Antimicrobial learning systems: an implementation blueprint for artificial intelligence to tackle antimicrobial resistance

Alex Howard, Stephen Aston, Alessandro Gerada, Nada Reza, Jason Bincalar, Henry Mwandumba, Tom Butterworth, William Hope, Iain Buchan**

The proliferation of various forms of artificial intelligence (AI) brings many opportunities to improve health care. AI models can harness complex evolving data, inform and augment human actions, and learn from health outcomes such as morbidity and mortality. The global public health challenge of antimicrobial resistance (AMR) needs large-scale optimisation of antimicrobial use and wider infection care, which could be enabled by carefully constructed AI models. As AI models become increasingly useful and robust, health-care systems remain challenging places for their deployment. An implementation gap exists between the promise of AI models and their use in patient and population care. Here, we outline an adaptive implementation and maintenance framework for AI models to improve antimicrobial use and infection care as a learning system. The roles of AMR problem identification, law and regulation, organisational support, data processing, and AI development, assessment, maintenance, and scalability in the implementation of AMR-targeted AI models are considered.

Lancet Digit Health 2024; 6: e79–86

*Joint senior authors

Department of Antimicrobial Pharmacodynamics and Therapeutics, Institute of Systems, Molecular and Integrative Biology (A Howard MBChB, S Aston PhD, A Gerada MD, N Reza MBChB, Prof W Hope PhD)**, Department of Health Data Science** (J Bincalar DBA) **and Department of Public Health, Policy and Systems, Institute of Population Health** (Prof I Buchan MD)**, University of Liverpool, Liverpool, UK; Liverpool University Hospitals NHS Foundation Trust, Liverpool, UK** (A Howard, S Aston, A Gerada, N Reza, J Bincalar, Prof W Hope)**; Malawi Liverpool Wellcome Programme, Kamuzu University of Health Sciences, Blantyre, Malawi** (Prof H Mwandumba PhD)**; Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, UK** (Prof H Mwandumba)**; Combined Intelligence for Public Health Action, NHS Cheshire and Merseyside, Warrington, UK** (T Butterworth BSc, Prof I Buchan)

Correspondence to: Alex Howard, Department of Antimicrobial Pharmacodynamics and Therapeutics, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool L7 8TX, UK **alexander.howard@liverpool.ac.uk**

See **Online** for appendix

Introduction

The evolution of antimicrobial resistance (AMR) presents challenges and opportunities for infection care. Management of AMR necessitates antimicrobial stewardship (AMS) to balance cure of infection with collateral damage, and infection prevention and control (IPC) to prevent AMR transmission. AMS and IPC teams are needed to implement active surveillance and behavioural and system interventions in inflexible health-care systems that are characterised by incomplete data capture, dwindling pipelines of new antimicrobials, one-size-fits-all antimicrobial treatment formularies, resource pressures, and poorly valued or implemented diagnostic innovations.

Artificial intelligence (AI) interventions targeted at AMR (AMR-AI) could be assets to AMS and IPC teams—their ability to rapidly process, interpret, and action data from clinical care could provide adaptability in changing health-care environments. Health-care systems, however, are accustomed to evaluating and implementing fixed interventions for fixed targets (eg, rigid licensing of new antimicrobial agents). The effective deployment of adaptable, learning AMR-AI models requires a bespoke implementation approach.

This Viewpoint outlines an AMR-bespoke systematic workflow-based approach (table) for assessing and implementing AMR-AI models as part of sociotechnical entities called antimicrobial learning systems (ALSs), and aims to identify AMR-specific opportunities, barriers, and solutions to facilitate their realisation (key terminology is listed in the appendix p 1).

AMR problem identification: the suitability of AMR for an AI solution

Health-care pathways (figure 1) are characterised by uncertainty that drives inappropriate antimicrobial prescribing and missed opportunities to prevent AMR transmission. AMR-AI models can reduce uncertainty by predicting and targeting AMR faster than microbiology laboratory diagnostics that are largely reliant on slow culture-based methods. These predictions can inform clinical decision support with AMR-AI models that could help appropriately target antimicrobial therapy and protective isolation to prevent uncontrolled AMR infection, unnecessary AMR selection pressure, and onward transmission.[1]

To perform these tasks, AMR-AI models could explore data networks of AMR transmission risk that cross geographical and health-care boundaries (eg, previous antimicrobial treatment, travel history, and hospital or nursing home admission), incorporating non-causally intuitive factors. Navigating these networks could better quantify the effect of antimicrobial prescription by detecting AMR generation, onward transmission, and subsequent death from AMR. Bioinformatics AMR-AI models could combine clinical, pharmacodynamic, genomic, and epidemiological information to develop more accurate and timely AMR diagnostics and set more accurate in vitro clinical breakpoints (eg, the area size of absent bacterial growth required around an antimicrobial disk to be reported as susceptible).[2]

These applications present several challenges. First, AMR health-care data is sparse (particularly in primary care settings where most antimicrobial prescribing occurs), potentially hindering predictive accuracy. Second, AMR-AI decision support tools can cause automation bias that overrules well founded clinician instincts (eg, delaying antimicrobial treatment on algorithmic advice despite clinical intuition pointing towards sepsis).[3] Third, automating and influencing parts of complex health-care AMR systems might create unintended downstream consequences (eg, changes in antimicrobial prescribing patterns causing drug shortages or affecting demand for therapeutic drug monitoring). Last, quantification of the effect of antimicrobial treatment raises ethical issues about utilitarian AMR-AI models prioritising the greatest good for the greatest number over individual needs.

	Workflow	Key questions
1	AMR problem identification	Is AI the right solution to the AMR problem?
2	Law and regulation	Is there a legal and regulatory basis to develop an AMR-AI model, and what governance will cover the intervention once it is in place?
3	Securing support	What departmental, organisational, national, and international appetite and resources exist for AMR-AI models?
4	AMR data processing	What AMR data availability, quality, security, linkage, and actionability issues need to be addressed?
5	AMR-AI development	What algorithmic methods are suitable for the AMR problem, and what level of performance is required?
6	AMR-AI assessment	How will the effect of AI on the AMR problem be measured?
7	AMR-AI maintenance	How will the effect of AI on the AMR problem be sustained?
8	AMR-AI scalability	Can use of the AMR-AI be adapted and expanded beyond the local and health-care settings to target a range of global settings and One Health AMR drivers?

AI=artificial intelligence. AMR=antimicrobial resistance.

Table: **A workflow-based framework for planning AMR-AI models that forms the structure of this Viewpoint**

The clinical effect of AMR-AI models therefore relies on identifying where in health-care pathways limited improved prediction could reduce AMR, prevent death from AMR, and improve the efficiency of tackling AMR. Identifying these targets, and their margins of error, requires a multidisciplinary, translational approach. For example, one that converges biological, organisational, ethical, and technical perspectives (eg, community doctors who use susceptibility prediction for oral antimicrobials might perceive margin of error differently to how intensive care physicians perceive margin of error for intravenous antimicrobials). Where individual and population consequences are in disagreement, holistic ethical approaches incorporating deontological right and wrong concepts might better calibrate these priorities.[4]

Decision support AMR-AI models that stop short of definitive recommendations (eg, only providing probability of antimicrobial susceptibility, credibility and confidence intervals, and management options rather than a recommended decision) can mitigate for, but not remove, automation bias; for example, an AMR-AI model providing a low probability of sepsis could cause a clinician to go against their intuition. Assessing a problem for AMR-AI decision support therefore requires consideration of the consequences of decision support intervention in their entirety (ie, from data input to the effects of the clinical decision) to predict clinician responses to outputs and their presentation (eg, a graphic user interface that states the probability of cure with amoxicillin is 90% *vs* one that states the probability of organism amoxicillin susceptibility is 0·9). Thought experiments involving clinicians and other target users can play out scenarios involving AMR-AI models to predict the end-to-end effects of their deployment.

Law and regulation: the basis for AMR-AI models

Decision support AMR-AI models that provide specific antimicrobial or infection control recommendations could present more medical device regulatory barriers than models that only provide probabilities of events or help to drive system efficiencies (eg, microbiology test triaging).[5] The global effect of AMR is less immediate than viral pandemics, but is likely to be more catastrophic; there is a strong argument for leveraging similar population health legislation to access health-care data for AMR control. AMR is less well understood by the public than viral pandemics, and data security lapses have increased public interest in data sharing, storage, and use. Legislation varies internationally, so suitable legal expertise in data protection (appendix p 2) is required.

AMR can disproportionately affect some populations (eg, deprived or marginalised communities), potentially causing prejudicial algorithmic biases that contravene human rights legislation. Such biases are rarely obvious, and poorly curated training datasets can create long-lasting effects (the so-called negative legacy effect). For example, algorithms trained in maternity hospitals with high AMR rates might erroneously attribute AMR risk to pregnancy in other datasets. In low AMR prevalence settings, algorithms become more accurate in predicting AMR among people at lowest risk because they dominate the training data.[6] Algorithmic bias therefore needs consideration throughout the implementation workflow. So-called AI fairness can be built into algorithms to detect infringements of individuals' rights and provide bias control.[7] Opinion and policy leaders in infection, technology, ethics, law, policy, and regulation should be engaged by AMR-AI implementation leaders to better clarify how AI and AMR interact with legal and regulatory frameworks.

Securing support: resource and buy-in

A range of regional health system assets (appendix p 3) and personnel (appendix p 5) are required for AMR-AI implementation.[8,9] Governments are increasingly recognising the public health risk of AMR, and the harm that inadequate data availability causes to patients—resultant centrally mandated targets for organisational AMS and IPC provide persuasive arguments to secure local resources. AMR is high on the priority list of international organisations that are potential routes of collaboration and funding; for example, WHO initiatives have spurred multinational data science-based AMR management programmes.

AMR-AI models are vulnerable to similar cultural, behavioural, and psychological factors to other AMR

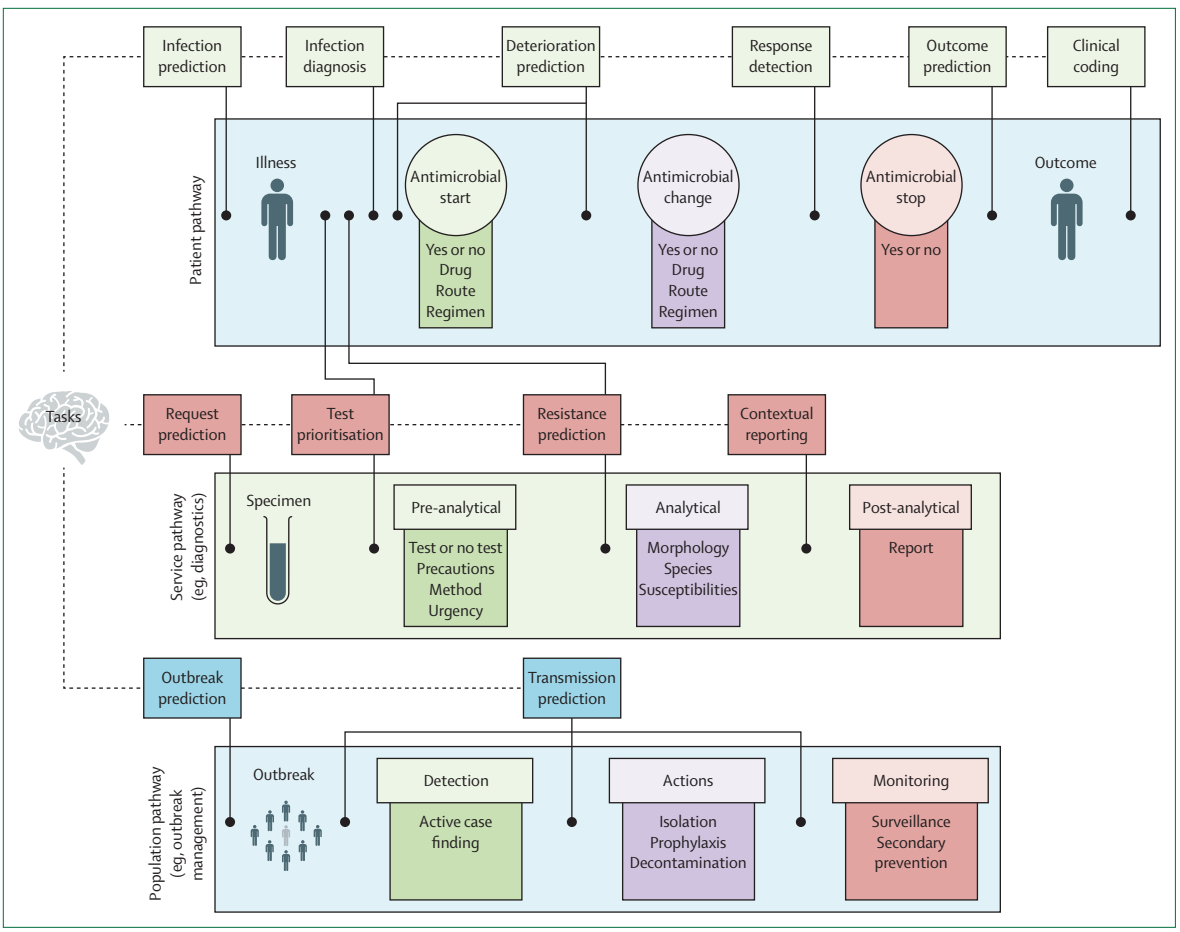

Figure 1: **Examples of tasks in inter-related infection health-care pathways that are often labour intensive and computationally challenging for humans, and therefore might present opportunities for antimicrobial resistance-artificial intelligence models to augment and improve systems**

interventions (eg, education targeted at handwashing and antimicrobial prescribing).[10–12] Managing these factors relies on local expertise and experience (eg, IPC nurses and AMS pharmacists) and bespoke ethnographic and qualitative data collection techniques (eg, questionnaires, interviews, and real-world observation; appendix p 6).[13]

AMR data processing: availability, quality, security, linkage, and actionability

The combinatorial complexity of patient–disease–organism–drug interactions can provide rich short-term AMR datasets that evolve over hours to days in secondary care. However, a sparsity of unbiased, quantifiable longer-term data inputs (predictors) and outputs (outcomes) in the community can present a challenge. Changing AMR rates over time can cause concept drift in training datasets that affects predictions.[14] Antimicrobial treatment pathways often span health care and community settings, fragmenting dataflows. Not all patients are screened for AMR, infections are often not microbiologically confirmed, and antimicrobial information might be undetected electronically (eg, in paper prescription charts). Behavioural and cultural drivers of AMR (eg, hand hygiene practice) and clinician intuition (eg, end-of-the-bed recognition of sepsis) might not readily produce machine-readable inputs.

AMR-AI models therefore require minimisation of missing data (eg, electronic handwriting recognition for paper prescription charts and automated population of organism species-specific intrinsic AMR) and data linkage between microbiology laboratory management information systems, electronic prescriptions, and other systems. To action these insights, data should be written back into front-end clinical systems (eg, a smartphone graphic user interface for an antimicrobial prescribing decision aid)—such interfaces require regular testing, version control, and release management.[15] Hardware that houses AMR-AI clinical interfaces should be suitable for decontamination to minimise fomite AMR transmission risk.

ALSs should be sociotechnical entities designed to embed AMR-AI models in population health and front-line patient care—these systems should be mapped to regional health systems to facilitate the technical and

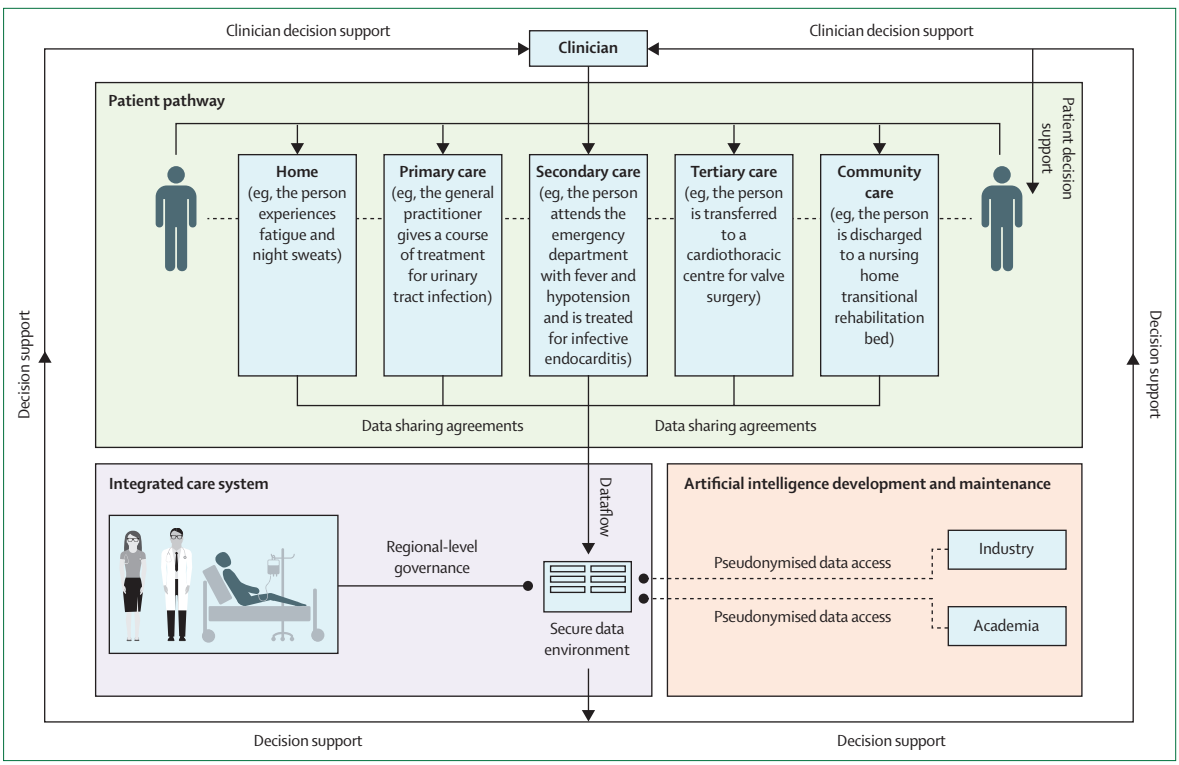

Figure 2: **A model for how regional integrated care systems should house antimicrobial learning systems' governance, security, and maintenance of dataflows to enable continuous delivery of antimicrobial resistance-artificial intelligence decision support to patient pathways**

governance aspects of the required afferent and efferent dataflows (figure 2).

AMR diagnostic data is unusual in that some antimicrobial susceptibilities are intentionally hidden from clinicians for AMS purposes, and raw data is parsed into strings of text for clinician interfaces—direct data extraction from microbiology laboratory information management systems is therefore essential for data completeness. Frequent systemic changes in response to the ever-expanding AMR problem might affect data quality, necessitating end-to-end validation. The number of validation cases or observations and factors or variables required depends on the application and sources of data variability (eg, susceptibility interpretation changes for *Escherichia coli* will have a greater effect than taxonomic changes for a rare commensal organism)—heuristic data validation and repair approaches are therefore required. The security and encryption of data transfer and storage, details of data sharing and data protection effect, robustness of pseudonymisation, and understanding of the consequences of automated deployment (eg, automated contact tracing for multidrug-resistant sexually transmitted infections) are crucial to maintain privacy.

AMR-AI development: algorithmic methods and performance indicators

Classification algorithms (eg, logistic regression, support vector machine, and decision trees) output categorical variables and are therefore useful for predicting antimicrobial susceptibility (ie, susceptible to standard dosing, susceptible at increased exposure, or resistant) or diagnosis (eg, excluding bacterial infection to prevent unnecessary antimicrobial treatment). AMR issues that influence the choice of classification algorithm are the application (eg, predicting AMR might be more challenging at individual level than at population level), data dimensionality (eg, genomic resistance data will have many rows and columns), variable interdependence (eg, plasmid colocation of aminoglycoside and fluoroquinolone resistance, or enzymes conferring resistance to multiple β-lactams), and generalisability (overfitting might prevent scalable application given the geographical variations in AMR).

Classification algorithms are often supervised (ie, use labelled data) and explainable (ie, the roles of variables in model performance are transparent), which theoretically enables translational infection-data science experts to better validate outputs, although this is an area of debate.[16,17] Bayesian inference underpins many explainable methods because it mimics human decision making—previous information (eg, previous antimicrobial treatment) is integrated with new data (eg, new laboratory results) to predict an event (eg, AMR). Causal, counterfactual (ie, alternative event) thought processes can identify care pathway targets for Bayesian AMR-AI models (figure 3).[18]

Regression algorithms can output continuous variable predictions (eg, antimicrobial minimum inhibitory concentrations). Clustering techniques, such as K-means, can identify epidemiological clusters (eg, methicillin-resistant *Staphylococcus aureus* outbreaks) and are often unsupervised or unexplainable particularly in neural networks—this might limit translational expert input and potentially affect uptake of unsupervised AMR-AI models among health-care professionals trained in transparency, accountability, and physiological aetiology.[19] In complex systems with a lower probability of harm (eg, microbiology laboratories), supervision and explainability might not be necessary or desirable.

The predictive performance required for AMR-AI models is antimicrobial–organism–patient–system combination specific and hinges on the consequences of incorrect predictions as much as accuracy. Utility functions quantifying these consequences need weighting carefully to avoid over-favouring unlikely but easily identifiable effects of antimicrobial under-prescribing (eg, death from sepsis). Translational infection-data science expertise can identify performance issues, their cause and solutions, and whether their consequences merit delaying or withdrawing deployment (eg, an AMR-AI model with serum aminoglycoside level prediction intervals that regularly straddle cutoffs associated with nephrotoxicity or ototoxicity). The provision of credibility and prediction intervals can indicate the confidence of AMR predictions to implementation teams and end users.[20,21]

AMR-AI models should be periodically retrained in line with data updates (eg, monthly). This process (and the amount of data required) should also be adaptive to the specific AMR problem (eg, predicting AMR in low-prevalence areas might require more data than in high-prevalence areas), changes in health-care standard operating procedures (eg, changes in antimicrobial formulary recommendations), key organism definitions (eg, changes in clinical breakpoint definitions), and changes in AMR rates. Translational infection-data scientist expertise can identify changes important enough to merit ad hoc retraining.

AMR-AI assessment: evaluating effects

Randomised controlled trials (RCTs) housed by the ALS should assess AMR-AI efficacy, safety, and fidelity (the degree to which AI uptake and use matches that intended; appendix p 7).[22] Clinical trial design and outcomes are important; for example, AMR-AI models that narrow the spectrum of prescribed antimicrobial activity are unlikely to elicit superior 7-day mortality rates, and cluster RCTs will be required for AMR-AI models deployed in primary care.[23] Pragmatic RCT designs could use electronic health-care records for outcome assessment, and simulation might be useful when routine randomisation is impractical or unethical.[24]

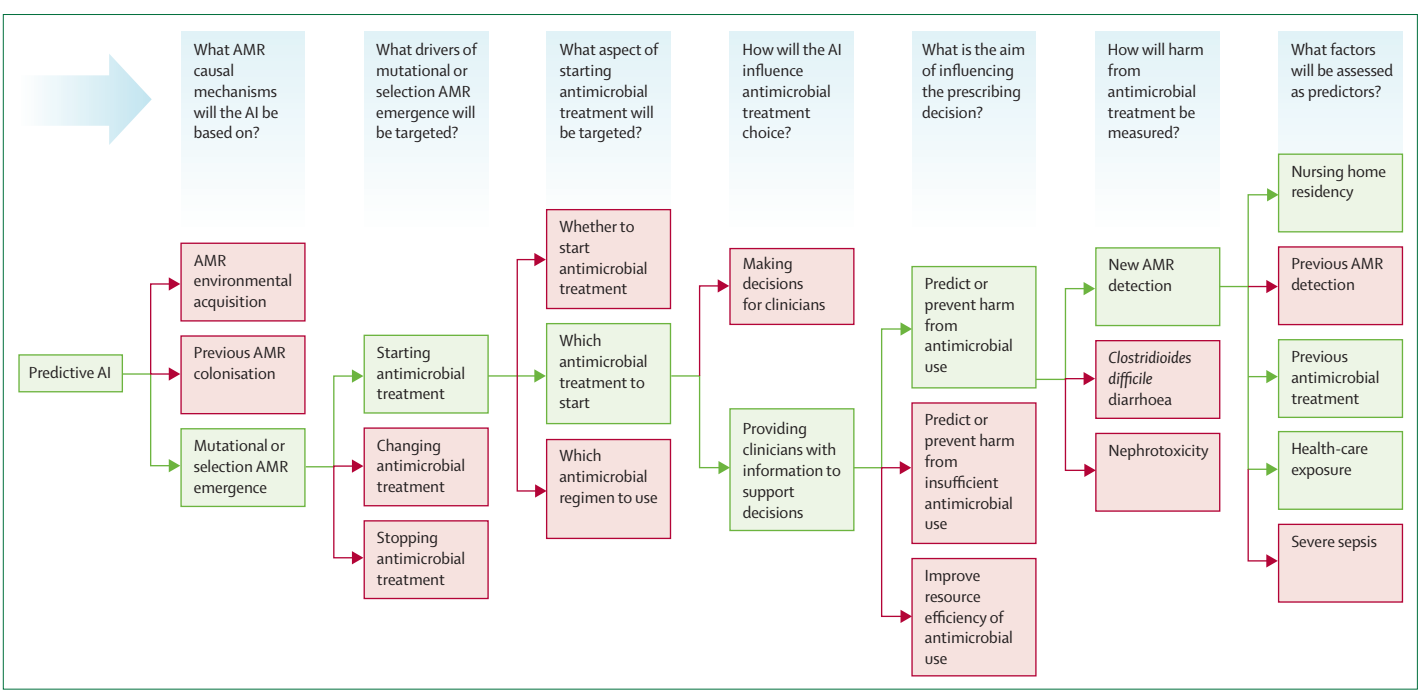

Figure 3: **A causal, counterfactual thought process for designing an explainable AMR-AI model using Bayesian statistics**
In this example, the boxes in green reflect a decision process moving from left to right as an AI tool predicts the AMR effect of different antimicrobial choices. AI=artificial intelligence. AMR=antimicrobial resistance.

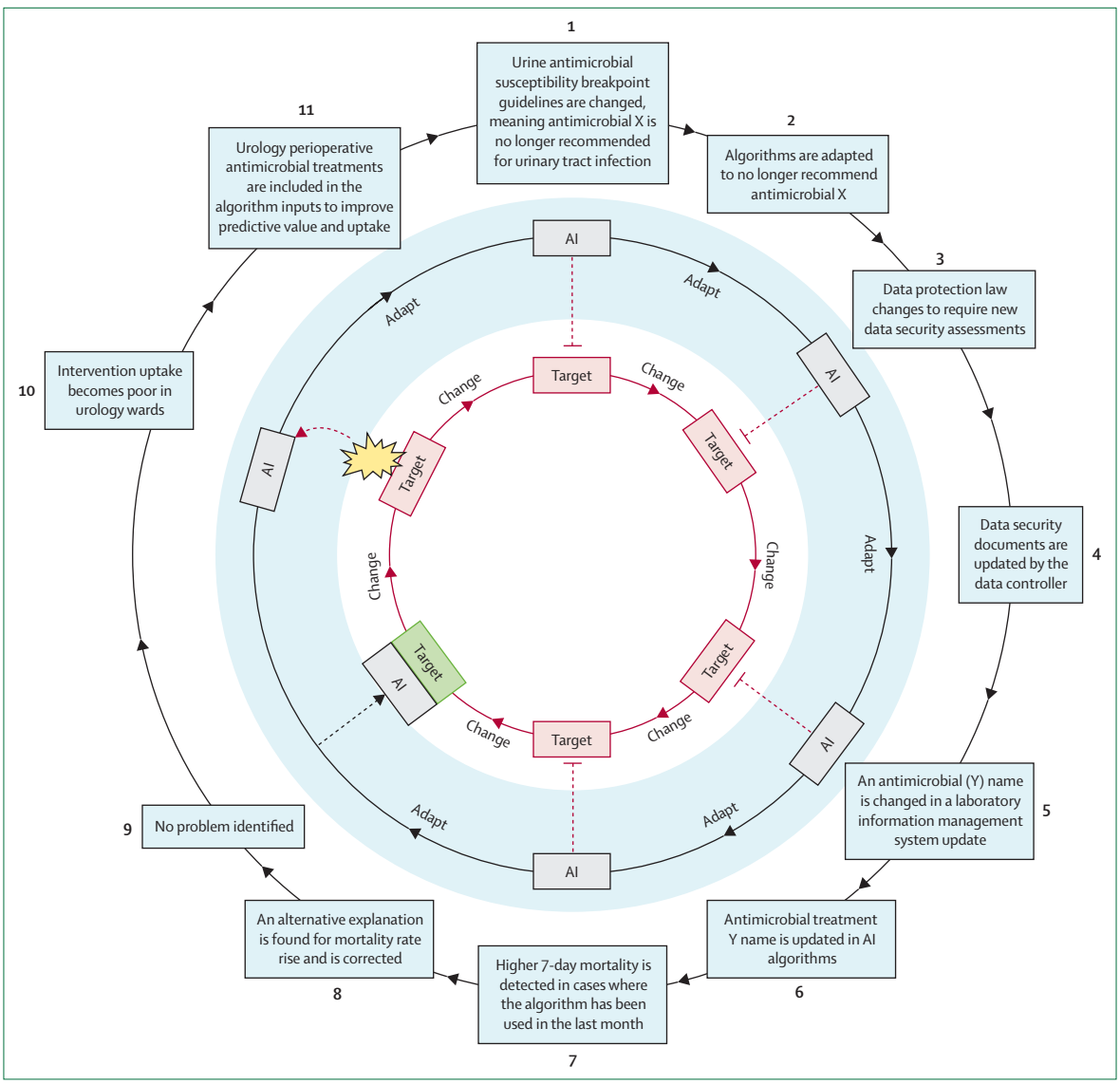

Figure 4: **Adaptive maintenance of a decision support AMR-AI model for urinary tract infection**
Continuous evaluation of health-care environments detects changes that endanger AI-problem fit (numbers 1, 3, 5, 7, and 10). Continuous tuning (numbers 2, 4, 6, 8, and 11) is therefore required to optimise fit and maintain AMR-AI intervention effectiveness.[25–27] AI=artificial intelligence. AMR=antimicrobial resistance.

AMR-AI maintenance: sustaining effects

AMR-AI models are more vulnerable to change than other health-care AI models because they have to account for organism, patient, drug, and system variability; maintaining their clinical efficacy relies on constant adaptive maintenance to sustain the fitness for purpose of the AMR-AI intervention for the AMR target in dynamic health-care environments (figure 4).[25–27]

The continuous and cumulative information generation of ALSs from routine care should be used to improve the predictive performance of AMR-AI models (eg, in under-represented groups). Over time, aspects of maintenance should be automated where possible for routine applications (eg, laboratory triage of non-sterile specimens).

AMR-AI scalability: global settings and One Health

ALSs can only affect the global AMR problem once they are interconnected; international collaboration and knowledge mobilisation is essential. The constituent AMR-AI models of ALSs should be written in open-source code run on open-source software to ensure operability in low-income and middle-income countries. Algorithms should be modular and transparent where possible, enabling local infection-data science

Search strategy and selection criteria

To support the concepts used in this Viewpoint, advanced searches of PubMed, Ovid MEDLINE, and Google Scholar literature databases were performed for articles in all languages, from the past 15 years. The targeted search strategies that were used are summarised in the appendix (pp 1–2). Manual result searches determined sources for inclusion based on relevance to the Viewpoint being presented.

translational experts to adapt inputs to local AMR risk factors and resistance rates. Global AMR data science projects should include the expansion of hardware availability, cloud computing, and data science training for clinicians to build translational expertise.

AMR is influenced by One Health factors, such as agricultural antimicrobial use and wastewater management—such domains typically have separate governance structures, dataflows, and interests. One Health data science collaborations and networks should find common ground to incentivise integration of these dataflows—without this, AMR-AI models deployed within health care might not exert enough of an effect to turn the global tide of AMR.

Conclusions

Bridging the implementation gap between AI innovation and tackling AMR presents technical, regulatory, organisational, and human challenges. Learning systems built on integrated dataflows, governance, and technologies have the potential to close this gap. Translational expertise between AMR and AI fields are essential to appropriately design, maintain, normalise, and globalise AMR-AI models in infection care and realise the potential for AI models to support clinician-driven AMR minimisation strategies.

Contributors

AH wrote and edited the manuscript and created the diagrams. SA, AG, NR, JB, WH, and IB commented on and edited the manuscript. HM and TB commented on the manuscript.

Declaration of interests

AH declares consulting work for Pfizer outside the submitted work. HM received support for research not related to the current manuscript from UK Research and Innovation (UKRI), the Bill and Melinda Gates Foundation, and National Institutes of Health; is on the data safety monitoring board for the IMPROVE trial; is an interview panel member for the Wellcome career development award; and is a council member of the International Society for Infectious Diseases. IB holds a senior investigator award with the National Institute for Health and Care Research, and received personal fees and other fees from AstraZeneca outside the submitted work. WH holds or has held research grants with UKRI, EU, F2G, Spero Therapeutics, Antabio, Pfizer, Bugworks, Phico Therapeutics, BioVersys, Global Antibiotic Research & Development Partnership (GARDP), and NAEJA-RGM; is or has been a consultant for Appili Therapeutics, F2G, Spero Therapeutics, NAEJA-RGM, Centauri, Pfizer, Phico Therapeutics, Pulmocide, Amplyx, Mundipharma Research, and VenatoRx; is a member of the Specialist Advisory Committee for GARDP; and is the Specialty National Co-lead for Infectious Diseases for the National Institute of Health Research. All other authors declare no competing interests.

Acknowledgments

We thank Peter Bower from the University of Manchester (Manchester, UK) for his expert advice. This research was funded in part by the Wellcome Trust [grant reference: 226691/Z/22/Z].

References

1 Shortliffe EH, Sepúlveda MJ. Clinical decision support in the era of artificial intelligence. *JAMA* 2018; **320:** 2199–200.
2 Doern GV, Brecher SM. The clinical predictive value (or lack thereof) of the results of in vitro antimicrobial susceptibility tests. *J Clin Microbiol* 2011; **49:** S11–14.
3 Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. *J Am Med Inform Assoc* 2012; **19:** 121–27.
4 Bolton WJ, Badea C, Georgiou P, Holmes A, Rawson TM. Developing moral AI to support decision-making about antimicrobial use. *Nat Mach Intell* 2022; **4:** 912–15.
5 Schönberger D. Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications. *Int J Law Inf Technol* 2019; **27:** 171–203.
6 Krawczyk B. Learning from imbalanced data: open challenges and future directions. *Prog Artif Intell* 2016; **5:** 221–32.
7 Chen RJ, Wang JJ, Williamson DFK, et al. Algorithmic fairness in artificial intelligence for medicine and healthcare. *Nat Biomed Eng* 2023; **7:** 719–42.
8 Burnett S, Benn J, Pinto A, Parand A, Iskander S, Vincent C. Organisational readiness: exploring the preconditions for success in organisation-wide patient safety improvement programmes. *Qual Saf Health Care* 2010; **19:** 313–17.
9 Garcia-Perez A, Cegarra-Navarro JG, Sallos MP, Martinez-Caro E, Chinnaswamy A. Resilience in healthcare systems: cyber security and digital transformation. *Technovation* 2023; **121:** 102583.
10 Ohta R. Social cognitive theory for antimicrobial stewardship. March 15, 2023. https://bjgp.org/content/social-cognitive-theory-antimicrobial-stewardship (accessed Aug 2, 2023).
11 Alradini F, Bepari A, Al Nasser BH, Al Gheshem EF, Al Ghamdi WK. Perceptions of primary health care physicians about the prescription of antibiotics in Saudi Arabia: based on the model of theory of planned behaviour. *Saudi Pharm J* 2021; **29:** 1416–25.
12 Heid C, Knobloch MJ, Schulz LT, Safdar N. Use of the health belief model to study patient perceptions of antimicrobial stewardship in the acute care setting. *Infect Control Hosp Epidemiol* 2016; **37:** 576–82.
13 Mosavianpour M, Sarmast HH, Kissoon N, Collet J-P. Theoretical domains framework to assess barriers to change for planning health care quality interventions: a systematic literature review. *J Multidiscip Healthc* 2016; **9:** 303–10.
14 Canovas-Segura B, Morales A, Martinez-Carrasco AL, et al. Improving interpretable prediction models for antimicrobial resistance. IEEE 32nd International Symposium on Computer-Based Medical Systems; June 5–7, 2019 (abstr 019).
15 Kumar A, Maskara R, Maskara S, Chiang I-J. Conceptualization and application of an approach for designing healthcare software interfaces. *J Biomed Inform* 2014; **49:** 171–86.
16 Loh HW, Ooi CP, Seoni S, Barua PD, Molinari F, Acharya UR. Application of explainable artificial intelligence for healthcare: a systematic review of the last decade (2011–2022). *Comput Methods Programs Biomed* 2022; **226:** 107161.
17 Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. *Lancet Digit Health* 2021; **3:** e745–50.
18 Kyrimi E, McLachlan S, Dube K, Neves MR, Fahmi A, Fenton N. A comprehensive scoping review of Bayesian networks in healthcare: past, present and future. *Artif Intell Med* 2021; **117:** 102108.
19 Gille F, Jobin A, Ienca M. What we talk about when we talk about trust: theory of trust for AI in healthcare. *Intell Based Med* 2020; **1–2:** 100001.
20 de Hond AAH, Leeuwenberg AM, Hooft L, et al. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. *NPJ Digit Med* 2022; **5:** 2.
21 Seoni S, Jahmunah V, Salvi M, Barua PD, Molinari F, Acharya UR. Application of uncertainty quantification to artificial intelligence in healthcare: a review of last decade (2013–2023). *Comput Biol Med* 2023; **165:** 107441.

22 Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health* 2011; **38:** 65–76.

23 Harrer S, Shah P, Antony B, Hu J. Artificial intelligence for clinical trial design. *Trends Pharmacol Sci* 2019; **40:** 577–91.

24 Bedding A, Scott G, Brayshaw N, et al. Clinical trial simulations – an essential tool in drug development. Hertfordshire, UK: Association of British Pharmaceutical Industry, 2014.

25 Aarons GA, Green AE, Palinkas LA, et al. Dynamic adaptation process to implement an evidence-based child maltreatment intervention. *Implement Sci* 2012; **7:** 32.

26 Chambers DA, Glasgow RE, Stange KC. The dynamic sustainability framework: addressing the paradox of sustainment amid ongoing change. *Implement Sci* 2013; **8:** 117.

27 Huddlestone L, Turner J, Eborall H, Hudson N, Davies M, Martin G. Application of normalisation process theory in understanding implementation processes in primary care settings in the UK: a systematic review. *BMC Fam Pract* 2020; **21:** 52.