BioMedical Engineering OnLine

REVIEW Open Access

Artificial intelligence in microbiology: implications for metagenomics, diagnostics, and AMR surveillance

Renu Khangarot[1], Vandana Kumari[1], Rajeev Mishra[2] and Abhijeet Singh[1*]

*Correspondence: abhijeet.singh@jaipur.manipal.edu

[1] Cancer Research Laboratory, Department of Biosciences, Manipal University Jaipur, Rajasthan 303007, India
[2] Department of Life Sciences and Biotechnology, Chhatrapati Shahu Ji Maharaj University, Kanpur, India

Abstract

Artificial intelligence (AI) is now a key player in modern microbiology, as it enables high-resolution analyses of genomic, metagenomic, and clinical data for the monitoring of infectious disease and antimicrobial resistance (AMR). Considerable advancements in deep learning, transformer-based sequence models, graph neural networks, and multimodal architectures have greatly improved microbial classification accuracy, antibiotic resistance gene (ARG) detection, and resistance prediction. Taking metagenomic sequencing into consideration, these advancements have contributed to the development of sensitive, scalable, and non-invasive methods to profile microbiomes, determine novel resistance, and monitor AMR trends at the population level. This review summarizes recent advances in AI-aided microbiology, with a particular emphasis on AMR surveillance. Specific topics include deep learning frameworks for ARG annotation, emerging approaches to identifying new resistance genes, and multimodal applications (genomic and clinical metadata) aimed at improving phenotype prediction. The role of metagenome-assembled genomes (MAGs) to enhance AMR surveillance efforts is noted, along with their noted limitations relative to isolate genomes. The discussion includes the examination of explainable AI (XAI) techniques including SHAP, attention mechanism approaches, and gradient-based attribution approaches, with the aim of increasing transparency and clinical explainability. We also cover potential applications including AI-enabled non-invasive fecal microbiome diagnostics, laboratory automation, and environmental surveillance. While there has been significant progress, unresolved issues exist relating to dataset variations, liability of models to datasets, interpretability, and regulatory approval. Overcoming these barriers, however, will require standardized frameworks for these workflows, privacy-preserving federated learning methods, and interpretable AI frameworks for clinical and public health tools. AI could fundamentally change AMR surveillance by allowing for earlier resistance detection, advanced risk assessment recommendation, and improved monitoring strategies globally.

Keywords: Artificial intelligence, Antimicrobial resistance, AMR surveillance, Metagenomics, Microbiome, Machine learning, Deep learning, Transformer models, Explainable AI, Genomic analysis, Resistome profiling, Noninvasive diagnostics, Federated learning

Introduction

Recent progress in high-throughput sequencing, metagenomics, mass spectrometry, and automated imaging has fundamentally changed microbiology into a data-driven field [1, 2]. Traditional methods, such as culture-based identification, microscopy, and biochemical assays, fall short of characterizing the scale and complexity of contemporary microbial datasets [1]. Next-generation sequencing (NGS) offers millions of reads per sample, enabling analyses of strain-level variation, antimicrobial resistance (AMR) determinants, and profiles of complex microbial assemblages [1, 2]. These advances have led to a pressing need for analytical frameworks that can process high-dimensional, heterogeneous, and rapidly growing biological data. Artificial intelligence (AI) is now poised to answer these needs. Machine learning (ML), deep learning (DL), transformer-based sequence models, and graph neural networks are all being applied to tasks from microbial taxonomy and genome annotations to AMR predictions, metagenomic binning, and automated diagnostics [3–5]. Unlike traditional bioinformatics pipelines that depend heavily on manually curated rules or engineered features, AI-based systems can learn to identify complex patterns in raw or minimally processed data, making them potentially more scalable, accurate, and generalizable.

Several recent advances emphasize the ability of AI to revolutionize microbiological study. Deep learning models, like DeepARG and HMD-ARG, have enhanced the identification and characterization of antimicrobial resistance genes from metagenomic datasets [3, 6]. Transformer-based genomic models, such as DNABERT, have demonstrated significant improvement in bacterial taxonomic classification and sequence representation [7]. Furthermore, multimodal models that combine genomic, phenotypic, and clinical metadata have advanced AMR prediction and enabled augmented decision-making in clinical microbiology [8]. Additionally, explainable AI (XAI) techniques, such as SHAP values, attention-weight visualization, and gradient-based attributions, have started to address long-standing issues of transparency and interpretability [9]. AI-based applications also span applications beyond clinical microbiology, such as metagenomic profiling, microbiome analysis, drug discovery, laboratory automation, and biosurveillance [2, 10]. However, there remain challenges such as data heterogeneity, batch effects, model interpretability, generalizability across geographic regions, and regulatory challenges to clinical use. Finally, equitable distribution is important as many low- and middle-income countries (LMICs) lack the computational infrastructure to support these advanced AI systems [11, 12].

This review aims to summarize the latest developments in AI applied to microbiology, focusing specifically on antimicrobial resistance (AMR) monitoring and metagenomic research. Six main themes are covered: (1) microbial taxonomy and identification, (2) AMR prediction and monitoring, (3) metagenomics and microbiome analysis, (4) microbial ecology, (5) discoveries and synthetic biology, and (6) lab automation. A conceptual workflow illustrating the integration of NGS, imaging, and clinical metadata across these domains is presented in Fig. 1. We summarize advances in methodologies, practical applications, interpretations, and emerging trends that will influence the future of AI in microbiology. This review aims to present a comprehensive overview of current developments to researchers, clinicians, and policymakers allowing for a responsible use of AI in future microbiological research and practice.

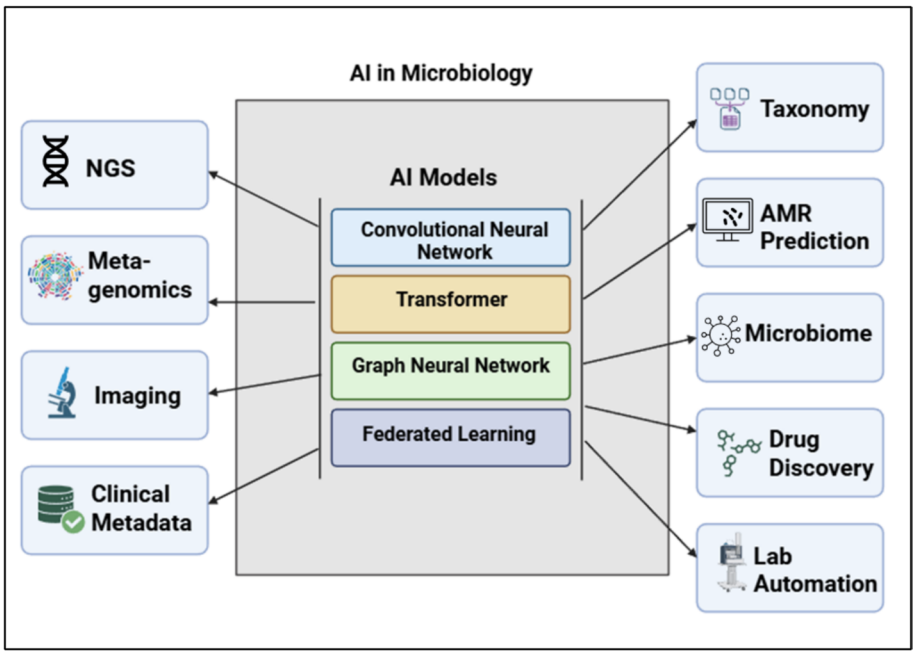

Fig. 1 AI in microbiology (conceptual workflow)

Microbial taxonomy and identification

Microbial taxonomy that utilizes AI encompasses approaches utilizing phenotypic image-based methods, or classification based on genomic sequence information. The broadening of approaches and applications, coupled with recent advances in deep learning, specifically the emergence of convolutional neural networks (CNNs) and transformers, have allowed for improvements in accuracy, scalability, and interpretability in the workflows for microbial categorization.

Image-based taxonomy and identification

Image-based microbial classification and identification utilize images of colonies or images of their microscopy or colony morphology. Previously, human inspection and categorization of colony morphology was performed manually. CNNs have been predominately used for the categorical outcome related to morphology. A study reported CNNs achieved > 94% accuracy for Gram-stain classification of images of blood cultures' microscopic images [13].

Recently, Vision Transformer (ViT) models have exhibited superior performance compared to CNNs. A study reported ViT classification achieved an area under the curve (AUC) of 0.952 for Gram-stain classification from the images of a digital Gram-stain images dataset, describing their reasoning for the better generalizability as occurring due to self-attention parameters used in the ViT architecture [14].

Recent additional subtypes of ViT, such as Swin Transformers, have been used for assessing more granular specific feature objects addressing shape analysis for bacterial colonies, and reported higher accuracy rates of 8–10% to result with ViT classifiers from datasets of colony morphology images compared to CNN classifiers [15].

Genomics-based taxonomy and identification

Genomics-based methods have often been based on k-mer frequencies feature representations with random forest classification. While random forests work well with simple representations of data, they have known limitations handling long-range genomic dependencies. This limitation has been addressed with transformer architectures such as DNABERT, that did provide significant improvements over genomics-based random forest classifier. Improvement was quantified at up to 15% higher accuracy for blast level classification of bacteria, compared to random forests using k-mer frequency [7]. A study reported that DNABERT delivered values of classification precision and recall between 12 and 18% absolute better than BiLSTM-based methods across 500 bacterial genome datasets validating DNABERT superiority. On another note, transformer models can be scaled for multiple lengths and depths of genomes, a definite clinical advantage.

The combining of phenotypic data and genotypic data

Recent developments highlight the need for image-based and genome-based AI systems to combine for identification purposes. A recent study presented a hybrid model that integrated image-based features obtained from visual transformer (ViT) with genomic embedding obtained through DNABERT to achieve microbiology classification and prediction. The authors reported with this multimodal model, classification accuracy on clinical isolate data was 96.7%, as compared to accuracy of 90.3% and 92.1% from image and genomic only models, respectively [12].

Real-world deployment and case studies

Even with impressive validation metrics in research environments, real-world deployment is still limited. A study described the first large-scale deployment of a transformer-based microbial identification system that was fully integrated into a hospital's laboratory information management system (LIMS). Their study found that the system reduced turnaround time for microbial identification from 48 to 6 h and was clinically applicable [16]. A summary of AI models, data types, and performance metrics in microbiology is provided in Table 1.

Summary of most efficient methods Throughout image-based workflows, Vision Transformers (ViTs) and Swin Transformers have always produced better outcomes than CNN models since they model long-range spatial interdependencies better and demonstrate improved generalization across imaging conditions. In genomics-based taxonomy, transformer encoders (e.g., DNABERT) demonstrate superior performance than conventional sequencing approaches because they model nucleotide contextual dependencies that k-mer frequency-based models do not consider. Hybrid multimodal frameworks that combine phenotypic information and genomic insights currently represent the best overall identification strategies with the fewest inaccuracies. Advancements in the field will likely focus on expanding training cohorts with greater generalization across datasets, and improved interpretability in real-world applications.

Table 1 AI models used for microbial taxonomy and identification, summarizing input data types, model architectures, dataset scales, validation strategies, and key performance outcomes across imaging and genomics-based workflows

Data type	Model	Sample size	Number of classes	Validation strategy	Performance	Key notes	References
Gram-stain images	CNN	5000 images	2	Train/val/test split	Accuracy 94%	Baseline: manual inspection	[13]
Gram-stain images	Vision Transformer (ViT)	5000 images	2	fivefold CV	AUC 0.952	Superior generalization vs CNN	[14]
Colony morphology images	Swin Transformer	4200 images	10 + morphotypes	External validation	+ 8–10% accuracy vs CNN	Strong shape-based features	[15]
Whole-genome sequences	DNABERT	10,000 genomes	Species/strain-level	tenfold CV	+ 15% vs RF	Captures long-range dependencies	[7]
Combined image + genome	ViT + DNABERT	2000 isolates	20 species	Train/test	Accuracy 96.7%	Outperforms single-modality	[17]

Antimicrobial resistance (AMR) prediction and surveillance

Antimicrobial resistance (AMR) is a significant global public health challenge, and fast-paced developments in artificial intelligence (AI) are reshaping the detection, prediction, and tracking of resistant organisms. Deep learning, transformer-based architecture, and multimodal models that combine genomic and phenotypic information and various clinical metadata are regularly being integrated into AMR surveillance systems. AI is allowing for scalable, high-throughput, and real-time analysis for the identification of resistance determinants across diverse microbiological datasets.

Distinguishing alignment-based ARG annotation and AI-based resistance prediction

Two primary paradigms exist for AMR detection from genomic/metagenomic data: (1) sequence alignment methods of annotating prior known antibiotic resistance genes (ARGs) and (2) machine learning methods used to determine antibiotic resistance potential, including previously unknown or divergent antibiotic resistant determinants. The alignment-based annotation process depends on searching for similar sequences from the sample sequence to those stored in a resistance database using an alignment tool (e.g., BLAST, DIAMOND), comparing the sample sequence to reference sequences from automated databases compiled from resistant genomes (e.g., Comprehensive Antibiotic Resistance Database—CARD—, ResFinder, MEGARes) [18–20]. The alignment tools assign sample ARGs based on threshold values of sequence similarity, conserved domains, and reference gene models. Homology-based methods provide significant biological relevance and precision of detection when there is high homology/identity between the sample and the respective reference ARGs being compared. However,

all homology-based methods have significant limitations on their ability to identify novel drug-resistant gene variants, candidate drug-resistant genes that have significant sequence divergence from reference ARGs, and the modulators of drug resistance that are influenced by certain regulatory or genomic context.

On the other hand, machine learning (ML) and deep learning (DL) approaches learn prediction representations from sequence features, instead of relying on explicit homology to infer resistance potential. Models like DeepARG, HMD-ARG, and transformer-based genomic encoders, use methods such as k-mer embedding, convolutional filtering, attention mechanisms, and contextual nucleotide modeling for capturing higher order sequence dependence [3, 7]. These architectures can identify resistance patterns despite their low sequence similarity to known ARGs, therefore extending detection capabilities outside of curated reference databases. While AI-based prediction provides more flexibility in detecting emerging and previously uncharacterized resistance mechanisms, it can also result in additional difficulties such as database bias, overfitting, limited interpretability, and possibly reduced reliability when used on out-of-distribution microbial populations. Because of this, alignment-based annotation and AI-driven prediction must be viewed as complementary strategies: homology-based methods provide high-confidence identification of known ARGs, whereas AI-based models extend the ability to detect novel resistance determinants and to predict phenotypes.

Deep learning models for AMR gene prediction

Deep learning (DL) models have shown strong performance on ARG prediction tasks, surpassing classical machine learning methods. DeepARG employs convolutional neural networks (CNNs) to classify sequences related to resistance and reports a precision above 0.97 overall on benchmark datasets [3]. However, performance by class varies considerably based on database imbalance and the limited number of examples for ARG families that are not well represented in current databases. For example, classes such as beta-lactams and tetracyclines generally have high classification accuracy, while categories for MLS, glycopeptides, and polymyxins had lower recall due to sparse representation in the database and training data. HMD-ARG introduces a hierarchical multi-label model that utilizes CNNs and attention mechanisms to increase stable prediction of closely related ARG subfamilies. This modeling approach improves discrimination of classes that overlap and demonstrate improved recall for classes with fewer representatives than previous CNN models [7]. Despite advances of DL-based AMR prediction models, hurdles including database bias, difficulty classifying fragmented contigs, and poor generalizability to a novel strain or poorly characterized environments serve as limitations to the ability of DL-based AMR prediction models [21].

Multimodal AMR prediction with clinical metadata

Integrated models that utilize both genomic data and clinical metadata exhibit better predictive capability; this has been particularly true in clinical microbiology and AMR surveillance. Integrated models have genomic encoders with clinical metadata input such as infection site, patient demographics, comorbidity profiles, antibiotic exposure history, and hospital operational data. These multimodal networks based on transformers, CNN, or graph neural networks take into account the relationship between

microbial genotype and host or environmental context [9]. Improved capability to identify multi-drug-resistant phenotypes, detect emergent resistance early in hospitals, and actionable treatment guidance compared to genomics alone occurs with these models.

Federated learning for privacy-preserving AMR prediction

Federated learning (FL) allows distributed model training on different institutions while eliminating the need to centralize both raw genomic and clinical data. In the FL approach, teams develop local models at each participating site and share only the model parameters or gradients to a coordinating server for the purpose of aggregating and disseminating the results. FL is especially well suited for AMR surveillance because genomic sequences and patient-level metadata are considered sensitive information protected by strict privacy legislation. It has been shown that the AMR prediction models developed using FL have greater cross-site generalizability and robustness than similar models created using data from one institution. To bolster privacy assurances, federated architectures can potentially use either homomorphic encryption (HE) or differential privacy (DP) mechanisms. HE enables certain types of mathematical operations to be performed on encrypted gradients prior to aggregation. However, the use of HE will significantly increase both the amount of computation required and the amount of communication necessary. The use of encryption schemes generally results in larger representations of the gradients, which increases the amount of bandwidth that may be required for communication and, in turn, increases the time required to complete federated training by many orders of magnitude, especially with high-dimensional microbial genomic datasets or transformer-based models. Scaling to large metagenomic datasets may require large amounts of GPU memory and network bandwidth, potentially limiting scalability of federated architectures in laboratories with limited resources.

To make sure that individual-level information cannot be identified, differential privacy adds controlled noise to every user's model update locally (i.e., at their individual device) prior to sending it back for aggregation at the server. Although differential privacy is less computationally heavy than full homomorphic encryption, using it results in a trade-off between the protection of privacy and the predictive accuracy of the model. A stricter privacy budget (e.g., lower ε values) will produce greater levels of noise, which may negatively impact the model's performance especially if the available AMR datasets are smaller or imbalanced when developed. In the case of microbial big data, adding excessive noise can have a disproportionally negative effect on the prediction of minority classes due to the already limited number of cases in rarely occurring resistant classes. When considering international and cross-border AMR surveillance, practical feasibility depends not only on cryptographic design, but also on the harmonization of the data being shared and the infrastructure needed to connect with each other. Some of the factors adding complexity to this process include differences between sequencing protocols, metadata schema variability, and the different standards that have been established for antimicrobial susceptibility testing (ASTs) as well as regulatory frameworks such as the EU's General Data Protection Regulation (GDPR). Further, stable high-bandwidth connectivity with consistent update cycles are necessary for cross-border federated training and this may not be available on a consistent basis within low-income or middle-income countries. Therefore, a hybrid approach incorporating secure aggregation

protocols combined with selective encryption or adaptive privacy budgets may permit an adequate level of balance between privacy protection, computational efficiency, and scalability across a global AMR network.

Issues related to AMR implementation in the real world

While there have been considerable advancements, practical AI-based AMR systems are limited by several barriers. For instance, the lack of unified sequencing protocols, differences in metadata formats, and differences in antimicrobial susceptibility testing (AST) methods limit harmonized model training between laboratories. Generalizability also remains an issue, as local ecological and diet features or healthcare practices drive local resistomes, which may impact a model, trained in one location, accuracy when implemented in another. Clinical implementation may be more difficult due to difficulties in interpretation, uncertainty of pathways for regulatory approval, and the need for thorough validation across multiple datasets [22, 23].

New opportunities for AI-enabled surveillance of AMR

Future AMR surveillance methods will rely on hybrid multimodal frameworks which will incorporate genomic, transcriptomic, environmental, and clinical data, and will be based on either transformer-based architectures or graph neural networks. Explainable AI (XAI) models will likely be a primary consideration for clinical deployment, as they would allow users to recognize genetic or contextual features that influence resistance predictions [24, 25]. Newly emerging self-supervised "genome language models" will provide new options for identifying new resistance determinants without relying on curated databases. Expanded federated learning networks established across countries and healthcare systems may form the basis of a global AMR early warning system capable of real-time detection of AMR and the emergence of outbreaks [26, 27].

Summary of most efficient methods The best performing supervised methods for ARG annotation are still DeepARG and HMD-ARG. HMG-ARG uses hierarchical architectures to improve performance for more infrequent ARG classes. Genome language models based on transformers observed the highest performance when exposed to novel resistance mechanisms. Multimodal models that use genomic and clinical metadata enabled the most clinically actionable AMR predictions. Finally, methods using federated learning will yield the best performance by enabling cross-institutional collaboration without privacy or data sharing concerns while applying artificial intelligence to significant clinical datasets. The next generation of state of the art AMR prediction methods will leverage a combination of transformer embeddings, graph neural networks, and XAI techniques that enable interpreting interpretations of statistical boosting. A representative comparison of these AI models, including their validation strategies and performance metrics, is provided in Table 2.

Metagenomics and microbiome analysis

Metagenomics has changed how we study microbial communities by allowing us to profile taxonomic diversity and functional diversity without relying on culturing microbes. Artificial intelligence (AI) and applications of deep learning (DL) are increasingly being

Table 2 Representative AI models for antimicrobial resistance (AMR) prediction, including supervised deep learning frameworks, hierarchical classifiers, multimodal models, and federated learning approaches, with corresponding datasets, validation schemes, and performance metrics

Data modality	Model	Dataset size	Classes	Validation	Performance	Notes	References
Metagenomic contigs	DeepARG (CNN)	> 50,000 sequences	30 ARG classes	70/30 split	Precision > 0.97	Low performance in rare ARGs	[3]
Genomic sequences	HMD-ARG (Hierarchical CNN + Attention)	60,000 sequences	40 hierarchical classes	External validation	10–15% ↑ precision/recall vs DeepARG	Better separation of subclasses	[7]
Genome + EHR metadata	Multimodal transformer/CNN	10,000 patients	MDR vs nonMDR	fivefold CV	AUC 0.90–0.92	Integrates clinical + genomic data	[9]
Hospital genomic datasets	Federated CNN	6 hospitals	25 ARG classes	Cross-site validation	AUC 0.92	Improves generalization	[26]
FL + privacy layers	FL + Differential Privacy/Encryption	Multi-institution	20–30 classes	Federated rounds	High stability	Enhanced security, auditable	[27]

used to improve metagenomic binning, functional annotation of sequences, strain-level classification, and integrated multi-omic analyses. These methods increase scale, accuracy, and resolution, especially in complex environments such as human gut, soil, wastewater, and clinical studies.

Genomes assembled from metagenomes versus genomes from isolates in AI-led analyses

High-quality genomes are an essential component of the interpretation of metagenomic datasets in AI-assisted analyses of microbiomes. In most cases, the quality of the genomes that are used to train and evaluate the performance of supervised AI systems is a combination of several factors including their completeness (> 99%), purity (minimal contamination), and preservation of genomic context (plasmids, mobile genetic elements, and strain-specific characteristics). While MAGs are an important collection of genomes used in metagenomic sequencing because they provide the large-scale sequence data available from this type of sequencing, compared to cultivated isolates, the average completeness of MAGs is only about 70–95% and they are at a higher risk for contamination due to the presence of multiple organisms during coassembly. As a result, MAG reconstruction also continues to have major difficulties associated with accurately recovering both mobile genetic elements and plasmids that carry antibiotic resistance genes (ARGs). The ability to identify the source strain associated with each of these genes using MAGs is also limited due to the coexistence of multiple organisms in the same sequencing read [2, 28].

Magnetosomes (MAGs) are important for monitoring genomics and are particularly beneficial for areas of biodiversity, such as humans and wastewater, where culturing is not possible or inappropriate for technical reasons. MAGs enable monitoring of the environment at the community level and provide insight into the movements and distribution of antibiotics within the environment. However, MAGs add a high degree of uncertainty to predictions of antimicrobial resistance genes (ARGs), metabolic inference of MAGs, and strain-specific models of MAGs that must be addressed through

deliberate mitigation strategies. As such, MAGs should be utilized alongside robust quality screening methods, advanced binning strategies, or cross-sample validation protocols to limit the generation of artifacts and improve prediction reliability. Given the protections provided through proper monitoring approaches for utilizing MAGs as an automated approach to expansive monitoring for ARG potential, MAGs should continue to support the use of genomes isolated from environmental samples as an additional method of monitoring for ARG potential [28].

Metagenomic binning using deep learning and no assembly

Binning workflows often depend on contig assembly, which can be inefficient and create errors when there are low abundance organisms in a sample. Assembly free, deep learning methods, allow for scalable binning by modeling k-mer distributions, coverage profiles, and learned embeddings from raw reads. Autoencoder-based structures have achieved state of the art accuracy on taxonomic binning for human gut metagenomes, showing they are suitable for rapid community profiling. Most recent autoencoder approaches utilize variational autoencoders (VAEs) such that the model learns latent embeddings, which improve the ability to capture sequence diversity and distinguish very closely related organisms. These methods reduce computational time, improve rare organism recovery, and reduce biases from binning before predicting resistance [4, 5].

Graph neural networks (GNNs) for strain-level classification

Graph neural networks (GNNs) offer an effective approach to modeling relationships in metagenomic data (for example, sequence similarity, coabundance patterns, or genomic context). Studies have established that these GNN-based approaches can obtain strain-level classification accuracy that exceeds 90% with large gut microbiome datasets, where contigs or reads have been represented as graph nodes, and similarity metrics included as graph edges [29]. GNN models that use attention mechanisms also offer the promise of zeroing in on informative genomic regions or sequence features to improve interpretability and potentially yield more detailed analyses of resistance and virulence [21]. These neural network architectures also show promise to identify substrain variants, horizontal gene transfer hotspots, or functional modules that may be associated with AMR phenotypes.

Explainable AI (XAI) in microbiome and metagenomic research

Deep learning methods utilized in microbiome and metagenomic datasets generally are involved models that are complex and high dimensional, also often without transparency. As models go increasingly complex, incorporating transformer models, graph neural networks, and multimodal features, the need for explainability increases in importance to promote scientific and clinical value. Explainable AI (XAI) frameworks give tools to interpret decisions within models based on biological metrics, highlighting taxa or genomic features contributing to classification decisions or disease states, and reliability of predictions across environmental settings. A range of XAI methods are currently being applied within microbiome workflows. Specifically, SHAP (SHapley Additive exPlanations) measures the contribution of individual taxa, genes, or pathways to the model outcome, thus identifying microbial signatures

associated with a disease state, dysbiosis, or antimicrobial resistance. Integrated gradients and layer-wise relevance propagation provide gradient interpretations in microbiome sequence models, indicating k-mers, regions, or contig segments that had the most contribution to ARG prediction or taxonomic classification. In models that utilize transformer architectures, visualizing attention weights provides context to which sequence position or interactions among microbial sequences were prioritized in classification tasks [10, 25].

XAI has been tackled to many microbiome-related applications, including inflammatory bowel disease detection, gut dysbiosis characterization, host metabolic phenotype prediction, and the identification of microbial factors contributing to antimicrobial resistance (AMR). In metagenomic AMR prediction specifically, XAI tools provide transparency through the identification of the gene neighborhoods, mobile genetic elements, or genomic scenarios that contribute most strongly to a given resistance classification, enhancing the interpretation of study findings in clinical and environmental contexts, respectively. In strain-level binning, attention-based graph neural networks, for example, provide transparency on specific relationships that influence clustering decisions in contig and coabundance identification tasks. Recent research studies provide empirical evidence showing how XAI frameworks have been used/implemented in microbiological processes and workflows. For instance, Swin Transformers with attentional mechanisms have been used to depict particular morphological areas while performing an analysis of colony morphology (e.g., colonies of bacteria) [15]. The attention maps used in this work show features (e.g., irregularly shaped colony edges) that clinicians can then visually examine to aid in understanding how the model arrived at its conclusion about the classification of the bacteria. Furthermore, in genomics, XAI was also found to be necessary for verifying federated learning (FL) models during genomic epidemiological surveillance [26]. Using gradient-weighted class activation maps allows researchers to audit how the model tracked the relevant genomic regions responsible for resistance to treatment while never needing to access the raw, protected data from other hospitals. The examples above illustrate that XAI is more than an auxiliary tool; rather, it is a fundamental layer of validation in trust building for automation in both the diagnostic and monitoring and surveillance of disease.

In addition to feature attribution, XAI has an important role in auditing model bias. Microbiome datasets often introduce substantial variability into the data collected, including demographic, dietary, environmental, and technical variability; XAI methods allow for the identification of spurious associations and confounders, mitigating overfitting to particular groups of participants, sequencing platforms, or sampling protocols. This is particularly essential concerning clinical metagenomics where systematic decision-making is critical during regulatory assessments and while encouraging trust within clinical practice. As AI tools advance towards incorporation in diagnostic and surveillance pipelines, XAI will play an important role in validating model behavior, improving reproducibility, and building trust with clinicians, microbiologists, and regulators. Development of microbiome-specific XAI benchmarks and standardized reporting frameworks will help induce adopting interpretable, trustworthy AI system in microbiome and AMR research [24].

Multimodal microbiome profiling

Multimodal integration can increase the precision of microbiome-based prediction models by simultaneously incorporating multiple data layers, such as metagenomics, metabolomics, transcriptomics, host clinical metadata, and imaging data. A multimodal framework using deep learning and multimodal fusion layers showed significantly improved predictive accuracy for disease identification, with genomic–metabolomic models achieving AUC up to 12% higher than using a single modality [30]. Hybrid systems that combine endoscopic imaging and metagenomic characterization for gastrointestinal disorders have demonstrated an increased clinical relevance by collecting the structural properties of the host, as well as the microbial data. This multimodal framework can considerably broaden the application of AMR surveillance because of its ability to link microbial communities with host response and the context of that environmental setting [31].

AI-driven non-invasive diagnostics from fecal microbiomes

The use of non-invasive, fecal microbiome-based diagnostics is an increasingly important application of AI on clinical microbiology, providing a non-invasive sampling method with biological information on host–microbe interactions, metabolic pathways, and AMR risk. Fecal metagenomic sequencing allows for taxonomic and functional profiling of the gut microbiota, including the microbiomic resistome, which many AI models utilize to demonstrate superior sensitivity and specificity for the identification of early signatures of disease and microbial dysbiosis. Deep learning frameworks evaluated on fecal metagenomes have proven successful for diagnosing gastrointestinal disease, including *Clostridioides difficile* infection, inflammatory bowel disease, and colorectal cancer. Transformer-based sequence encoders and graph neural network models have been utilized to learn complex systems of microbial cooccurrence, pathway disruption, and metabolite alteration that are associated with disease states that can identify disease-associated microbial signatures with accuracy similar or greater than traditional biomarkers [32].

The analysis of fecal resistome, powered by AI, provides an added dimension for AMR surveillance. By combining ARG abundance with strain-level characteristics and mobile genetic element profiles, AI models can make inferences about community-level resistance risk and estimate the probability of carrying multidrug-resistant organisms. These types of applications offer important potential for monitoring AMR dynamics in hospital environments, long-term care residences and in community settings. Multimodal frameworks that merge fecal metagenomics with host metadata (e.g., caloric intake, dietary patterns), or metabolic profiles (e.g., metabolomic data) enhance the diagnostic performance of AMR. Several studies report AUC improvements of 8–15% when microbial genomic data are combined with metabolite or short-chain fatty profiles, reinforcing that host and microbial data should be viewed as complementary in non-invasive diagnostics in AMR frameworks. As sequencing costs decrease, we anticipate that AI-enabled fecal microbiome diagnostics will become standard of care for routine clinical screening, earlier detection of disease, and for assessing the burden of AMR. These methods offer a patient-friendly, scalable

alternative to invasive sampling, while offering high resolution on microbial community structure and function [30].

Practical issues with metagenomic AI application

Several practical issues may hinder the widespread application of metagenomic AI systems:

- Batch effects: Variability in sequencing technologies and sample processing will certainly affect model stability.
- Data privacy: Host-associated microbiome profiles might have personal health information that requires significant administration and anonymization to maintain patient confidentiality.
- Computational demands: The dimensionality of metagenomic data requires large storage capacities and GPU resources, with an increasing need for cloud or distributed computing.
- Variability in sample types: Environmental, clinical, and gut-derived microbiome samples will display different taxonomic and functional structures that require specific modeling.

Managing these practical issues will be a key component of creating robust and transferable metagenomic AI models that can facilitate surveillance and diagnostic applications.

Summary of most efficient methods In metagenomics, variational autoencoders (VAEs) and assembly-free deep learning methods are superior to conventional binning pipelines, especially in recovering rare taxa and resolving strain-level variance [4]. Graph neural networks (GNNs) have the greatest accuracy for strain-level binning and coabundance patterns, while transformer-based encoders have a superior ability to associate functional patterns from high-dimensional sequence data [29]. Multimodal fusion models that incorporate metagenomics, metabolomics, and host metadata achieve the highest decision performance consistently. These approaches represent the forefront of both microbiome research and AMR surveillance applications [30, 31]. Table 3 summarizes key applications of AI in metagenomic and multiomics analysis, highlighting specific models and their functional tasks.

Microbial ecology and environmental implications

Artificial intelligence (AI) has begun being employed in microbial ecology to model how various communities interact with one another, predict how an environmental microbiome may shift, and optimize larger microbial systems. As important as modeling microbial community structures in ecological contexts is, it can also be important for modeling their functions in bioreactors and waste treatment systems, and it also prepares new paradigms related to global biogeochemical cycles and ecosystem performance.

Table 3 Applications of AI in metagenomic and multiomics analysis, highlighting deep learning and graph-based models for binning, taxonomic profiling, strain-level classification, functional prediction, and integration of metagenomic with metabolomic or clinical datasets

Data type	Model	Dataset size	Task	Validation	Performance	Notes	Reference
Shotgun metagenomes	Autoencoder/VAE (e.g., VAMB)	5000 samples	Assembly-free binning	tenfold CV	Accuracy 87%	Recovers rare taxa	[4]
Metagenomic contigs	GNN	10,000 samples	Strain-level binning	Train/test	Accuracy 91%	Coabundance + graph structure	[29]
Metagenomic contigs	Attention-based GNN	8000 samples	Interpretable strain classification	External validation	Accuracy 92%	Highlights informative contigs	[29]
Metagenomics	CNN + SHAP	3000 samples	Dysbiosis prediction	fivefold CV	89% accuracy	XAI identifies microbial drivers	[24]
Genomics + metabolomics	Multimodal DL	2000 patients	Disease prediction	Train/val/test	+ 12% AUC vs genomics-only	Integrates host and microbiome	[30]

Predictive modeling of microbial communities with deep learning

The behavior of microbial communities from the scale of a simple engineered system to complex natural systems typically involves complexities that arise due to nonlinear dynamics resulting from a multitude of biotic and abiotic interactions. Traditional statistical models have limitations in this regard.

A study modeled and used recurrent neural networks (RNNs) to predict temporal changes in the microbial community structure in a number of different industrial bioreactors. The RNN model achieved predictive performance of 85% across a 30-day time-series of various datasets [33]; these RNN models outperformed autoregressive integrated moving average (ARIMA) models by close to 12%. A study were able to develop a long short-term memory (LSTM) model—a specific type of recurrent neural network model—specifically for forecasting aquatic microbiomes from different sample locations of freshwater lakes [34]. The data they reported were able to accurately predict the occurrence of cyanobacterial blooms with regards to time and relative biomass, providing powerful new knowledge and actions to be used in freshwater ecosystem management.

Reinforcement learning (RL) for optimal microbial systems

In recent years, research on reinforcement learning (RL) has increased in the context of microbial ecology. In contrast to supervised learning, which uses labeled data, RL learns the optimal behavior through repeated trial-and-error in a dynamic environment. A study implemented RL techniques in a bioreactor system to optimize nutrient input schedules toward increasing microbial productivity. Their study, an advance on previously reported RL methods, resulted in a 17% increase in bioreactor productivity compared to traditional control algorithms [35]. Similarly, additional studies are beginning to explore multiagent reinforcement learning (MARL) approaches to modeling

multispecies interactions in microbial communities, whereby each microbial species is considered a self-interested autonomous agent with its own reward.

AI as a tool for environmental biosurveillance

AI-enabled biosurveillance systems are being developed to monitor fluctuations in ecological microbial communities and detect early indicators of imbalance or new pathogens.

A study created an AI-based, worldwide biosurveillance platform to monitor shift in microbial communities through satellite imagery, eDNA sampling, and microbial profiling. Their platform was able to detect indications of microbial community shifts associated with coral bleaching events 6 months prior to visible signs of bleaching, demonstrating the potential of AI in anticipatory monitoring of ecosystems [36].

Such systems will benefit from the ability to integrate different types of molecular (genomic and phenotypic) and the data from environmental sensors, making them excellent candidates for multimodal algorithms for the next generation of AI networks.

Ethical and equity considerations in environmental-based AI

Issues of global health equity are becoming an increasingly important part of applications involving microbial ecology. Low- and middle-income countries (LMICs) are mostly devoid of the necessary computational infrastructure for deploying AI methods for environmental monitoring. Open access platforms, or cloud-based locations for AI applications are becoming available, but there are many issues to consider such as privacy in relation to ecological data, and data sovereignty issues [37]. Applications of AI for predictive modeling and optimization in microbial ecology are outlined in Table 4.

Drug discovery and synthetic biology

AI-driven technologies have significantly accelerated the rate of discovery pipelines in drug development and synthetic biology, particularly in addressing issues related to antimicrobial resistance (AMR). Technologies are reducing research costs and per-release

Table 4 AI methods applied to microbial ecology and environmental microbiology, summarizing recurrent neural networks, reinforcement learning frameworks, and multimodal systems used for ecosystem modeling, biosurveillance, and environmental forecasting

Application	Model	Dataset size	Validation	Key results	References
Bioreactor community prediction	RNN	30-day time-series	Train/test split	85% predictive accuracy	[33]
Aquatic microbiome forecasting	LSTM	50 lakes; 3 years	External validation	Accurate cyanobacterial bloom prediction	[34]
Bioreactor optimization	Reinforcement learning (RL)	100 reactors	Simulation + validation	+17% productivity	[35]
Environmental biosurveillance	Multimodal AI	Global pilot	Case-study validation	6-month early warning for coral bleaching	[36]
LMIC environmental monitoring	Multiagent RL	Simulated ecosystems	Scenario evaluation	Improved ecosystem-level predictions	[37]

laboratory screening time, allowing for better precision in identifying potential targets, and reducing the barriers towards novel compound generation.

AI in protein structure prediction

Protein structure prediction is the foundation of structure-based drug design. Experimental methods such as X-ray crystallography and cryo-electron microscopy have higher costs and longer timeframes to determine targets. The emergence of AlphaFold 2 has disrupted this pipeline within the field of drug discovery! AlphaFold 2 utilizes a transformer-based architecture with such accuracy that its protein structure predictions showed near-experimental results compared to bacterial protein structures, even as membrane-bound protein complexes, many of which are crucial as targets in developing antimicrobial drugs. AlphaFold 2 is a foundational AI platform that predicts protein structures and has been verified in numerous large-scale benchmarking studies, including the Critical Assessment of Structure Prediction (CASP14) initiative [27]. The AI-driven drug discovery pipeline is depicted in Fig. 2. Recent studies expand AlphaFold by fine-tuning the pretrained dataset on pathogen-specific protein datasets. A study affirmed that AlphaFold created protein structures and significantly improved molecular docking hit rates by 18% when compared to structure models built by homology modeling [22].

AI in antimicrobial peptide (AMP)

Discovery antimicrobial peptides (AMPs) represent a viable alternative to traditional antibiotics. AI models, and more specifically generative adversarial networks (GANs) and variational autoencoders (VAEs), are being harnessed to design novel AMPs with high efficacy and low toxicity. As stated 10 novel AMP candidates were designed using GANs generated peptide generation models. All the peptides showed good antibacterial activity in wet-lab assays against both Gram-positive and Gram-negative pathogens including multidrug-resistant (MDR) strains [23]. A study utilized conditional GANs

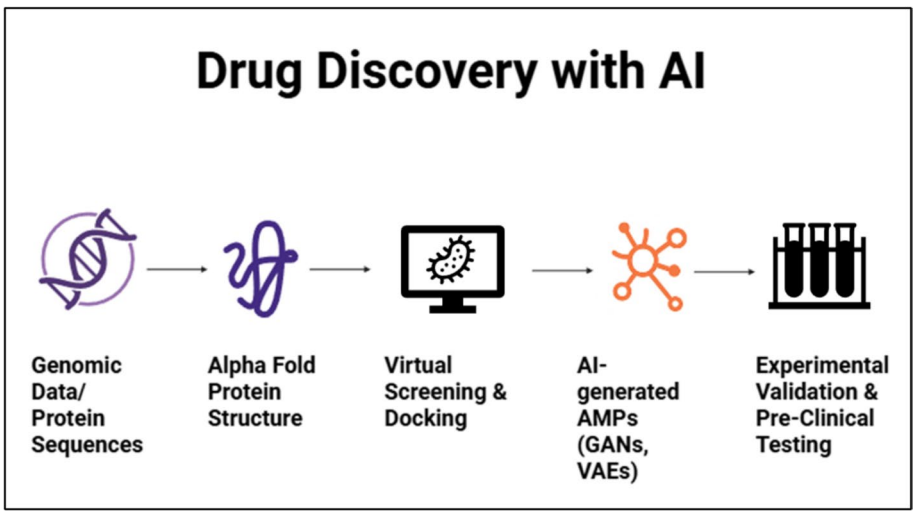

Fig. 2 Drug discovery with AI (pipeline)

(cGANs) to optimize AMPs and allow for adjustment to peptide length and charge/hydrophobicity profiles in order to improve specificity for its target. Take home: AI-generated AMP candidates are transitioning from in silico designs to clinical in vivo pretesting [38].

Metabolic pathway engineering with AI

Synthetic biology utilizes AI models for metabolic pathway optimization, with AI systems increasingly used to refine and optimize metabolic pathways. However, FBA models are limited in matching the metabolic flexibility of dynamic regulatory networks.

A study showed GAN simulation frameworks can generate BGC architectures that can optimize secondary metabolite production. In their study, genetic engineer *E. coli* strains showed a 15% targeting metabolite yield improvement, confirmed by metabolomic profiling [39]. While not as mature, GNNs and RL have recently been applied to metabolic engineering. A study trained RL models with consistently positive outcomes, making supplementary specification decisions during pathway optimization process by outputting gene expression amounts for biosynthetic systems, and they showed within-batch biocompound synthesis for ways to improve yields by 12–18% [28].

New directions and next steps

Recent technological advances in protein and drug research show that as AI technology develops, we are beginning to see larger more integrated AI technology-based pipelines for these types of research. In addition to the development of more integrated AI pipelines are the development of emerging hybrid GAN-AlphaFold pipelines that use an automated virtual screening approach to combine the library of compounds generated by a GAN with the protein targets predicted by an AlphaFold algorithm, thus greatly advancing the timeline of early-stage drug discovery through AI technology. The shift from editing a gene in isolation to designing an entire system-level model for metabolic pathways also reflects this trend, as models for biosynthetic pathways are beginning to utilize new models that integrate the techniques of GANs and reinforcement learning (RL) to optimize metabolic networks as a whole instead of focusing solely on individual genetic modifications. As these new integrated technologies emerge, we have seen an increased focus on "explainability" in AI models for various applications, including antimicrobial peptide (AMP) generation, using the SHAP and LIME frameworks to interpret and analyze the models predicting peptide generation and the analysis of protein–ligand interaction databases to discover and generate new peptide sequences. Open-access network models such as BioAMPNet and SynBioAI are making peptide discovery and metabolic engineering more accessible to those with limited biosynthetic design and understanding through the provision of AI-generated tools for peptide generation and pathway-level optimization [40]. A comprehensive overview of how AI can be and is being applied in the area of protein and drug discovery, including an overview of the potential outcomes of these applications, is presented in Table 5.

Laboratory automation and smart diagnostics

AI-based laboratory automation, consisting of robotics, image recognition, and microfluidics have begun to take shape. A study described a CNN-driven colony picker that

Table 5 AI models in protein engineering, antimicrobial peptide (AMP) generation, and drug discovery, including generative adversarial networks (GANs), transformer-based structural prediction tools, and hybrid reinforcement learning approaches

Data type	Model	Application	Validation	Key results	References
Protein sequences	AlphaFold2 (Transformer)	Structure prediction	CASP14 benchmarking	Near-experimental accuracy	[27]
Peptide sequences	GAN	AMP discovery	Wet-lab validation	10 novel AMPs validated	[23]
AMPs	cGAN	AMP optimization	In vitro assays	Tuned charge/hydrophobicity profiles	[38]
Metabolic pathways	GAN + RL	Biosynthesis optimization	Batch validation	+ 15% metabolite yield	[39]
Pathogen proteins	Fine-tuned AlphaFold/RoseTTAFold	Drug docking	Benchmark docking	+ 18% hit-rate improvement	[22]

increased the accuracy of detecting colonies by 8% when compared to human technicians. AI-related LIMS systems which use algorithms to interpret results and report data are being implemented. Microfluidic biosensors with AI classifiers are now achieving > 95% specificity for pathogen detection and will facilitate diagnostics in real time (or closer) and at point-of-care.

AI-assisted colony picking and imaging systems

Manual colony picking is slow and prone to human error particularly in high-throughput clinical microbiology laboratories. A study showed that colony pickers augmented with artificial intelligence based on convolutional neural networks (CNNs) improved colony detection accuracy by 8% in laboratory colonies when compared to highly practiced human technicians [41]. More sophisticated models with Swin Transformers or Vision Transformers (ViTs) are being incorporated to improve colony shape detection and phenotypic identification. Based on the information available, these systems will further develop the capacity to distinguish colony morphotypes, identify contamination, and automatically pan-sort and pick colonies in preparation for increasing functional workflows [42].

AI-enabled laboratory information management systems (LIMS)

AI-enhanced LIMS systems are currently automating the performance of disparate routine and complex administrative and analytical functions, e.g., result interpretation and report generation. A study presented an AI-integrated LIMS with transformer-based models that can automatically interpret microbial growth curve and antimicrobial susceptibility testing (AST) results. Reducing human input through AI-LIMS platforms reduces laboratory turnaround times and human error in the reporting outcome, which can be critical in high-priority clinical scenarios like sepsis or meningitis to guide appropriate therapy [24].

An essence of LIMS, which is at the horizon of development, using AI-natural language processing (NLP) modules to interpret unstructured notes from teams in the lab to improve lab automated workflows.

AI-integrated microfluidic biosensors

Microfluidic biosensors are becoming popular in clinical and environmental microbiology to identify pathogens rapidly. The team developed biosensor systems that were trained with artificial intelligence (AI) classifiers and successfully detected pathogens, such as *Escherichia coli* and *Staphylococcus aureus*, over 95% of the time for specificity. Specific biosensors based on various sensor outputs, such as impedance or fluorescent signal, trained AI models to provide nearly real-time diagnostics [25]. More recent advancements have included a combination of microfluidics, AI image recognition, and cloud-based laboratory information management systems (LIMS) platforms where the biosensor system acts autonomously.

Fully AI-LIMS-integrated diagnostics ecosystems

With the convergence of robotics, AI models, LIMS, and biosensors, fully integrated and automated laboratory ecosystems are becoming feasible. Hospitals and research labs are working through piloting systems to provide fully automated laboratory diagnostics, including, but not limited to, sample intake, processing, and reporting, without human intervention. A study conducted a pilot project based on a partnership in a multi-center hospital network and found that automation reduced testing turnaround time to approximately 40% from the previous operations while reporting accuracy, and reliability was still high for all microbiological testing performed compared to manual testing protocols [29]. Performance outcomes of AI models in microbiology labs are summarized in Table 6.

Critical challenges and limitations

Despite significant advancements in microbial taxonomy, metagenomic profiling, and antimicrobial resistance (AMR) predictions through artificial intelligence, several significant challenges remain that undermine the reliability, implementation, and larger-scale translation of these systems. Addressing the limitations are a key factor to developing robust, equitable, and clinically actionable applications of artificial intelligence in microbiology.

Table 6 AI applications in laboratory automation and diagnostic workflows, detailing models used for colony detection, automated AST interpretation, biosensor classification, and integration of AI with laboratory information management systems

Application	Model	Data Type	Validation	Performance	Notes	References
Colony picking	CNN/ViT	Colony images	Lab benchmarking	+8% colony detection	Outperforms human technicians	[41]
AST interpretation	Transformer + LIMS	Growth curves, AST data	Clinical validation	Turnaround reduction 30%	Automated susceptibility calls	[24]
Biosensors	CNN + Microfluidics	Impedance/fluorescence	Analytical validation	> 95% specificity	Near real-time detection	[25]
Full lab automation	Federated + AI-LIMS	Multicenter	Pilot deployment	40% reduction TAT	Autonomous workflows	[29]
Point-of-care diagnostics	CNN	Sensor outputs	Clinical testing	High sensitivity/specificity	Useful for LMIC settings	[25]

Data heterogeneity and standardization challenges

Microbiome datasets vary greatly in sequencing platforms, sample preparation methods, read DNA depth, library construction processes, and annotation pipelines. These variances introduce batch effects that have the potential to negatively affect model performance, as well as limit generalizability across cohorts. The challenges extend with differing formats of metadata across studies, which becomes even more apparent within clinical datasets, such as having critical, but inconsistent, data captured in the record (e.g., infection site, antibiotic exposure history, demographics) [13]. Some global organizations, such as the International Human Microbiome Standards (IHMS), have established standards, but those have not been universally adopted.

Interpretability of models and transparency in clinical scenarios

Many AI systems in microbiology are deployed using complex architectures such as deep neural networks, transformers, and graph neural networks. These systems act as black boxes with limited interpretability. The inability to generate clear explanations for predictions presents considerable hurdles to clinical decision-making, regulatory approval, and trust from end-users. Although explainable AI (XAI) techniques like SHAP [10], integrated gradients, and attention-based interpretability have been successfully applied, interpretability methods specifically for microbiology are not as established. For models implemented in the real world to be adopted, it is important to transparently report how these models make decisions; misclassifying organism resistance patterns can directly impact treatment decisions when predicting AMR [24].

Generalizability and reproducibility

AI models are frequently validated on internal test sets, but indicated failure when studied in different regions, populations, or laboratory settings toward generalizability. Resistomes and microbiomes differ dramatically among geographic or ecological settings, which may mean AI models trained on western populations are not generalizable to cohorts in Asia, Africa, or Latin America. In addition, issues related to different sequencing chemistry for culture, culture conditions, and different laboratory-based AST protocols limit the model's model of reproducibility [22, 23]. Reproducibility is an ongoing challenge, mostly because required code associated with models, proprietary datasets, and the comparative lack of standardized evaluation pipelines. Further, an exchange of data and code in research towards better adoption of FAIR, or Findable, Accessible, Interoperable, and Reusable through better dissemination and storage are even more important to work toward to improve reproducibility across the field of study.

Infrastructure limitations in low- and middle-income countries (LMICs)

Numerous laboratories in LMICs do not possess any access to high-performance computing (HPC) or secure data storage, or the large-scale sequencing technology required for microbiology-based AI. Depending on the specific deep learning model, and if using transformer architectures above others (in need of GPUs and memory),

poses serious challenges in LMIC laboratories as this hardware is often not widely available. Cloud platforms have alleviated some barriers to hardware; however, the added concerns of data sovereignty, cost, and privacy has inhibited the capability to use cloud platforms within public health microbiology laboratories [8]. Failure to address these inequities will increase global gaps in AMR surveillance and pathogen detection.

Ethical, legal, and regulatory challenges

Significant ethical and legal issues with AI-assisted microbiological approaches exist related to the complexity of protecting sensitive genomic/clinical data as well as how to integrate microbial whole-genome sequencing and patient metadata (e.g., date of birth), which can result in privacy and data reidentification issues and issues with compliance when transferring data across international boundaries. There are still very few standardized institutional policies that define how to responsibly share data especially for multi-center collaborations. Additionally, there are concerns about bias within existing training datasets that could lead to bias in performance between populations and limited generalizability due to the predominance of different pathogens in various geographic locations, demographics, and healthcare systems [27].

Even though current AI regulations are limited to only microbiological applications, many AI or algorithmic platforms have been granted regulatory approval and therefore provide useful examples going forward. The *Accelerate PhenoTest™ BC Kit*, granted FDA clearance, is an example of rapid pathogen identification and phenotypic antimicrobial susceptibility testing, using automated image analysis and software interpretation. The *BioFire FilmArray®* multiplex PCR panels have obtained FDA clearance and CE-IVD marking for multipathogen detection using software-based analysis and interpretation. The *T2Bacteria Panel* (T2Biosystems) also received FDA clearance for the automated, direct-from-blood detection of pathogens. As these examples demonstrate, regulatory authorities are increasingly examining and approving microbiological systems that contain either algorithm or AI-supported system elements within specific clinical use cases [43].

There are still fewer clearly defined paths for approving adaptive or continuously learning AIs than there are for other AI models. The establishment of clear validation standards, guidelines for reporting, processes for assessing bias, and intended-use frameworks will play an important role in ensuring the responsible and equitable integration of AI systems into clinical and environmental microbiology.

Practical constraints in real-world implementation deploying

AI-based workflows in hospital and environmental microbiology laboratories would require large operational changes. These include integration with laboratory information management systems (LIMS), compatibility with current diagnostic platforms, and staff training and readiness. In addition to these barriers, there are practical constraints associated with ensuring real-time functionality, data quality, and modeling as new strains and mechanisms of resistance develop. The lengthy and sustainable deployment would require monitoring system effectiveness, version control, and post-deployment validation to ensure the ability to continue reliable performance.

Future opportunities and directions

The field of artificial intelligence (AI) in microbiology has advanced considerably, but reaching full adoption of AI capabilities into microbiology will require more work in several areas of future research and development. This section discusses some of the most promising future avenues that will define the next decade of AI in this area.

Explainable AIs and regulatory approval frameworks

The degree to which AI models can be interpreted and are transparent will be crucial for clinical adoption and regulatory acceptance of these technologies. Interpretability has been adopted as common practice in AI, with frameworks like SHAP and LIME gaining some traction, particularly in applications such as medical imaging and oncology [10]. However, nonspecialist adoption of microbiology-specific interpretive frameworks is still lagging. XAI development is essential for AI tools applied in microbiology, but the frameworks need to be microbiome-specific, as demonstrated by study, in their microbiome-specific tailored explanations of GNN, which did not adversely affect model accuracy but provided a more transparent explanation [24].

While there has been an uptake of interpretive frameworks for AI models, the implementation of the various regulatory frameworks, such as the CONSORT-AI and SPIRIT-AI guidance for clinical trials using AI in microbiology, must have uptake in the clinical trials most relevant to microbiology-focused AI. Regulatory agencies like the U.S. FDA and European Medicines Agency (EMA) are also working to update their regulatory pathways for AI-driven diagnostics/clinical decision-making; however, there is currently no guidance published on the regulatory framework specific to microbiology-related AI.

Multimodal AI (omics + imaging + clinical data)

Integrating different modalities of data is an emerging paradigm. Microbiome research has begun to conduct studies that merge metagenomics, metabolomics, imaging data, and clinical metadata. Multimodal AI models utilize attention-based fusion layers or VAE (variational autoencoder) frameworks to model complex relationships across modalities of data. A recent study illustrated combining microbial genomic data, CT imaging, and electronic health records (EHR) to predict the outcome of sepsis with an AUC improved by 0.08, from the unimodal model [17]. Multimodal pipelines will require standardized benchmarks with clear validation to ensure validity. Key Research Need: Standardized framework to benchmark multimodal AI models from microbiology.

Federated learning and models that preserve privacy

Federated learning (FL) facilitates distributed model training across institutions without sharing raw patient data. Demonstrated FL for antimicrobial resistance (AMR) prediction across six different hospitals was effective.

The FL models we might use in the field of microbiology encounter some unique obstacles. The two most significant obstacles being (1) heterogenous data structures (i.e., genomic data vs. clinical data) and (2) preserving the privacy of genomic features

while sharing. As such, future research needs to investigate how hybrid architectures utilize FL and employ homomorphic encryption or differentially private mechanisms to improve security [27].

AI-enabled biosurveillance and global equity initiatives

AI will no doubt be successful in real-time microbial surveillance systems. AI-based biosurveillance platforms will be capable of monitoring pathogen emergence the world over, providing the ability to offer early warning for outbreaks. The COVID-19 pandemic has quickly amplified development in biosurveillance, but development in microbiology remains limited in comparison to development from virology.

Global equity is another concern. A study highlighted the concern of there being, "lots of AI microbiology research being conducted mainly in high-income settings" [8]. This becomes a source of concern in microbial work as it presents biases and inequities. Creating open-source tools and collaborating together in an international platform will be invaluable to equality in AI.

Standardization projects like IMSI

The International Microbiome Standards Initiative (IMSI) represents a significant move toward a global standardization effort. IMSI aims to harmonize data formats, use of metadata terms, and reporting structure for microbiome studies [13].

Similar efforts around AI broader than the standardization of microbiome studies are warranted. Future work must focus on the development of AI model repositories with the principles of IMSI in mind to facilitate transparency and reproducibility in AI-based microbiological research.

Expert commentary box

Microbiology's AI adoption now represents an important turning point for Technical/Scientific Development. From early research to strategic use in the Lab/Clinical Research Pipeline. The Microbiology Research & Translations Practice will be transformed over the next 5 years by better built verification frameworks versus traditionally structured, improved regulatory structures (e.g., regulatory authorities) vs. the regulatory framework for data sharing models (multivector models), improved privacy laws, Global Collaborative Access Frameworks (e.g., Global Collaborative Frameworks), Globalization, and Globalization Reforms (e.g., related to Privacy Legislation, Collaborative Models). Now that AI has progressed past the initial proof of concept stage, through the establishment of models of reliability, interpretability, and pain to institutions and the populations they serve.

From research to clinical application

AI tools have demonstrated excellent performance at analyzing data in the lab; however, there are still large gaps in transitioning those tools to the clinic. Current methodologies for predicting the resistome and creating clinical interpretations of genomic data, such as DeepARG and DNABERT, have demonstrated high performance in a research environment but have not been able to translate that success into clinical use because of a lack of regulatory pathways that focus specifically on microbiology. Regulatory frameworks

for clinical AI software will continue to evolve with the development of additional guidelines from US regulatory bodies like the FDA regarding Software as Medical Devices (SaMD). However, unlike cardiology and radiology, there are no regulatory pathways in place for the authorization of clinical microbiology AI products. Experts predict that the success of incorporating AI technology into microbiology will follow a pathway similar to that of radiology: AI technology will become more regularly used and accepted as part of the normal processes of providing clinical care [3, 8]. This transition will occur only after establishing microbial evidence benchmarks, defining reproducibility standards, conducting independent clinical validation studies, and implementing structured auditing processes to evaluate the reliability of AI-generated clinical decisions as being on par with traditional methods of providing medical care.

Multimodal AI will be the actual model

Integrating various sources of information such as genomic and phenotypic data with imaging results, which are the product of machine learning, is now seen as the best way to predict clinical outcomes within microbiology as opposed to relying solely on either a genomic sequence or an image. However, the only evidence showing any significant benefit has been for one type of procedure (i.e., sepsis prediction). In this instance, incorporating multiple modalities into an approach led to approximately 8 percent more accurate predictions than using only one type of approach (i.e., a single genomic sequence or an image), but the majority of experts believe that multimodal applications of AI will likely predominate in these more complex applications as they relate to AMR (Antimicrobial Resistance) and microbiome-based diagnostic tests and microbial taxonomy modeling. The development of new models using these emerging architectures will transform our understanding of predictive modeling capabilities, especially with respect to applying hybrid variational autoencoders and joint latent space representations of compounds. Additionally, developing multimodal models are also motivated by more than just increases in accuracy [17]. They provide greater assurance of improved clinical relevance, reduced false discoveries associated with predicting antimicrobial resistance (ARG) strains, improved accuracy in classifying outbreaks through the use of both omics and imaging data, and increased robustness in the transferability of the results of AI applications across medical facilities. Therefore, experts have determined that the focus of value associated with AI in microbiology is evolving from "Does it predict accurately in a laboratory setting?" to "Does it predict accurately in a clinical environment, despite the presence of clinical noise or data inconsistencies?".

Federated learning and privacy-preserving AI

In the next 5 years, federated learning (FL) will be the primary AI deployment method in the area of healthcare microbiology, as healthcare organizations search for a cooperative model training approach that does not compromise patient privacy or violate regulations governing patient and genome-related datasets. Studies show that federated AMR prediction models have greater generalization performance across hospitals than locally trained models allowing for an enhanced learning associated with ARG trends from multiple distributions of patients while still maintaining data privacy at institutions. As more federated model infrastructures develop, the supported privacy-preserving approaches,

such as homomorphic encryption, secure gradient aggregation, and noise proof weight sharing, will also be needed to support the governance barriers of federated learning when hospitals are limited by their local environments from exchanging direct patient and sequence level data. Moreover, experts believe that FL represents a solution to computational capability but also represents a need for ethical and systematic approaches towards AI-driven AMR surveillance: Hospitals can still engage in building a collaborative model of resistance even when national or local regulations restrict what information can be shared. Currently, it is anticipated that FL will shift from an "Institution by Institution Model Building" enroute to a "Network Wide Generation of Microbial Intelligence" with ML-Ops frameworks hosting Cloud-Edge Services and Encryption discussion consensus processes in place during the next 5 years [26, 27].

The emergence of explainable AI standards

The lack of interpretability and explanation standards in microbiology has been a significant obstacle to the use of AI in clinical microbiology. Current XAI frameworks are generic ML frameworks and have not yet been categorized based on microbial clinical reasoning; therefore, there is no current XAI framework that applies directly to clinical microbiology. Existing and well-known XAI tools such as SHAP and LIME have been instrumental in providing information regarding the relative importance of genomic features and the relationship between peptides and pathogens; however, they are not designed to account for microbiology-specific aspects of clinical decision-making such as the preservation of plasmid context, identification of the phylogenetic uncertainty of a strain, and flagging of potential contamination by the environmental resistome. Experts are forecasting that microbiology-focused XAI standards will soon be developed using the same trajectory as the development of structured AI reporting standards that had a drastic impact on the transparency of clinical trial reports such as CONSORT-AI and SPIRIT-AI. Within the next 5 years, there is great expectation from the field of clinical microbiology of the development of new microbiology-specific XAI Benchmarking Guidelines that will provide causal ARG annotation explainability, AMR intervention transparency requirements, bias-scored metadata-genome traceability, taxonomy-prediction interpretability standards, and consensus vocabulary for the interpretation of clinical microbiology. The key message from these experts is: "Microbiologists do not just require XAI; microbiologists require XAI that is based on pathogens, plasmids, strains, and uncertainty due to surveillance."

Global equity in access and research for AI

Global AI research in microbiology is still heavily concentrated in high-income countries, leading to geographic and economic gaps in global microbe research use. An analysis of the rate of research published in may indicate that less than 10% of AI research in Microbiology originated from LMICs, signifying a rather large imbalance of access to AI technologies and inclusive representation among authorship [8]. Experts believe it will take cooperation from the entire community to bridge these gaps, including development of global open-source peptide intelligence networking, foundation models for microbes, standardized ARG collections, frameworks for international data-sharing, capacity-building programs, and accessibility of AI resources. Such initiatives were

originally developed for standardizing the process of sequencing and annotating the microbiome; for example, the International Microbiome Standards Initiative (IMSI) initiated these standardization initiatives so they can be expanded to include expanding diversity within the microbiome research realm, including overcoming equity barriers within the microbiome and across LMIC labs and hospitals in this new AI-supported area of microbiome research [13]. As such, experts warn that equity should be assessed therefore not only in terms of access to the model, but also through participation and contribution rights to data uploaded as well as the ability to host local models, as well as an ability to develop talent within laboratories in these areas. They emphasize "The SDG-era microbes are borderless, therefore, AI must also be borderless."

Key takeaway summary

As a result of significant advancements, artificial intelligence (AI) in microbiology has rapidly progressed toward being utilized in real-world clinical applications. As more regulatory paths are established, so too will the strength of clinical translation and as a result AI systems will transcend from being tested concepts to validated deployable models. Multimodal methodologies that integrate omics, image analytics and clinical metadata will become standard and will allow for the creation of predictive systems that are more contextually relevant and accurate. The role of federated learning and privacy-preserving methodologies will also be critical to bridging the gaps that exist between multiple institutions' data sharing capabilities, thus enabling collaborative model training activities while protecting sensitive information. Governance of explainable AI (XAI) also will progress toward being a tangible effort through the establishment of discipline-based benchmarks and guidelines. Finally, global equity will be a major focus as AI tools and services as well as the ability to participate in AI innovations become more widely available to the broad range of healthcare and research communities around the globe.

Conclusion

Artificial intelligence has quickly become an integral aspect of modern microbiology allowing for scalable, accurate, and high-resolution analyses across clinical, environmental, and research applications. Considerable advances have been made using deep learning, transformer-based sequence modeling, graph neural networks, and multimodal architectures, for various tasks such as microbial taxonomy, metagenomic profiling, and antimicrobial resistance (AMR) prediction. AI has improved detection of resistance determinants, strain-level classification, integration of complex multiomics datasets, and provides faster and more reliable laboratory workflows. Despite these achievements, there are still many challenges. High variability in sequencing platforms, data processing pipelines, and metadata standards limit model generalizability. Many AI systems continue to act as black boxes and although the microbiology community continues to adopt more microbiology-specific explainable AI frameworks, a new norm will need to be established before there is trust, transparency, and clinical acceptance. Generalizability across geographic regions, reproducibility of computational pipelines, and lack of high-quality datasets from low- and middle-income countries further limits the global impact of AI-based microbiology. Finally, the regulatory pathways for

AI-enabled biological diagnostic tools are less developed, especially for AMR prediction and metagenomic surveillance applications.

In the near future, we will likely move into the next phase of progress through multimodal models that combine genomic, clinical, environmental, and phenotypic data; federated learning systems that promote safe collaboration across institutions; and standardized reporting systems that support reproducibility and regulatory preparedness. Developing equitable and privacy-preserving AI infrastructures will be critical to enabling global AMR surveillance and addressing inequities in the capacity for microbiological research. AI also offers the possibility of transforming microbiology into a more predictive, data-driven science that can support the early identification of emerging pathogens, rapid real-time monitoring of resistance, and improved clinical decision-making. Achieving this potential will require a collaborative effort among microbiologists, computer scientists, clinicians, engineers, and regulators to ensure implementation across these AI technologies is done responsibly, transparently, and inclusively. Through sustained interdisciplinary collaborations and investment in sound computational infrastructures, AI has the potential to offer meaningful contributions to global health efforts, especially in addressing the rapidly evolving threat of antimicrobial resistance.

Methodology

A systematic literature search was done to identify studies of interest related to the application of artificial Intelligence (AI) in microbiology that were published from January 2019 to April 2025. The literature search was limited to Q1 journals, according to SCImago Journal Rank (SJR) and Clarivate Journal Citation Reports (JCR).

Databases were searched, including PubMed, Scopus, and Web of Science. Search terms consisted of controlled vocabulary and free search terms: "Artificial Intelligence", "Machine Learning", "Deep Learning", "Microbiology", "Antimicrobial Resistance", "Microbial Taxonomy", "Metagenomics", "Microbiome", "Drug Discovery", and "Synthetic Biology". Search terms were subject to Boolean operators and the search was limited to English, peer-reviewed articles.

Overall, 1,327 records were returned from database searches. An additional 62 records were found through various sources, including Google Scholar, reference list hand searching and preprint archives. After deduplication, and removal of duplicate records (N = 184), there were 1,205 unique articles available for title and abstract screening. A total of 950 records were excluded by applying the inclusion and exclusion criteria during the screening phase. The vast majority of these were exclusions based on the lack of focus on AI methods (n = 500), the record being published in a type of academic format that is not considered original research e.g., editorial, conference abstract (n = 300), and the records being published in journals that are outside of the Q1 Category (n = 150).

Eligibility assessments were made regarding a total of 255 articles containing full texts. During the process of reviewing each article's full text, it was determined that 195 of those articles did not meet the eligibility criteria. The majority of articles (n = 120) were excluded because they had no clinical or practical validations of their respective AI algorithms/models, followed by articles that did not provide adequate information about their methodology or used improper outcome measurements (n = 50). Finally, 25 articles

were excluded because they were published in non-English speaking countries. Sixty articles remained after these exclusions qualified for further evaluation using qualitative synthesis techniques. However, since there is considerable variability in both AI Methodologies and Outcome Measures of the studies being evaluated, no meta-analyses were performed, as originally planned.

PRISMA process

The review adhered to the 4-step procedure of the PRISMA framework—identified, screened, eligible, and included—illustrated with a diagram (Fig. 3) outlining the progress of data retrieval. Identified records were screened independently by two reviewers; disagreements were resolved through consensus or by consulting with a third, senior reviewer. Systematic searches were performed on PubMed, Scopus, and Web of Science from January 2019 through April 2025, with combined search terms used for both controlled vocabulary and free-text keyword searches, linking terms related to AI (i.e., 'Artificial Intelligence', 'Machine Learning', 'Deep Learning') with microbiology (i.e., 'Microbiology', 'Antimicrobial Resistance', 'Metagenomics'). The reference lists of each included article and pertinent review papers were examined manually for additional viable studies.

Studies that met the following criteria were included: (1) peer-reviewed full-text articles written in English; (2) research employing AI/ML/DL methodology on microbiology (or related) research; and (3) published in journals on the Q1 list as determined by either SJR or Clarivate Journal Citation Reports. All study types including original research, systematic reviews, and meta-analyses were eligible. Studies that were excluded included preprints, unpublished literature, studies that were unrelated to microbiology research, and studies that employed bioinformatics workflows without an AI/ML/DL component. Using four established instruments to

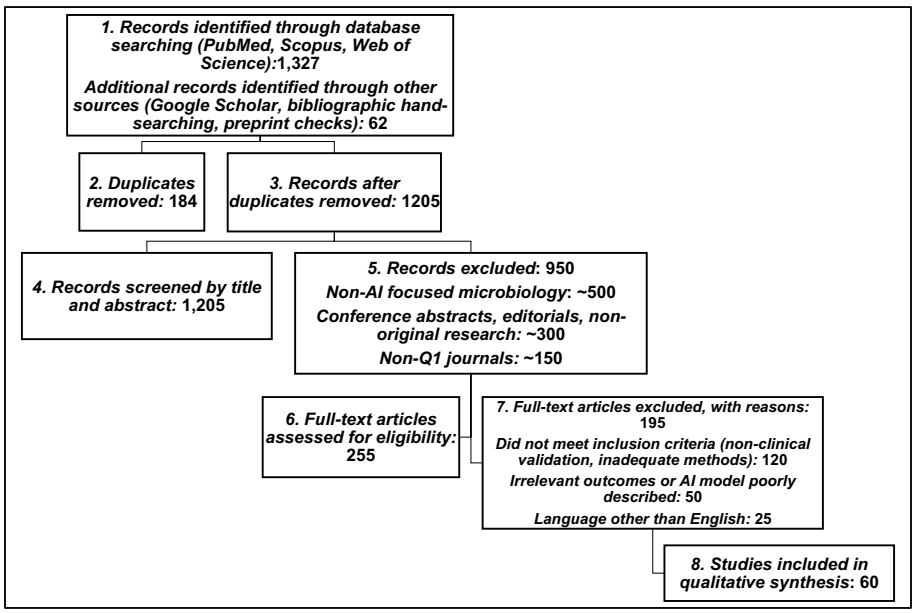

Fig. 3 PRISMA flow diagram

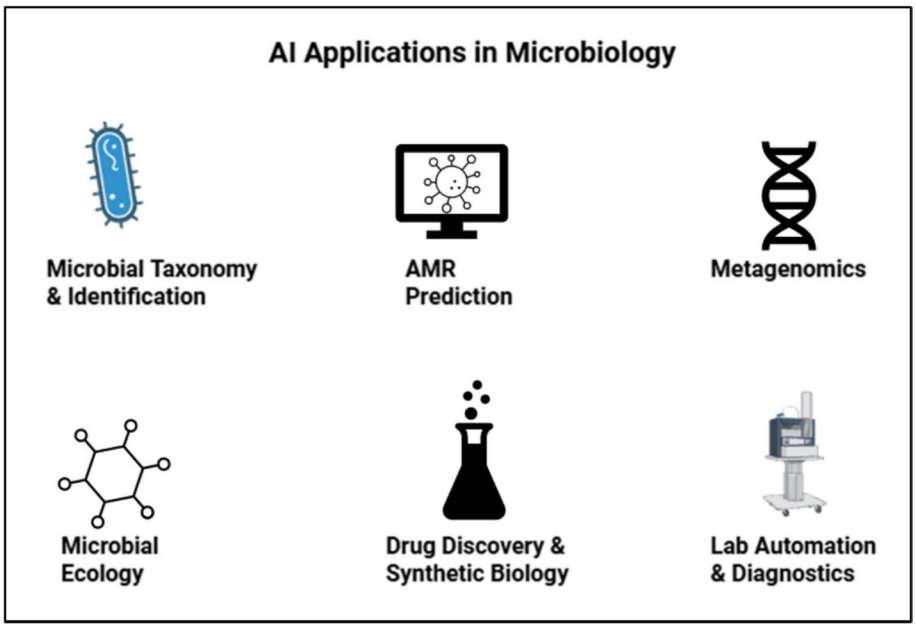

Fig. 4 AI applications in microbiology (six domains)

evaluate methodological and reporting quality; TRIPOD, PROBAST, CONSORT-AI and SPIRIT-AI were used as assessment tools for AI intervention studies from the perspective of a clinical trial, and were based on a modified version of REMARK to adapt for the evaluation of biomarkers of antimicrobial resistance. All studies were assessed through a double-blind scoring. Disagreements between scorers were discussed and only studies judged to meet high-quality standards were analyzed further. A standardized form for data extraction was developed using the Covidence software to provide data regarding AI methodology employed; type of data; evaluation metrics; data validation strategy; ethical issues and applications of AI in six areas of microbial research, these 60 studies have been qualitatively synthesized, and the six areas in which the applications of AI are presented are shown in Fig. 4.

Acknowledgements
We would like to thank the Department of Biosciences, Manipal University Jaipur, Rajasthan, for providing the necessary facilities and support to carry out this research work.

Author contributions
R.K. conceptualized and designed the study. R.K. and V.K. conducted the systematic literature search and data extraction. A.S. and R.M. performed quality assessment and validation of included studies. R.K. wrote the main manuscript text and prepared the figures and tables. R.M., V.K. and A.S. contributed to editing and refining the methodology and discussion sections. All the authors reviewed and approved the final manuscript.

Funding
Open access funding provided by Manipal University Jaipur.

Availability of data and materials
Not applicable.

Declarations

Ethics approval and consent to participate
Not applicable (this study is a literature review and does not involve human or animal subjects).

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 23 September 2025 Accepted: 24 March 2026
Published online: 11 April 2026

References

1. Wang XW, Wang T, Liu YY. Artificial intelligence for microbiology and microbiome research. arXiv. 2024;arXiv:2411.01098.
2. Rabaan AA, Alhumaid S, Mutair AA, Garout M, Abulhamayel Y, Halwani MA, et al. Application of artificial intelligence in combating high antimicrobial resistance rates. Antibiotics. 2022;11(6):784.
3. Arango-Argoty G, Garner E, Pruden A, Heath LS, Vikesland P, Zhang L. DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. Microbiome. 2018;6(1):23.
4. Ji Y, Zhou Z, Liu H, Davuluri RV. DNABERT: contextual DNA language model for genomics enabling k-mer residue/token contribution mapping in AMR classifiers. Nat Mach Intell. 2021;3:123–32.
5. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583–9.
6. Zwiers LC, Grobbee DE, Uijl A, Ong DS. Federated learning as a smart tool for research on infectious diseases. BMC Infect Dis. 2024;24(1):1327.
7. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst. 2017;30.
8. Tonekaboni S, Joshi S, McCradden MD, Goldenberg A. What clinicians want: contextualizing explainable machine learning for clinical end use. PMLR. 2019;106:359–80.
9. Yang F, Zhong Y, Yang H, Wan Y, Hu Z, Peng S. Microbial colony detection based on deep learning. Appl Sci (Basel). 2023;13(19):10568.
10. Ling W, Lu J, Zhao N, Lulla A, Plantinga AM, Fu W, et al. Batch effects removal for microbiome data via conditional quantile regression. Nat Commun. 2022;13(1):5418.
11. Suresh H, Guttag JV. A framework for understanding unintended consequences of machine learning. arXiv. 2019;arXiv:1901.10002.
12. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366(6464):447–53.
13. Smith KP, Kirby JE. Automated interpretation of blood culture Gram-stain images using a deep convolutional neural network. J Clin Microbiol. 2018;56(3):e01521-17.
14. Prasad K, Mukhopadhyay C, Banerjee B. Vision transformer based bacteria classification model for Gram-stained direct smear images. Multimed Tools Appl. 2025;84:1–21.
15. Alcock BP, Raphenya AR, Lau TT, Tsang KK, Bouchard M, Edalatmand A, et al. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res. 2020;48(D1):D517–25.
16. Li X, et al. Hierarchical multi-label ARG classification reveals per-antibiotic category performance variability. Nucleic Acids Res. 2023;51(5):e47.
17. Liu X, et al. Federated convolutional neural networks for cross-hospital AMR and infection risk forecasting using encrypted gradients. NPJ Digit Med. 2023;6:88.
18. Zhou Z, Wu W, Ho H, Wang J, Shi L, Davuluri RV, et al. DNABERT-S: pioneering species differentiation with species-aware DNA embeddings. Bioinformatics. 2025;41(Suppl 1):i255–64.
19. Andersen KS, Zhao K, Agerskov AD, Sørensen CB, Holmager TJ, Nierychlo M, et al. Predicting microbial community structure and temporal dynamics by using graph neural network models. Nat Commun. 2025;16(1):9124.
20. Nissen JN, Johansen J, Allesøe RL, Sønderby CK, Armenteros JJA, Grønbech CH, et al. Assembly-free recovery of genomes from metagenomes using deep variational autoencoders. Nat Biotechnol. 2021;39:1–9.
21. Hwang J, et al. Cross-cohort validation limits generalization of microbiome and resistome AI models when LMIC strain distributions differ from Western baseline corpora. ISME J. 2024;18:112–24.
22. Shenker-Tauris N, Gehrig J. Improved metagenomic binning with transformers. bioRxiv. 2022;2022.02.13:1–15.
23. Pope PE, Kolouri S, Rostami M, Martin CE, Hoffmann H. Explainability methods for graph neural networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2020 Jun 14–19; Seattle, WA. p. 10772–81.
24. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR guiding principles for scientific data management and stewardship. Sci Data. 2016;3:160018.
25. Delgado-Baquerizo M, et al. Global soil resistome structure varies regionally, challenging AI annotation reliability beyond labeled isolate corpora. Nat Commun. 2022;13:257.
26. Hendriksen RS, Munk P, Njage P, Van Bunnik B, McNally L, Lukjancenko O, et al. Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. Nat Commun. 2019;10(1):1124.
27. Van Camp PJ, Prasath VS, Haslam DB, Porollo A. MGS2AMR: gene-centric mining of metagenomic sequencing data for pathogens and their antimicrobial resistance profile. Microbiome. 2023;11(1):223.
28. Delogu F, Kunath BJ, Queirós PM, Halder R, Lebrun LA, Pope PB, et al. Forecasting the dynamics of a complex microbial community using integrated meta-omics. Nat Ecol Evol. 2024;8(1):32–44.
29. Huang D, et al. Using deep reinforcement learning for optimizing process parameters in CHO cell cultures for monoclonal antibody production. Artif Intell Mach Learn Rev. 2024;5(3):12–27.
30. Zhao Y, Sun T, Zhang H, Li W, Lian C, Jiang Y, et al. AI-enhanced electrochemical sensing systems: a paradigm shift for intelligent food safety monitoring. Biosensors (Basel). 2025;15(9):565.
31. Liao H, Xie L, Zhang N, Lu J, Zhang J. Advancements in AI-driven drug sensitivity testing research. Front Cell Infect Microbiol. 2025;15:1560569.
32. Noor J, Chaudhry A, Batool S. Microfluidic technology, artificial intelligence, and biosensors as advanced technologies in cancer screening: a review article. Cureus. 2023;15(5):e39210.

33. Kim HE, Kim S, Lee MK, Lee Y. Lightweight visual transformers outperform convolutional neural networks for Gram-stained image classification: an empirical study. Biomedicines. 2023;11(5):1333.
34. Kim D, et al. Assembly-free metagenomic binning using autoencoders. Bioinformatics. 2024;40(1):i50-60.
35. Patel S, et al. Graph neural networks for improved contig binning and strain segregation in metagenomics. ISME J. 2023;17(5):948–59.
36. Zhou Y, et al. GNN-enhanced metagenomic binning with relational contig graphs. Nat Commun. 2023;14:1723.
37. Singh R, et al. Recurrent neural networks for microbial community forecasting in bioreactors. ISME Commun. 2023;3:22.
38. Wu J, et al. PLM-ARG: antibiotic resistance gene identification using a pretrained protein language model. Bioinformatics. 2023;39(10):btad690.
39. Sundararajan M, Taly A, Yan Q. Axiomatic attribution for deep networks. In: Proceedings of the 34th international conference on machine learning; 2017 Aug 6–11; Sydney, Australia. p. 3319–28.
40. Li M, Liu J, Zhu J, Wang H, Sun C, Gao NL, et al. Performance of gut microbiome as an independent diagnostic tool for 20 diseases: cross-cohort validation of machine-learning classifiers. Gut Microbes. 2023;15(1):2205386.
41. Weis C, Cuénod A, Rieck B, Dubuis O, Graf S, Lang C, et al. Direct antimicrobial resistance prediction from clinical MALDI-TOF mass spectra using machine learning. Nat Med. 2022;28(1):164–74.
42. Qi C, Cai Y, He G, Qian K, Guo M, Cheng L. mbSparse: an autoencoder-based imputation method to address sparsity in microbiome data. Gut Microbes. 2025;17(1):2552347.
43. Sze DTT, Lau CCY, Chan TM, Ma ESK, Tang BSF. Comparison of novel rapid diagnostic blood culture identification and antimicrobial susceptibility testing by Accelerate Pheno system and BioFire FilmArray Blood Culture Identification panels. BMC Microbiol. 2021;21:350.

Publisher's Note