole-genome sequencing, MALDI-TOF instruments, and comprehensive electronic health records- is often scarcest where the need is greatest. Without deliberate efforts to build locally representative datasets, validate models across diverse populations, and design tools that function under data scarcity, AI could widen rather than narrow existing disparities in AMR outcomes. International collaborations that support data sharing and capacity building, together with regulatory frameworks that require evidence of performance across population subgroups, will be essential if the benefits of AI-based AMR control are to reach those most affected by the problem.

These challenges should not be read as arguments against AI, but rather as a specification of the conditions under which AI will succeed. Models must be trained on well-characterized, sufficiently large datasets; evaluated with methods that respect biological structure; validated externally and prospectively; and embedded in workflows with clear accountability. Meeting these conditions requires collaboration across microbiology, clinical medicine, bioinformatics, and machine learning, and it requires that the field hold itself to the same evidentiary standards applied to any other medical technology. The history of clinical prediction is full of models that worked in development and failed in practice; the distinguishing feature of those that succeeded was rigorous validation, not algorithmic sophistication, and there is no reason to expect AI-based resistance prediction to be any different.

6. Future Directions

Several developments are likely to shape the next phase of AI-based AMR control. Foundation models, large neural networks pretrained on broad biological data, are beginning to be applied to proteins and small molecules, and they promise to improve the transfer of knowledge across prediction tasks, such as antibiotic activity, toxicity, and pharmacokinetics, within a single framework. In antibiotic discovery, generative models may move beyond screening existing libraries toward proposing entirely new chemical entities optimized for resistance-evasive properties, with the explainability tools developed in recent work providing a bridge between machine suggestions and medicinal chemistry. In peptide discovery, the combination of metagenomic mining with generative design could accelerate the development pipeline from hit to lead, addressing the stability and delivery problems that currently limit clinical translation.

On the diagnostics side, the convergence of whole-genome sequencing, MALDI-TOF spectrometry, and electronic health records creates an opportunity for multimodal models that integrate complementary signals: genotype, phenotype, patient history, and clinical context. The recent emphasis on population-structure-aware evaluation should be extended to all such models, with phylogenetic partitioning and external cohorts becoming standard components of validation rather than occasional additions. Prospective, randomized evaluations, in which clinicians are randomly assigned to decision support with or without AI recommendations, will be essential to establish real-world impact on patient outcomes, antibiotic consumption, and resistance emergence; the first such trials, including the INSPIRE program in US community hospitals, have already demonstrated meaningful reductions in empiric broad-spectrum antibiotic use [11], and scaling these results to other settings and outcomes remains an active priority. Federated learning, which allows models to be trained across institutions without pooling raw patient data, offers a path toward the large, diverse datasets that generalizable models require, while respecting the privacy and governance concerns that currently limit data sharing.

The integration of AI into surveillance systems represents a further frontier. One Health surveillance, which spans human, animal, and environmental data, generates datasets far too large for manual interpretation, and ML is well suited to detecting early signals of emerging resistance, linking resistant clones across reservoirs, and forecasting the spread of high-risk lineages. Automated pipelines that couple sequencing, resistance prediction, and epidemiological reporting could shorten the interval between the appearance of a new resistance mechanism and an effective public health response, transforming surveillance from a descriptive activity into a predictive one. The same tools can support stewardship programs by identifying wards, units, or patient populations where resistance is rising, enabling targeted interventions before the problem becomes endemic.

Realizing these possibilities will require more than algorithmic advances. It will require data sharing across institutions and countries, standardized formats for phenotypes and metadata, transparent reporting of model development and limitations, and regulatory frameworks that can keep pace with iterative model updates. Equally important is the human dimension: clinicians must understand what a model can and cannot do, and computational tools must be designed as decision aids that augment, rather than replace, clinical judgment. Training in data literacy for clinicians and in clinical reasoning for computational scientists should be part of the infrastructure of the field, not an afterthought. If these conditions are met, AI offers a realistic path toward slowing the loss of antibiotics and preserving their effectiveness for future generations.

7. Conclusion

Drug-resistant bacteria represent one of the defining health challenges of the twenty-first century, and the scale of the associated mortality, estimated at 4.95 million deaths in 2019, underscores the urgency of new approaches. This review has surveyed the ways in which artificial intelligence is being applied to control resistant pathogens, spanning the discovery of novel antibiotics and antimicrobial peptides, the genomic and clinical prediction of resistance, and the personalization of antibiotic treatment. The achievements to date are substantial: deep learning has identified structurally novel antibiotics with activity against the most problematic resistant species, machine learning has extracted resistance information from mass spectra and electronic records with clinically relevant accuracy, and personalized models have outperformed standard empirical therapy in retrospective analyses. At the same time, the field faces well-documented challenges, most notably the risk that models trained on structured bacterial populations will overestimate their own generalizability, and the scarcity of external and prospective validation.

The most productive stance is neither uncritical enthusiasm nor reflexive skepticism, but rigorous, evidence-driven development in which computational innovation is paired with careful clinical and epidemiological validation. Models must be judged by their performance in the settings where they will be used, under evaluation designs that cannot be fooled by population structure, and against outcomes that matter to patients, such as cure rates, adverse events, and the conservation of antibiotic effectiveness. If the field holds itself to these standards, artificial intelligence is likely to become a cornerstone of the global effort to control drug-resistant bacteria, complementing stewardship, infection prevention, and new drug development in a coordinated, data-driven response. The technologies described in this review are advancing rapidly; the challenge now is to build the evidence base and the institutional structures that will allow them to fulfill their promise.

CRediT authorship contribution statement

Brady Wilson: Writing – original draft, Methodology, Formal analysis, Conceptualization. **Mia Carter:** Writing – review & editing, Methodology, Conceptualization.

Declaration of generative AI and AI-assisted technologies in the writing process

The free versions of Grammarly were used for spell-checking and grammar corrections. After using these tools, the author(s) reviewed and edited the content as needed and took full responsibility for the content of the publication.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interests

The authors declare that they have no competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We thank the reviewer for their knowledgeable contributions to the manuscript.

REFERENCES

[1] Murray, C. J. L., Ikuta, K. S., Sharara, F., Swetschinski, L., Robles Aguilar, G., Gray, A., Han, C., Bisignano, C., Rao, P., Wool, E., Johnson, S. C., Browne, A. J., Chipeta, M. G., Fell, F., Hackett, S., Haines-Woodhouse, G., Kashef Hamadani, B. H., Kumaran, E. A. P., McManigal, B., ... Naghavi, M. (2022). Global burden of bacterial antimicrobial resistance in 2019: A systematic analysis. *The Lancet, 399*(10325), 629–655. https://doi.org/10.1016/S0140-6736(21)02724-0

[2] Stokes, J. M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N. M., MacNair, C. R., French, S., Carfrae, L. A., Bloom-Ackermann, Z., Tran, V. M., Chiappino-Pepe, A., Badran, A. H., Andrews, I. W., Chory, E. J., Church, G. M., Brown, E. D., Jaakkola, T. S., Barzilay, R., & Collins, J. J. (2020). A deep learning approach to antibiotic discovery. *Cell, 180*(4), 688–702.e13. https://doi.org/10.1016/j.cell.2020.01.021

[3] Liu, G., Catacutan, D. B., Rathod, K., Swanson, K., Jin, W., Mohammed, J. C., Chiappino-Pepe, A., Syed, S. A., Fragis, M., Rachwalski, K., Magolan, J., Surette, M. G., Coombes, B. K., Jaakkola, T., Barzilay, R., Collins, J. J., & Stokes, J. M. (2023). Deep learning-guided discovery of an antibiotic targeting Acinetobacter baumannii. *Nature Chemical Biology, 19*(11), 1342–1350. https://doi.org/10.1038/s41589-023-01349-8

[4] Wong, F., Zheng, E. J., Valeri, J. A., Donghia, N. M., Anahtar, M. N., Omori, S., Li, A., Cubillos-Ruiz, A., Krishnan, A., Jin, W., Manson, A. L., Friedrichs, J., Helbig, R., Hajian, B., Fiejtek, D. K., Wagner, F. F., Soutter, H. H., Earl, A. M., Stokes, J. M., ... Collins, J. J. (2023). Discovery of a structural class of antibiotics with explainable deep learning. *Nature, 626*(7997), 177–185. https://doi.org/10.1038/s41586-023-06887-8

[5] Ma, Y., Guo, Z., Xia, B., Zhang, Y., Liu, X., Yu, Y., Tang, N., Tong, X., Wang, M., Ye, X., Feng, J., Chen, Y., & Wang, J. (2022). Identification of antimicrobial peptides from the human gut microbiome using deep learning. *Nature Biotechnology, 40*(6), 921–931. https://doi.org/10.1038/s41587-022-01226-0

[6] Yu, Y., Wheeler, N. E., & Barquist, L. (2025). Biased sampling driven by bacterial population structure confounds machine learning prediction of antimicrobial resistance. *PLOS Biology, 23*(12), e3003539. https://doi.org/10.1371/journal.pbio.3003539

[7] Weis, C., Cuénod, A., Rieck, B., Dubuis, O., Graf, S., Lang, C., Oberle, M., Brackmann, M., Søgaard, 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

[8] Yelin, I., Snitser, O., Novich, G., Katz, R., Tal, O., Parizade, M., Chodick, G., Koren, G., Shalev, V., & Kishony, R. (2019). Personal clinical history predicts antibiotic resistance of urinary tract infections. *Nature Medicine, 25*(7), 1143–1152. https://doi.org/10.1038/s41591-019-0503-6

[9] Neuman, I., Shvartser, L., Teppler, S., Friedman, Y., Levine, J. J., Kagan, I., Bishara, J., Kushinir, S., & Singer, P. (2024). A machine-learning model for prediction of Acinetobacter baumannii hospital acquired infection. *PLOS ONE, 19*(12), e0311576. https://doi.org/10.1371/journal.pone.0311576

[10] Stracy, M., Snitser, O., Yelin, I., Amer, Y., Parizade, M., Katz, R., Rimler, G., Wolf, T., Herzel, E., Koren, G., Kuint, J., Foxman, B., Chodick, G., Shalev, V., & Kishony, R. (2022). Minimizing treatment-induced emergence of antibiotic resistance in bacterial infections. *Science, 375*(6583), 889–894. https://doi.org/10.1126/science.abg9868

[11] Gohil, S. K., Septimus, E., Kleinman, K., Varma, N., Avery, T. R., Heim, L., Rahm, R., Cooper, W. S., Cooper, M., McLean, L. E., Nickolay, N. G., Weinstein, R. A., Burgess, L. H., Coady, M. H., Rosen, E., Sljivo, S., Sands, K. E., Moody, J., Vigeant, J., ... Huang, S. S. (2024). Stewardship prompts to improve antibiotic selection for urinary tract infection: The INSPIRE randomized clinical trial. *JAMA, 331*(23), 2018–2028. https://doi.org/10.1001/jama.2024.6259