The Journal of Antibiotics (2021) 74:838–849
https://doi.org/10.1038/s41429-021-00471-w

JARA Japan Antibiotics Research Association
The Society for Actinomycetes Japan

REVIEW ARTICLE

A review: antimicrobial resistance data mining models and prediction methods study for pathogenic bacteria

Xinxing Li[1] · Ziyi Zhang[1] · Buwen Liang[1] · Fei Ye[2] · Weiwei Gong[2]

Received: 11 May 2021 / Revised: 27 May 2021 / Accepted: 16 July 2021 / Published online: 14 September 2021

Abstract

Antimicrobials have paved the way for medical and social development over the last century and are indispensable for treating infections in humans and animals. The dramatic spread and diversity of antibiotic-resistant pathogens have significantly reduced the efficacy of essentially all antibiotic classes and is a global problem affecting human and animal health. Antimicrobial resistance is influenced by complex factors such as resistance genes and dosing, which are highly nonlinear, time-lagged and multivariate coupled, and the amount of resistance data is large and redundant, making it difficult to predict and analyze. Based on machine learning methods and data mining techniques, this paper reviews (1) antimicrobial resistance data storage and analysis techniques, (2) antimicrobial resistance assessment methods and the associated risk assessment methods for antimicrobial resistance, and (3) antimicrobial resistance prediction methods. Finally, the current research results on antimicrobial resistance and the development trend are summarized to provide a systematic and comprehensive reference for the research on antimicrobial resistance.

Introduction

Background and motivation

Antimicrobials are effective in treating bacterial infectious diseases and are the cornerstone of modern medicine [1]. However, the unregulated use and misuse of antimicrobials in healthcare and animal farming, as well as the lack of new and novel antimicrobials [2–4], have led to an increasingly serious problem of antimicrobial resistance [5]. This poses a huge clinical and public health threat and seriously affect the life and health of animals and humans [6, 7]. Although the use of antimicrobials in livestock and poultry has contributed to health and productivity improvements, it has also contributed to the evolution of drug-resistant strains of bacteria [8, 9]. Many studies have shown that the use of antimicrobials in animals accounts for more than half of the total consumption and that resistance to antimicrobials of animal origin can be transmitted to humans directly or indirectly through the food chain [10]. It is expected that by 2030, more than 200,000 tons of antimicrobials will be used in animals worldwide [11], and by 2050, the number of people dying from drug-resistant infections is expected to reach 10 million annually worldwide and cost ~$100 trillion [12, 13].

Countries have developed policies to address drug resistance in the context of local environmental conditions, economic and demographic situations, and general trends in resistance rates [14], but the problem remains serious due to previous and current overuse and unregulated use. The main problems that limit drug resistance research are (1) the diversity of data collection [15] and the unavailability of some data, especially the lack of data from low- and middle-income countries [16], which will hinder inter-regional comparability; (2) existing studies mostly focus on biogenetics, lacking in-depth association analysis, which makes it difficult to predict the development trend of drug resistance under multiple factors [17]; (3) traditional methods have limitations in handling large-scale data and cannot reveal causal relationships from massive datasets.

With the increase in the number of resistant microorganisms and the shift from targeted identification by culture and/or using polymerase chain reactions to high-throughput

✉ Weiwei Gong
gongweiwei@rails.cn

[1] College of Information and Electrical Engineering, China Agricultural University, Beijing, PR China

[2] Transportation and Economics Research Institute, China Academy of Railway Sciences Corporation Limited, Beijing, PR China

metagenomics for detecting and identifying antimicrobial resistance, large-scale complex data are generated. Machine learning and data mining techniques [18–20] can well analyze the diverse and fragmented antimicrobial resistance datasets generated by genomic surveillance and reveal the mechanistic changes. Based on artificial intelligence technology [21], the existing massive data can be deeply mined and dynamically analyzed [22] to determine the overall antimicrobial resistance pattern, to explore new resistance evaluation and prediction problems from the macroscopic resistance level, and to lay the foundation for formulating corresponding policies and guiding drug use at the national or regional level.

This paper reviews the data storage and analysis techniques for antimicrobial resistance, antimicrobial resistance assessment methods and the associated risk assessment methods for antimicrobial resistance, and antimicrobial resistance prediction methods. It aims to explore the feasibility of machine learning methods and data mining techniques in solving the problem of antimicrobial resistance and to provide a systematic and comprehensive reference for the study of antimicrobial resistance.

Paper organization

This paper reviews the application of computer technology to various aspects of the drug resistance problem from two perspectives: macroscopic country-regional level and microscopic genetic level. This review is organized as follows:

Section 2 analyzes the methods for monitoring antimicrobial resistance, mainly summarizing the antimicrobial resistance gene database and monitoring tools over the past five years, and pointing out the characteristics and development trends of antimicrobial resistance.

Section 3 describes the existing antimicrobial resistance assessment methods and the associated risk assessment methods for antimicrobial resistance, focus on describing the application of machine learning methods in predicting antimicrobial resistance.

Section 4 discusses the current status of research on antimicrobial resistance, summarizes some of the problems that need to be addressed urgently, and suggests possible solutions.

Section 5 looks forward to the future trends and research directions to provide a systematic and comprehensive reference for the research of antimicrobial resistance.

Antimicrobial resistance monitoring

The increase in antimicrobial-resistant bacteria has led to an urgent need for rapid detection of resistance in clinical samples and improved global surveillance, and the availability of more accurate and comprehensive resistance surveillance data is a prerequisite for resistance assessment, particularly in low- and middle-income countries.

Macro-level antimicrobial resistance monitoring systems

With increased concern, governments and health organizations around the world have called for increased surveillance and coordinated action to combat antimicrobial resistance, and several countries around the world have now developed antimicrobial resistance surveillance systems and programs, as shown in Table 1.

In 1995, Denmark established the Danish Integrated Antibiotic Resistance Monitoring and Research Program (DANMAP) [23], the first system to monitor antibiotic resistance data in animals, food, and humans. In 1996, the National Antibiotic Resistance Monitoring System for Enteric Bacteria (NARMS) [24] was established in the U.S. In 1998, Europe established the European Antimicrobial Resistance Monitoring Network (EARS-Net) [25]. In 1999, Japan established the Japanese Veterinary Antibiotic Resistance Monitoring System (JVARM) [26] to monitor resistance to *E. coli* and *Salmonella* sp. in food animals (cattle, pigs, chickens). In 2003, Canada established the Canadian Integrated Program for Antimicrobial Resistance Surveillance (CIPARS) [27]. Germany also conducted the German national surveillance program (GERM-Vet) and its complementary program (BfT-GermVet) [28] to determine the antimicrobial susceptibility of bacterial pathogens in German. Nowadays, China Antimicrobial Resistance Surveillance System (CARSS) [29] has been established, which can help to obtain scientifically valuable and stratified bacterial resistance data and to grasp the national bacterial resistance epidemic trends.

To a certain extent, these national systems have realized the monitoring of key areas of antimicrobial drug use from the perspective of human medicine, veterinary medicine and food, obtained a large amount of data with research significance, and regularly published bacterial resistance trends and antibiotic use on the Internet, laying the data foundation for further monitoring.

Despite the recognition of the urgency of antimicrobial resistance, it is clear that there are general gaps in antimicrobial resistance surveillance in different countries and that existing national surveillance networks may be limited by low data volumes, bacterial handling, and geographic location. Besides, these existing networks are often geared toward high-level policy development and planning and may not be useful for regional surveillance or local decision making. To optimize antimicrobial surveillance in the short and long term, additional data sources and technologies are needed to achieve effective surveillance of antimicrobial

Table 1 Macro-level antimicrobial resistance monitoring stations

Monitoring stations	Function	URL	Year	References
DANMAP	Monitoring antimicrobial drug consumption and antimicrobial resistance in animals, food, and humans.	http://www.danmap.org	1995	[23]
NARMS	Data on antibiotic susceptibility in *Salmonella*, *Campylobacter*, *Escherichia coli* (*E. coli*), and *Enterococcus* sp. were collected.	https://www.cdc.gov/narms/index.html	1996	[24]
EARS-Net	Collect comparable, representative, and accurate antimicrobial resistance data for analysis of temporal and spatial trends in antimicrobial resistance in Europe to provide timely antimicrobial resistance data for decision making.	https://www.ecdc.europa.eu/en/about-us/partnerships-and-networks/disease-and-laboratory-networks/ears-net	1998	[25]
JVARM	Monitor drug resistance of *E. coli* and *Salmonella sp.* in food animals (cattle, pigs, chickens).	http://www.maff.go.jp/nval/tyosa_kenkyu/taiseiki/monitor/e_index.html	1999	[26]
CIPARS	Monitoring the use of antibiotics for pathogenic bacteria in food.	http://www.phac-aspc.gc.ca/cipars-picra	2003	[27]
BfT-GermVet	Determination of antibiotic susceptibility of bacterial pathogens in German animals.	http://vetline.de/17079309/150/3130/69483	2004	[28]
CARSS	Obtain scientifically valuable, stratified bacterial resistance data and grasp national trends in bacterial resistance prevalence.		2020	[29]

resistance data on a global scale to eliminate disparities between different countries and regions.

Antimicrobial resistance databases and tools

Antimicrobial resistance genes are often identified or predicted based on the best hits from sequence searches of existing databases. Accurate identification of correlations between antibiotic treatment outcomes and resistance determinants, and antimicrobial resistance gene profiles, will facilitate the development of targeted treatment regimens. The success of this approach depends heavily on the comprehensiveness and quality of public antimicrobial resistance gene databases and surveillance tools, which have an important role in the development of bioassays and computational tools. A relatively large number of genetic databases and tools have been developed worldwide, providing a large amount of relevant resistance gene data to support further analysis and prediction of drug resistance, as shown in Table 2.

Assessment and prediction of antimicrobial resistance

Antimicrobial resistance is a growing worldwide problem, but communicating this challenge to policymakers and non-experts is complicated by the diversity of bacterial pathogens and the diversity of antibiotics used to treat these pathogens. Even experts in antibiotic pharmacodynamics have difficulty inferring the severity of antimicrobial resistance if they lack information on the extent of antibiotic use and the effectiveness of antibiotic therapy.

Machine learning techniques discover patterns directly from the observation of the data, without the constraints of formulas or even human theories. By learning from the natural world, machine learning techniques can often have better accuracy than traditional methods [49]. The advent of large-scale datasets provided by next-generation sequencing and electronic health records has made it possible to apply machine learning to the study and treatment of antimicrobial resistance.

DRI

The Drug Resistance Index (DRI) was developed as a measure of antibiotic treatment efficacy that encapsulates the multiple relationships between pathogen susceptibility and antibiotic use into one easy-to-understand metric. The DRI represents the overall level of resistance and can be analyzed from both a temporal and spatial perspective.

Laxminarayan and Klugman constructed DRIs for two pathogens, *Acinetobacter* and *E. coli*, using data on the

Table 2 Antimicrobial resistance databases and surveillance tools

Database & Tools	Description	URL	References
DeepARG-DB	Covers ARG predictions with high confidence and extensive manual inspection.	http://bench.cs.vt.edu/deeparg	[30]
Bacterial Antimicrobial Resistance Reference Gene Database	Contains the hierarchy of AMR proteins, manually managed boundaries, and associated hierarchy names.	https://www.ncbi.nlm.nih.gov/pathogens/isolates#/refgene/	[31]
AMRFinder	Identification of antimicrobial resistance genes in whole genome sequences.	https://www.ncbi.nlm.nih.gov/pathogens/antimicrobial-resistance/AMRFinder/	
BLDB	Provides up-to-date structural and functional information on the enzyme superfamily.	http://bldb.eu/	[32]
u-CARE	Contains 52 reported resistant antibiotics, 107 genes, transcription factors, and single nucleotide polymorphisms (SNPs) involved in multiple drug resistance of this pathogen.	http://www.e-bioinformatics.net/ucare	[33]
MUBII-TB-DB MUBII	Analysis of gene sequences associated with antimicrobial resistance in tuberculosis.	https://umr5558-bibiserv.univ-lyon1.fr/mubii/mubii-select.cgi	[34]
DR mutation library	Predictive library of *Mycobacterium tuberculosis* resistance to 15 anti-tuberculosis drugs.		[35]
PATRIC	Provides information on antimicrobial resistance at the genomic and genetic levels.	https://www.patricbrc.org/	[36]
ARDB	Provide information on antimicrobial resistance	http://ardb.cbcb.umd.edu/	[37]
MEGARes	Analysis of large-scale ecological sequence datasets	https://megares.meglab.org	[38]
ARG-ANNOT	Detection of existing and hypothetical new antimicrobial resistance genes in the bacterial genome.		[39]
CARD	Provides reference DNA and protein sequences, assay models on the molecular basis of bacterial antimicrobial resistance.	https://card.mcmaster.ca	[40]
ResFinder	Identification of antimicrobial resistance genes obtained from fully sequenced isolates.	https://cge.cbs.dtu.dk/services/ResFinder/	[41]
PointFinder	Detection of chromosomal point mutations associated with antimicrobial resistance.	http://www.genomicepidemiology.org/	
ARP	Identification of new antibiotic adjuvants and removal of known naturally occurring antibiotic scaffolds.		[42]
ResistanceOpen	Aggregate antimicrobial resistance index.	https://resistanceopen.org/	[43]
ResCap	The first targeted sequence capture, specifically developed to analyze resistomes.		[44]
SEAR	Detects antimicrobial resistance genes obtained horizontally from raw sequencing data, providing gene information, abundance estimates, and reconstructed sequences of antimicrobial resistance genes, and also provides web links to additional information for each gene.	http://computing.bio.cam.ac.uk/sear/SEAR_WEB_PAGE/SEAR.html	[45]
VRprofile	Detection of virulence and antimicrobial resistance genes in newly sequenced pathogenic bacterial genomes.	http://bioinfo-mml.sjtu.edu.cn/VRprofile	[46]
CASTB	Epidemiological analysis and drug resistance prediction.	http://castb.ri.ncgm.go.jp/CASTB	[47]
ARIBA	Identification of AMR-related genes from short read sequences.	https://github.com/sanger-pathogens/ariba	[48]

proportion of resistant isolates tested in the United States and antimicrobial consumption to fit trends in antimicrobial resistance and antimicrobial consumption in the United States [50]. Klein et al. calculated the DRIs for 41 countries for which antimicrobial use and resistance data were available and estimated the combined resistance rate of pathogenic bacteria in these countries [51]. The calculation of DRIs in both of these papers was based on Eq. (1).

$$DRI = \sum_k \rho_k^t q_k^t \tag{1}$$

Where ρ_k^t is the proportion of resistance among all included pathogens to drug k for time t, and q_k^t is the proportion of drug k used for treatment of those pathogens in all drugs included in the index for time t.

Chitanand et al. indexed multidrug resistance based on bacteriological analysis to determine the prevalence of multiple antimicrobial resistance in isolates from different river basins [52]. Multiple Antibiotic Resistance (MAR Index) Index of the samples was calculated by the Eq. (2).

$$MAR\ Index = \frac{y}{nx} \tag{2}$$

Where, y is the total number of resistance scored, n is the number of isolates, x is the total number of antibiotics tested.

Ciccolini et al. analyzed the use weights and frequencies of nearly 20 antibiotics for the treatment of *E. coli* and *Klebsiella sp*. and based on an exponential method combining three simple aggregated metrics (relative frequency of etiological agents, level of resistance and relative frequency of antibiotic use), established a DRI by which to assess the appropriateness of antibiotics for patients [53].

$$DRI_i(D, C) = \sum_{j=1}^{20} P_j(D, C) R_{i,j} \tag{3}$$

Where the sum is over all antibiotic regimens, $P_j(D,C)$ is the relative frequency of prescription of antibiotic regimen j among patients in diagnostic group D that were prescribed a C-compliant therapy, and $R_{i,j}$ is the prevalence of resistance to antibiotic regimen j among bacterial species i.

Chen et al. established DRIs to prioritize hospital drug use by studying the rate of microbial resistance to antibiotics and evaluating the effectiveness of treating major pathogenic bacteria [54]. Hughes et al. developed two complementary indices, the Empirical Coverage Index and the Empirical Choice Index, to summarize the clinical impact of antimicrobial resistance and drug availability on empiric therapy, and validated the rationale for the indices in the intensive care unit of a Toronto hospital [55].

However, antimicrobial resistance is a complex system that is susceptible to multiple factors. Using only the frequency of antimicrobial use or resistance rate to establish DRI is not very reliable, and integrating multi-factors is an important criterion to make DRI scientific. Machine learning solves this problem by performing a series of feature engineering processes on the raw data, from data preprocessing to feature scaling, feature coding, feature selection, and feature extraction, to remove impurities and redundancies from the raw data and distill it into features for use by algorithms and models.

Li et al. established a DRI for anti-*E. coli* drug risk status based on principal component analysis, combining AMR, antimicrobial use data, and environmental factors (water, soil) to reveal the effects of antimicrobial use, contamination level, and bacterial AMR, bridging the gap between judging AMR by resistance rate alone [56]. Using three sources of data (ResistanceMap, the WHO 2014 report on antimicrobial resistance, and contemporary publications), Collignon et al. created two global indices of antimicrobial resistance for 103 countries using data from 2008 to 2014: *E. coli* resistance and aggregated resistance, and found no significant association between antimicrobial consumption and higher antimicrobial resistance through logistic regression models [57].

DRI combines various antimicrobials consumption and the proportion of resistance of different pathogens in a single indicator. Although several methods have been proposed to calculate DRI, there is still no unified standard to measure the relationship between antimicrobial use and pathogen resistance, and the methods proposed above do not have sufficient scientific validity for the selection of characteristics in the calculation of DRI [58]. Feature engineering can select a subset of features containing all the important information to some extent [59, 60], but may suffer a combinatorial explosion as the number of data increases. Better results may be produced if an end-to-end network [61, 62] of deep learning is used, incorporating this step of feature extraction into the algorithm and allowing the algorithm to learn by themselves without human intervention.

Antimicrobial resistance risk assessment

There are currently three main methods for antimicrobial resistance risk assessment [63]: qualitative risk assessment, semi-quantitative risk assessment, and quantitative risk assessment. Qualitative risk assessment uses descriptive terms of different levels to indicate the degree of risk; semi-quantitative risk assessment is similar to qualitative risk assessment in that each model parameter of the risk assessment pathway needs to be scored and then integrated; and quantitative risk assessment parameters and variables are expressed as quantitative values, with the final result indicating the magnitude of risk and the uncertainty in the assessment process.

Qualitative risk assessment is gradually being discarded due to its unscientific nature. Collineau et al. developed a semi-quantitative risk assessment model [64], using data from the Swiss AMR monitoring program, 208 combinations of animal species/bacteria/antimicrobial classes were identified as relevant hazards.

Quantitative microbial risk assessment is currently the dominant method for evaluating human health risks associated with pathogenic microorganisms. Collineau et al. developed a quantitative microbial risk assessment model to assess the extent to which *Salmonella Heidelberg*, which is resistant to ceftiofur, in broiler *chicken* products poses a risk to public health in Canada [65]. Hoa et al. evaluated occurrence of antimicrobial-resistant bacteria in the VAC environment in northern Vietnam, with quantitative analysis of antibiotic pollution [66]. Beaudequin et al. proposed a quantitative microbial risk assessment in the context of wastewater reuse represented as a Bayesian network (BN), which allows a comprehensive visualization of the main influencing factors in the exposure pathway [67]. Pouillot et al. performed a quantitative risk assessment of the risk of waterborne *Cryptosporidium* sp. infection and Cryptosporidiosis in immunocompromised and immunocompromised French populations based on second-order Monte Carlo simulations [68].

In addition, Monaco et al. used SMRT sequencing to assess the profile of resistance to HIV quasispecies in 38 HIV-positive patients with treatment failure and found a complex resistance pattern in the quasispecies with treatment failure [69]. Hendriksen et al. used metagenomic analysis of untreated sewage to characterize the bacterial resistome from 79 sites in 60 countries, and found that global AMR gene diversity and abundance vary by region [70]. Okada et al. developed a modified computational molecular dynamics simulation (MP-CAFEE) to evaluate drug resistance due to ALK compound mutations [71]. Cassini et al. modeled disease outcomes for five infection types to estimate the incidence of infection for 16 combinations of antimicrobial-resistant bacteria from data from the European Antimicrobial Resistance Surveillance Network (EARS-Net). The results showed that all age groups were affected by antimicrobial-resistant bacterial infections [72].

Prediction of antimicrobial resistance

Whole genome sequencing technology can be used to analyze antimicrobial resistance genes and their mutations. In addition to analyzing known resistance mechanisms, it can also predict potential resistance mechanisms, such as predicting antimicrobial resistance in bacteria [73, 74], and predicting antimicrobial susceptibility [75, 76]. Compared to polymerase chain reaction, whole genome sequencing can detect more genes and predict drug resistance phenotypes more effectively and accurately [77].

While the operational processes of WGS are relatively undemanding, the management and analysis of these large datasets require specialized expertize and software tools. Machine learning has more prominent advantages for handling large-scale data sets, and many studies have explored machine learning algorithms for antimicrobial resistance, highlighting its role in predicting resistance phenotypes based on the whole-genome sequence [78].

Using 25 whole genome sequences, Rishishwar et al. established an *Staphylococcus aureus* (*S. aureus)* discriminator based on 14 genetic parameters that can distinguish vancomycin-intermediate *S. aureus* from vancomycin-susceptible *S. aureus* based on specific genomic markers of antimicrobial susceptibility [79]. Davis et al. custom built AdaBoost (adaptive boosting) machine learning classifiers for identifying carbapenem resistance in *Acinetobacter baumannii*, methicillin resistance in *S. aureus*, and beta-lactam and co-trimoxazole resistance in *Streptococcus pneumoniae* with accuracies ranging from 0.88 to 0.99 based on bacterial genomes with AMR metadata collected in the PATRIC database [80]. Her and Wu used genetic algorithms (GA) to find the best subset of auxiliary gene clusters and CARD annotations (acc/CARD) to predict drug-resistant or sensitive strains based on *E. coli* genomic and AMR metadata from the PATRIC database, and evaluated model performance using four machine learning methods (SVM, RF, Adaboost, NB), respectively, and found that predicting AMR activity from genomic information may be a nonlinear problem [81]. Valizadehaslani et al. developed an XGBoost regression model based on the genomic sequence of bacteria to predict the minimum antimicrobial dose required to treat bacterial infections and to identify the content of specific mutations or alterations in genes that cause drug resistance. [82]. Maguire et al. sequenced and assembled 97 genomes isolated from broiler chickens on farms in British Columbia between 2005 and 2008 and developed logistic regression models to predict AMR phenotypes with a prediction accuracy of 0.92–0.99 [83].

And machine learning algorithms can learn antimicrobial resistance mechanisms from DNA sequences without any prior information. Using machine learning to solve the AMR problem is still in its infancy, but promising progress has been made [84].

Lakin et al. proposed a Meta-MARC machine learning classifier that uses a hierarchical, DNA-based hidden Markov model to accurately and diversely detect sequences that differ from known antimicrobial resistance genes [85]. Khaledi et al. sequenced the genomes and transcriptomes of 414 clinical *P. aeruginosa* isolates, generated predictive models based on a support vector machine classifier with linear kernel functions, and identified biomarkers of

resistance to four commonly used antimicrobial drugs [86]. Kulshrestha et al. constructed a decision tree classifier to identify resistance patterns based on results from patients who underwent antimicrobial susceptibility testing and used it to predict resistance to various antimicrobials [87]. Martínez-Agüero et al. used five methods (logistic regression, k-nearest neighbor, decision tree, random forest, and multilayer perceptron) to determine the susceptibility of bacteria to certain antimicrobials and trends in antimicrobial resistance based on clinical and demographic characteristics from patients and data from cultures and antimicrobial maps [88]. Nguyen et al. developed an XGBoost-based machine learning model [89] to predict MICs for 15 antimicrobials and explored the genomic regions identified by the model for predicting MICs using feature selection. Liu et al. used support vector machine (SVM) and set covering machine (SCM) models to learn and predict drug resistance for five drugs (Tetracycline, Ampicillin, Sulfisoxazole, Trimethoprim, and Enrofloxacin), and the training accuracy and testing accuracy of the SVM and SCM models for the five drugs were above 90% [90].

Traditional machine learning techniques are feasible in terms of principles and results, but have limitations and are not robust in dealing with raw data. Deep learning methods are representation learning methods with multiple levels of representation, where one level of representation (starting from the original input) is transformed into a higher, slightly more abstract level of representation through non-linear modules respectively [91]. The key aspect of deep learning is that these feature layers are not designed by human engineers; they are learned from the data using a learning procedure.

Deep learning has proven to be the most powerful machine learning approach to date for many applications particularly in the case of predicting DNA sequence affinities, the deep learning model surpasses all known binding site prediction approaches [92]. Stokes et al. trained a deep neural network capable of predicting molecules with antimicrobial activity [93]. Arango-Argoty et al. applied deep learning to identify antimicrobial resistance genes and constructed two deep learning models DeepARG-SS and DeepARG-LS for short read sequences and full gene length sequences to predict antimicrobial resistance genes with higher precision and recall relative to next-generation sequencing methods.

Besides, Ruppé et al. developed and validated an annotation method (called pairwise comparative modeling) on the basis of a three-dimensional structure (homology comparative modeling), leading to the prediction of 6095 antimicrobial resistance determinants in a catalog of 3.9 million proteins from the human intestinal microbiota [94]. Duranti et al. predicted putative mobile elements such as episomes, conjugative transposons, and prophages through homology searches against in-house-generated databases, including genes retrieved from the National Center for Biotechnology Information database [95]. Collignon et al. investigated the relationship between AMR and its potential influencing factors and predicted *E. coli* resistance to several different drugs based on univariate and multivariate analyses.

A summary of the above methods is shown in Table 3.

Discussions and conclusions

Antimicrobial resistance is a major public health threat. Monitoring and understanding the prevalence, mechanisms, and spread of antimicrobial resistance are priorities for personal care and global infection control strategies. We primarily review the research addressing antimicrobial resistance over the past five years, including antimicrobial resistance databases and surveillance tools, antimicrobial assessment methods, and antimicrobial prediction methods. A summary of the literature analysis revealed that there is insufficient research on drug resistance, mainly in the following areas.

[1] A large amount of resistance-related data is lacking. Although some progress has been made in constructing comprehensive antimicrobial resistance gene databases, the diversity of data collection, lack of uniformity across databases, and long update intervals hinder the realization of their potential.

[2] Surveillance of bacterial diseases, drug use, and resistance in livestock remain incomplete. Although a large number of drug-resistant bacteria and corresponding genes have been identified in various animal foods, the trend of antimicrobial resistance in animals has not been taken seriously, and achieving better antimicrobial stewardship on farms remains a challenge.

[3] There is a lack of fully functional models to predict which resistance genes will spread at local levels in the healthcare setting and globally between countries. These models may need to incorporate not only antimicrobial resistance gene sequences and mechanisms but also genomic context, host bacterial species, and geographic location.

[4] Antimicrobial resistance is spreading much faster than new compounds are entering clinical practice, and this has created a public health crisis. Although novel antimicrobial resistance gene discovery techniques such as functional metagenomics exist, there are still significant limitations to the types of antimicrobial resistance genes that can be detected by these techniques, and innovative approaches that can identify other antimicrobial resistance gene mechanisms are urgently needed.

We suggest encouraging research on antimicrobial resistance at the national level and promoting microscopic pathogenic experiments in local laboratories to fill the data gap. At the same time, local surveillance of antimicrobial resistance should be strengthened and data collection should be

Table 3 Prediction methods of antimicrobial resistance

Models & Methods	Strains	Result	References
Machine learning	*Staphylococcus aureus*	The model is able to discriminate based on specific genomic markers of antibiotic susceptibility rather than overall sequence relatedness.	[79]
AdaBoost (adaptive boosting) machine learning classifiers	*Acinetobacter baumannii* *Staphylococcus aureus* *Streptococcus pneumoniae*	Identifying carbapenem resistance in *Acinetobacter baumannii*, methicillin resistance in *Staphylococcus aureus*, and beta-lactam and co-trimoxazole resistance in *Streptococcus pneumoniae* with accuracies ranging from 88–99%.	[80]
Machine learning	*E. coli*	Accurate prediction of drug resistance activity of *E. coli* strains and identification of key factors for drug resistance.	[81]
Machine learning	*Campylobacter jejuni* *Neisseria gonorrhoeae* *Klebsiella pneumoniae* *Salmonella enterica*	Their proposed method produces comparable or even more accurate results when compared to existing methods for the goal of dose prediction, and it can provide additional insight for scientists who study AMR mechanisms.	[82]
Machine learning	*Nontyphoidal Salmonella*	Model accuracy of 0.92 to 0.99.	[83]
Extreme gradient boosting (XGBoost)-based machine learning models	*Nontyphoidal Salmonella*	The MIC prediction models had an overall average accuracy of 95% within 1 2-fold dilution step (confidence interval, 95% to 95%), an average very major error rate of 2.7% (confidence interval, 2.4–3.0%), and an average major error rate of 0.1% (confidence interval, 0.1–0.2%).	[89]
Support vector machine (SVM) and set covering machine (SCM)	*Actinobacillus Pleuropneumoniae*	Training accuracy (mean cross-validation score) and test accuracy above 90% for all 5 drugs.	[90]
Deep Learning		Prediction of antimicrobial resistance genes achieves high precision (>0.97) and recall (>0.90).	[30]
Deep neural network	*Mycobacterium tuberculosis* *carbapenem-resistant Enterobacteriaceae*	They discovered a molecule from the Drug Repurposing Hub halicin that displays bactericidal activity against a wide phylogenetic spectrum of pathogens including Mycobacterium tuberculosis and carbapenem-resistant Enterobacteriaceae.	[93]
Pairwise comparative modeling	*Human Intestinal Microbiota*	Prediction of 6,095 antimicrobial resistance determinants (ARDs) in a catalog of 3.9 million proteins from the human intestinal microbiota.	[94]
Molecular diagnostics	*Pseudomonas aeruginosa*	Sensitivity and prediction accuracy of 0.8 and 0.9 can be achieved.	[86]
Univariate and multivariate analyses	*E. coli*	Examining the association between antimicrobial resistance and potential contributing factors.	[57]

standardized. In the data modeling process, based on feature engineering and deep learning end-to-end networks, converting data attributes into data features can be free from noise and more accurately uncover potential trends in the data.

Prospect

The spread of drug resistance continues to accelerate, and we need to assess the extent of the impact of this problem on health and economic outcomes, including assessing the impact of resistance at the societal level, as well as developing methods to assess the severity of disease in patients with infectious diseases and implementing measures that can improve outcomes for patients with drug-resistant infections. These include techniques that emphasize the prevention of the emergence and spread of drug resistance through the judicious use of antimicrobials, appropriate infection control measures, strategies to minimize delays in the use of antimicrobial therapy, and the promotion of early detection of drug-resistant pathogens.

Effective action against AMR requires a coordinated response from government, industry, and international agencies and scientists, as well as the involvement of clinicians, pharmacists, patients, veterinarians, and farmers. A limitation of the current study is that the bioinformatics aspects of AMR surveillance have not been investigated on large biological or pharmacological datasets, as few studies have been conducted on this topic. In the future, we recommend an overview of antimicrobial resistance surveillance systems based on bioinformatics such as genetic, chemical, and biological behavior or structure and macroscopic resistance rates and resistance profile characteristics to develop strategies to improve antimicrobial stewardship, discover sources of antimicrobial drugs, and develop effective and sustainable alternatives to address microbial diseases in humans and livestock.

Acknowledgements The authors would like to acknowledge the financial supports provided by the National Natural Science Foundation of China (No. 61802411).

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Hutchings MI, Truman AW, Wilkinson B. Editorial overview: antimicrobials: tackling AMR in the 21st century. Curr Opin Microbiol [Internet]. 2019;51:iii–v. https://doi.org/10.1016/j.mib.2019.11.004.
2. Duong A, Jaelin M. 6 factors that have caused antibiotic resistance the number of new antibiotics being identified has slumped to an all time low. InfectionControl.tips. 2015;1–2.
3. Laxminarayan R, Matsoso P, Pant S, Brower C, Røttingen JA, Klugman K, et al. Access to effective antimicrobials: a worldwide challenge. Lancet 2016;387:168–75.
4. Woolhouse M, Ward M, Van Bunnik B, Farrar J. Antimicrobial resistance in humans, livestock and the wider environment. Philos Trans R Soc B Biol Sci. 2015;370:1–7.
5. Walther C, Rossano A, Thomann A, Perreten V. Antibiotic resistance in Lactococcus species from bovine milk: presence of a mutated multidrug transporter mdt(A) gene in susceptible Lactococcus garvieae strains. Vet Microbiol. 2008;131:348–57.
6. Smith R, Coast J. The true cost of antimicrobial resistance. BMJ. 2013;346:1–5.
7. Stewardson AJ, Allignol A, Beyersmann J, Graves N, Schumacher M, Meyer R, et al. The health and economic burden of bloodstream infections caused by antimicrobial-susceptible and non-susceptible Enterobacteriaceae and Staphylococcus aureus in European hospitals, 2010 and 2011: A multicentre retrospective cohort study. Eurosurveillance. 2016;21:5–16.
8. Liu YY, Wang Y, Walsh TR, Yi LX, Zhang R, Spencer J, et al. Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. Lancet Infect Dis [Internet]. 2016;16:161–8. https://doi.org/10.1016/S1473-3099(15)00424-7.
9. Coetzee J, Corcoran C, Prentice E, Moodley M, Mendelson M, Poirel L, et al. Emergence of plasmid-mediated colistin resistance (MCR-1) among Escherichia coli isolated from South African patients. South Afr Med J. 2016;106:449–50.
10. Sharma C, Rokana N, Chandra M, Singh BP, Gulhane RD, Gill JPS, et al. Antimicrobial resistance: its surveillance, impact, and alternative management strategies in dairy animals. Front Vet Sci. 2018;4:1–27.
11. Van Boeckel TP, Brower C, Gilbert M, Grenfell BT, Levin SA, Robinson TP, et al. Global trends in antimicrobial use in food animals. Proc Natl Acad Sci USA. 2015;112:5649–54.
12. Neill JO'. Antimicrobial resistance: tackling a crisis for the health and wealth of nations The Review on Antimicrobial Resistance Chaired by Neill JO'. 2014;(December).
13. Brogan DM, Mossialos E. A critical analysis of the review on antimicrobial resistance report and the infectious disease financing facility. Glob Health [Internet]. 2016;12:1–7. https://doi.org/10.1186/s12992-016-0147-y.
14. Lulijwa R, Rupia EJ, Alfaro AC. Antibiotic use in aquaculture, policies and regulation, health and environmental risks: a review of the top 15 major producers. Rev Aquac. 2020;12:640–63.
15. Wernli D, Jørgensen PS, Harbarth S, Carroll SP, Laxminarayan R, Levrat N, et al. Antimicrobial resistance: the complex challenge of measurement to inform policy and the public. PLoS Med. 2017;14:1–9.
16. de Kraker MEA, Stewardson AJ, Harbarth S. Will 10 million people die a year due to antimicrobial resistance by 2050? PLoS Med. 2016;13:1–6.
17. Mack SG, Turner RL, Dwyer DJ. Achieving a predictive understanding of antimicrobial stress physiology through systems biology. Trends Microbiol [Internet]. 2018;26:296–312. https://doi.org/10.1016/j.tim.2018.02.004.
18. Wainberg M, Merico D, Delong A, Frey BJ. Deep learning in biomedicine. Nat Biotechnol 2018;36:829–38.
19. Camacho DM, Collins KM, Powers RK, Costello JC, Collins JJ. Next-generation machine learning for biological networks. Cell [Internet]. 2018;173:1581–92. https://doi.org/10.1016/j.cell.2018.05.015.

20. Yang JH, Wright SN, Hamblin M, McCloskey D, Alcantar MA, Schrübbers L, et al. A white-box machine learning approach for revealing antibiotic mechanisms of action. Cell. 2019;177:1649–.e9.
21. Yu K-H, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng [Internet]. 2018;2:719–31. https://doi.org/10.1038/s41551-018-0305-z.
22. Webb S. Deep learning for biology. Nature. 2018;554:555–7.
23. Hammerum AM, Heuer OE, Emborg H, Bagger-skjøt L, Jensen VF, Rogues A, et al. Danish integrated antimicrobial resistance monitoring and research program. Emerg Infect Dis. 2007;13:1632–9.
24. Gilbert JM, White DG, McDermott PF. The US National Antimicrobial Resistance Monitoring System. Future Microbiol [Internet]. 2007;2:493–500. https://doi.org/10.2217/17460913.2.5.493.
25. Aracil-García B, Oteo-Iglesias J, Cuevas-Lobato Ó, Lara-Fuella N, Pérez-Grajera I, Fernández-Romero S, et al. Rapid increase in resistance to third generation cephalosporins, imipenem and co-resistance in Klebsiella pneumoniae from isolated from 7140 blood-cultures (2010–2014) using EARS-Net data in Spain. Enfermedades Infecc Microbiol Clin (Engl ed) [Internet]. 2017;35:478–84. https://doi.org/10.1016/j.eimce.2017.08.007.
26. Okura M, Sato M, Noda K. Overview of national antimicrobial resistance monitoring system in Europe and the United States. J Vet Epidemiol. 2012;16:152–6.
27. Conly JM. Antimicrobial resistance programs in Canada 1995–2010: a critical evaluation. Antimicrob Resist Infect Control. 2012;1:1–5.
28. Schwarz S, Alesik E, GAarestrup FM, Luebke-Becker A, Wallmann J, Werckenthin C, et al. The BfT-GermVet monitoring program—aims and basics. Berl Munch Tierarzt Wochenschr. 2007;120:357–62.
29. System AC of CARS. Technical programme of China antimicrobial resistance surveillance system, 2020 edition. Chin J Infect Chemother [Internet]. 2020;20:560–4. http://kns.cnki.net/KCMS/detail/detail.aspx?FileName=KGHL202005025&DbName=CJFQTEMP.
30. 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;6:1–15.
31. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, et al. Validating the AMRFINder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. Antimicrob Agents Chemother. 2019;63:1–20.
32. Naas T, Oueslati S, Bonnin RA, Dabos ML, Zavala A, Dortet L, et al. Beta-lactamase database (BLDB)–structure and function. J Enzym Inhib Med Chem [Internet]. 2017;32:917–9. https://doi.org/10.1080/14756366.2017.1344235.
33. Saha SB, Uttam V, Verma V. u-CARE: user-friendly comprehensive antibiotic resistance repository of Escherichia coli. J Clin Pathol. 2015;68:648–51.
34. Flandrois JP, Lina G, Dumitrescu O. MUBII-TB-DB: a database of mutations associated with antibiotic resistance in Mycobacterium tuberculosis. BMC Bioinform. 2014;15:1–9.
35. Coll F, McNerney R, Preston MD, Guerra-Assunção JA, Warry A, Hill-Cawthorne G. et al. Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences. Genome Med [Internet]. 2015;7:1–10.
36. Antonopoulos DA, Assaf R, Aziz RK, Brettin T, Bun C, Conrad N, et al. PATRIC as a unique resource for studying antimicrobial resistance. Brief Bioinform. 2018;20:1094–102.
37. Liu B, Pop M. ARDB—antibiotic resistance genes database. Nucleic Acids Res. 2009;37:443–7.
38. Lakin SM, Dean C, Noyes NR, Dettenwanger A, Ross AS, Doster E, et al. MEGARes: an antimicrobial resistance database for high throughput sequencing. Nucleic Acids Res. 2017;45:D574–80.
39. Kumar GS, Roshan PB, M DS, Rafael L-R, Marie K, Luce L, et al. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Chemother. 2014;58:212–20.
40. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, et al. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res. 2020;48:D517–25.
41. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother. 2012;67:2640–4.
42. Cox G, Sieron A, King AM, De Pascale G, Pawlowski AC, Koteva K, et al. A common platform for antibiotic dereplication and adjuvant discovery. Cell Chem Biol [Internet]. 2017;24:98–109. https://doi.org/10.1016/j.chembiol.2016.11.011.
43. Macfadden DR, Fisman D, Andre J, Ara Y, Majumder MS, Bogoch II, et al. A platform for monitoring regional antimicrobial resistance, using online data sources: ResistanceOpen. J Infect Dis. 2016;214:S393–8.
44. Lanza VF, Baquero F, Martínez JL, Ramos-Ruíz R, González-Zorn B, Andremont A, et al. In-depth resistome analysis by targeted metagenomics. Microbiome. 2018;6:1–14.
45. Rowe W, Baker KS, Verner-Jeffreys D, Baker-Austin C, Ryan JJ, Maskell D, et al. Search engine for antimicrobial resistance: A cloud compatible pipeline and web interface for rapidly detecting antimicrobial resistance genes directly from sequence data. PLoS One. 2015;10:1–12.
46. Li J, Tai C, Deng Z, Zhong W, He Y, Ou HY. VRprofile: gene-cluster-detection-based profiling of virulence and antibiotic resistance traits encoded within genome sequences of pathogenic bacteria. Brief Bioinform. 2018;19:566–74.
47. Iwai H, Kato-Miyazawa M, Kirikae T, Miyoshi-Akiyama T. CASTB (the comprehensive analysis server for the Mycobacterium tuberculosis complex): a publicly accessible web server for epidemiological analyses, drug-resistance prediction and phylogenetic comparison of clinical isolates. Tuberculosis [Internet]. 2015;95:843–4. https://doi.org/10.1016/j.tube.2015.09.002.
48. Hunt M, Mather AE, Sánchez-Busó L, Page AJ, Parkhill J, Keane JA, et al. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. Microb Genom. 2017;3:1–11.
49. Durrant JD, Amaro RE. Machine-learning techniques applied to antibacterial drug discovery. Chem Biol Drug Des. 2015;85:14–21.
50. Laxminarayan R, Klugman KP. Communicating trends in resistance using a drug resistance index. BMJ Open. 2011;1:1–8.
51. Klein EY, Tseng KK, Pant S, Laxminarayan R. Tracking global trends in the effectiveness of antibiotic therapy using the Drug Resistance Index. BMJ Glob Heal. 2019;4:1–7.
52. Chitanand MP, Kadam TA, Gyananath G, Totewad ND, Balhal DK. Multiple antibiotic resistance indexing of coliforms to identify high risk contamination sites in aquatic environment. Indian J Microbiol. 2010;50:216–20.
53. Ciccolini M, Spoorenberg V, Geerlings SE, Prins JM, Grundmann H. Using an index-based approach to assess the population-level appropriateness of empirical antibiotic therapy. J Antimicrob Chemother. 2015;70:286–93.
54. Chen Y, Song B, Shan X, Qin Y, Wang L, Wang H, et al. Assessing antibiotic therapy effectiveness against the major bacterial pathogens in a hospital using an integrated index. Future Microbiol. 2017;12:853–66.
55. Hughes JS, Hurford A, Finley RL, Patrick DM, Wu J, Morris AM. How to measure the impacts of antibiotic resistance and antibiotic development on empiric therapy: New composite indices. BMJ Open. 2016;6:1–12.

56. Li X, Liang B, Xu D, Wu C, Li J, Zheng Y. Antimicrobial resistance risk assessment models and database system for animal-derived pathogens. Antibiotics. 2020;9:1–16.
57. Collignon P, Beggs JJ, Walsh TR, Gandra S, Laxminarayan R. Anthropological and socioeconomic factors contributing to global antimicrobial resistance: a univariate and multivariable analysis. Lancet Planet Heal. 2018;2:e398–405.
58. Vandenbroucke-Grauls CMJE, Kahlmeter G, Kluytmans J, Kluytmans-Van Den Bergh M, Monnet DL, Simonsen GS, et al. The proposed Drug Resistance Index (DRI) is not a good measure of antibiotic effectiveness in relation to drug resistance. BMJ Glob Heal. 2019;4:1–3.
59. Effrosynidis D, Tsikliras A, Arampatzis A, Georgios S. Species distribution modelling via feature engineering and machine learning for pelagic fishes in the Mediterranean Sea. Appl Sci. 2020;10:1–23.
60. Kennedy A, Nash G, Rattenbury NJ, Kempa-liehr AW. Modelling the projected separation of microlensing events using systematic time-series feature engineering. Astron Comput [Internet]. 2021;35:100460. https://doi.org/10.1016/j.ascom.2021.100460.
61. Wang Q, Yuan Z, Du Q, Li X. GETNET: a general end-to-end 2-D CNN framework for hyperspectral image change detection. IEEE Trans Geosci Remote Sens. 2019;57:3–13.
62. Li Y, Xu Z, Han W, Cao H, Umarov R, Yan A, et al. HMD-ARG: hierarchical multi-task deep learning for annotating antibiotic resistance genes. Microbiome. 2021;9:1–12.
63. Bo-yu A, Man H, Xiang-yue X, Wen-jin M, Ling-li H, Fu-nan W, et al. Research progress of risk assessment on veterinary antimicrobial resistance. Chinese J Antibiot. 2021;46:27–33.
64. Collineau L, Carmo LP, Endimiani A, Magouras I, Müntener C, Schüpbach-Regula G, et al. Risk ranking of antimicrobial-resistant hazards found in meat in Switzerland. Risk Anal. 2018;38:1070–84.
65. Collineau L, Chapman B, Bao X, Sivapathasundaram B, Carson CA, Fazil A, et al. A farm-to-fork quantitative risk assessment model for Salmonella Heidelberg resistant to third-generation cephalosporins in broiler chickens in Canada. Int J Food Microbiol [Internet]. 2020;330:108559. https://doi.org/10.1016/j.ijfoodmicro.2020.108559.
66. Hoa PTP, Managaki S, Nakada N, Takada H, Shimizu A, Anh DH, et al. Antibiotic contamination and occurrence of antibiotic-resistant bacteria in aquatic environments of northern Vietnam. Sci Total Environ [Internet]. 2011;409:2894–901. https://doi.org/10.1016/j.scitotenv.2011.04.030.
67. Beaudequin D, Harden F, Roiko A, Mengersen K. Utility of Bayesian networks in QMRA-based evaluation of risk reduction options for recycled water. Sci Total Environ [Internet]. 2016;541:1393–409. https://doi.org/10.1016/j.scitotenv.2015.10.030.
68. Pouillot R, Beaudeau P, Denis JB, Derouin F. A Quantitative risk assessment of waterborne cryptosporidiosis in france using second-order Monte Carlo simulation. Risk Anal. 2004;24:1–17.
69. Monaco DC, Zapata L, Hunter E, Salomon H, Dilernia DA. Resistance profile of HIV-1 quasispecies in patients under treatment failure using single molecule, real-time sequencing. Aids. 2020;34:2201–10.
70. 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;10:1–12.
71. Okada K, Araki M, Sakashita T, Ma B, Kanada R, Yanagitani N, et al. Prediction of ALK mutations mediating ALK-TKIs resistance and drug re-purposing to overcome the resistance. EBioMedicine [Internet]. 2019;41:105–19. https://doi.org/10.1016/j.ebiom.2019.01.019.
72. Cassini A, Högberg LD, Plachouras D, Quattrocchi A, Hoxha A, Simonsen GS, et al. Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. Lancet Infect Dis. 2019;19:56–66.
73. Zhao S, Tyson GH, Chen Y, Li C, Mukherjee S, Young S, et al. Whole-genome sequencing analysis accurately predicts antimicrobial resistance phenotypes in Campylobacter spp. Appl Environ Microbiol. 2016;82:459–66.
74. Tyson GH, McDermott PF, Li C, Chen Y, Tadesse DA, Mukherjee S, et al. WGS accurately predicts antimicrobial resistance in Escherichia coli. J Antimicrob Chemother. 2015;70:2763–9.
75. Walker TM, Kohl TA, Omar SV, Hedge J, Del Ojo Elias C, Bradley P, et al. Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. Lancet Infect Dis. 2015;15:1193–202.
76. Gordon NC, Price JR, Cole K, Everitt R, Morgan M, Finney J, et al. Prediction of staphylococcus aureus antimicrobial resistance by whole-genome sequencing. J Clin Microbiol. 2014;52:1182–91.
77. Moran RA, Anantham S, Holt KE, Hall RM. Prediction of antibiotic resistance from antibiotic resistance genes detected in antibiotic-resistant commensal Escherichia coli using PCR or WGS. J Antimicrob Chemother. 2017;72:700–4.
78. Pesesky MW, Hussain T, Wallace M, Patel S, Andleeb S, Burnham CAD, et al. Evaluation of machine learning and rules-based approaches for predicting antimicrobial resistance profiles in gram-negative bacilli from whole genome sequence data. Front Microbiol. 2016;7:1–17.
79. Rishishwar L, Petit RA, Kraft CS, Jordana IK. Genome sequence-based discriminator for vancomycin-intermediate Staphylococcus aureus. J Bacteriol. 2014;196:940–8.
80. Davis JJ, Boisvert S, Brettin T, Kenyon RW, Mao C, Olson R, et al. Antimicrobial resistance prediction in PATRIC and RAST. Sci Rep [Internet]. 2016;6:1–12. https://doi.org/10.1038/srep27930.
81. Her HL, Wu YW. A pan-genome-based machine learning approach for predicting antimicrobial resistance activities of the Escherichia coli strains. Bioinformatics. 2018;34:i89–95.
82. Valizadehaslani T, Zhao Z, Sokhansanj BA, Rosen GL. Amino acid K-mer feature extraction for quantitative antimicrobial resistance (AMR) prediction by machine learning and model interpretation for biological insights. Biology. 2020;9:1–92.
83. Maguire F, Rehman MA, Carrillo C, Diarra MS, Beiko RG. Identification of primary antimicrobial resistance drivers in agricultural nontyphoidal *Salmonella enterica* serovars by using machine learning. mSystems. 2019;4:1–17.
84. Feretzakis G, Loupelis E, Sakagianni A, Kalles Di, Lada M. Christopoulos C, et al. Using machine learning algorithms to predict antimicrobial resistance and assist empirical treatment. *Stud Health Technol Inform*. 2020;272:75–8.
85. Lakin SM, Kuhnle A, Alipanahi B, Noyes NR, Dean C, Muggli M, et al. Hierarchical Hidden Markov models enable accurate and diverse detection of antimicrobial resistance sequences. Commun Biol [Internet]. 2019;2. https://doi.org/10.1038/s42003-019-0545-9.
86. Khaledi A, Weimann A, Schniederjans M, Asgari E, Kuo T, Oliver A, et al. Predicting antimicrobial resistance in Pseudomonas aeruginosa with machine learning-enabled molecular diagnostics. EMBO Mol Med. 2020;12:1–19.
87. Kulshrestha S, Panda S, Nayar D, Dohe V, Jarali A. Prediction of antimicrobial resistance for disease-causing agents using machine learning. In: Proceedings of the second international conference of intelligence computer controlled system. 2019;972–5.
88. Martínez-Agüero S, Mora-Jiménez I, Lérida-García J, Álvarez-Rodríguez J, Soguero-Ruiz C. Machine learning techniques to

identify antimicrobial resistance in the intensive care unit. Entropy. 2019;21:1–24.
89. Nguyen M, Wesley Long S, McDermott PF, Olsen RJ, Olson R, Stevens RL, et al. Using machine learning to predict antimicrobial MICs and associated genomic features for nontyphoidal Salmonella. J Clin Microbiol. 2019;57:1–15.
90. Liu Z, Deng D, Lu H, Sun J, Lv L, Li S, et al. Evaluation of machine learning models for predicting antimicrobial resistance of Actinobacillus pleuropneumoniae from whole genome sequences. Front Microbiol. 2020;11:1–7.
91. Lecun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521:436–44.
92. Alipanahi B, Delong A, Weirauch MT, Frey BJ. Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. Nat Biotechnol [Internet]. 2015;33:831–8. https://doi.org/10.1038/nbt.3300.
93. Stokes JM, Yang K, Swanson K, Jin W, Cubillos-Ruiz A, Donghia NM, et al. A deep learning approach to antibiotic discovery. Cell [Internet]. 2020;180:688–702.e13. https://doi.org/10.1016/j.cell.2020.01.021.
94. Ruppé E, Ghozlane A, Tap J, Pons N, Alvarez AS, Maziers N, et al. Prediction of the intestinal resistome by a three-dimensional structure-based method. Nat Microbiol [Internet]. 2019;4:112–23. https://doi.org/10.1038/s41564-018-0292-6.
95. Duranti S, Lugli GA, Mancabelli L, Turroni F, Milani C, Mangifesta M, et al. Prevalence of antibiotic resistance genes among human gut- derived bifidobacteria. Appl Environ Microbiol. 2017;83:1–14.