Code Red

How Artificial Intelligence Can Help Win the Fight Against Drug-Resistant Infections

Northeastern University

PATH

The Rockefeller Foundation

About this report

This report captures the key ideas, debates, and recommendations that emerged from a convening of international experts held in October 2025 at The Rockefeller Foundation's Bellagio Center. Bringing together leading voices from across science, policy, industry, and global health, the convening explored the current landscape of artificial intelligence in the fight against antimicrobial resistance (AMR), and the most promising opportunities for impact. The report aims to frame how researchers, funders, and policymakers should approach artificial intelligence (AI)–based tools in the AMR context and highlights the emerging questions that will shape the field in the years ahead.

Acknowledgements

We would like to thank The Rockefeller Bellagio Team, and Nadia Gilardoni in particular, without whom none of this would have been possible. This report captures the key ideas, debates, and recommendations that emerged from a convening of international experts held in October 2025 at The Rockefeller Foundation's Bellagio Center, led by Prof. Samuel Scarpino, Prof. Dame Sally Davies, Prof. Bilal Mateen, and Raghu Dharmaraju. We are also grateful to The Rockefeller Health team's leadership, Naveen Rao and Manisha Bhinge, for their support in designing this convening and for funding it (FO-Grant-0000187) as well as the time of the convenors.

Our sincere thanks go to all the experts who joined us in Bellagio and generously shared their years of hard-won wisdom to inform this report: James Anderson (International Federation of Pharmaceutical Manufacturers & Associations), Julian Beach (Medicines and Healthcare products Regulatory Agency), Raghu Dharmaraju (ARTPARK at the Indian Institute of Science), Dylan George (Gates Foundation), Valeria Gigante (World Health Organisation), Marianne Holm (Novo Nordisk Foundation), Rajaram (Laurus Labs Limited), Wengong Jin (Northeastern University), Tariq Khokhar (Wellcome Trust), Andrew Leach (LifeArc), Jessica Lee (Verily Health Sciences), Anuradha Poonepalli (Health Sciences Authority Singapore), Melanie Saville (PATH), Tavpritesh Sethi (Indraprastha Institute of Information Technology Delhi), Henry Skinner (AMR Action Fund), and Emily Yu (The Rockefeller Foundation).

Finally, we extend our gratitude to Maureen O'Hagan for copyediting the report and to Jen Fox (Atomic Fox Design) for bringing our words to life through the design of this report.

Suggested citation

Mateen, BA. Davies, S. Musselman, B. Dharmaraju, R. Scarpino, S. (Eds). *Code Red: How Artificial Intelligence Can Help Win the Fight Against Drug-Resistant Infections*. 2026. PATH, USA.

Copyright

Contents

The Editors

Prof. Samuel V. Scarpino

Samuel V. Scarpino is a Professor of Public Health and Health Sciences and Director of AI + Life Sciences in the Institute for Experiential AI at Northeastern University. Named an External Professor at the Santa Fe Institute in 2020 recognition of his contributions to complex systems science, he previously served as Vice President of Pathogen Surveillance at The Rockefeller Foundation and co-founded the data science initiative Global.health. His research has appeared in *Nature, Science, and the Proceedings of the National Academy of Sciences*, and has been covered by The New York Times, Wired, and National Geographic.

Prof. Dame Sally Davies

Dame Sally Davies is the 40th Master of Trinity College (Cambridge), and the first woman to hold the post. Appointed the UK Government's Special Envoy on Antimicrobial Resistance (AMR) in 2019, she previously served as Chief Medical Officer for England and Senior Medical Adviser to the UK Government from 2011–2019. During that period, she was a member of the WHO Executive Board (2014–2016) and co-convener of the United Nations (UN) Inter-Agency Coordination Group on AMR. In 2020, she was appointed to the UN Global Leaders Group on AMR, where she continues to advocate for global action on the topic.

Ms. Brianna Musselman

Brianna Musselman is a conflict mediation expert, having facilitated resolution and dialogue in both academic and international development settings. Her work focuses on the intersection of digital health and primary health care. With more than a decade of experience leading global health initiatives across complex and resource-constrained environments, she recently led the project management unit for the US Centers for Disease Control and Prevention (CDC) Technical Assistance Platform. Brianna currently serves as the Deputy Director (Operations) for the AI Initiative at PATH.

Prof. Bilal A. Mateen

Bilal A. Mateen is the inaugural Chief AI Officer at PATH and a member of the UN's Independent International Science Panel on AI. Trained as a physician, he has spent his career at the intersection of technology and global health, holding an honorary professorial chair in machine learning for health and leading a portfolio of more than $250 million in grant-funded digital health and AI initiatives worldwide. Prior to joining PATH, he led digital technology funding at the Wellcome Trust, overseeing a nine-figure annual budget spanning infectious disease (including AMR), mental health, and climate and health research.

Preface

By Prof. Dame Sally Davies,
Special Envoy on Antimicrobial Resistance, United Kingdom

More than one million people die every year because the drugs that should save them no longer can. A year ago, I feared becoming one of them.

Following cancer surgery, I developed a serious infection. As a doctor, I have spent much of my career warning the world about antimicrobial resistance. Lying in a hospital bed, I understood for the first time—not as a scientist, but as a patient—what it means to depend on antibiotics working. I recovered. But for a growing number of people around the world, that's not what happens.

Without effective action, the cumulative impact of AMR will swell to almost 40 million lives lost by 2050. And the burden falls hardest on people in low- and middle-income countries (LMICs) without access to clean water, safe surgery, or fully functioning health systems.

AMR is not a future threat. It is a present catastrophe we have chosen to underprioritize.

To be clear, this is not primarily a scientific failure. We understand the drivers of antimicrobial resistance. We know what good stewardship looks like. What we have lacked is the alignment of science, economics, policy, and political will to act at the speed the crisis demands. The pipeline of new antibiotics is the starkest evidence of that failure. Companies that have done everything right have nonetheless gone bankrupt, because the market for antibiotics is structurally broken. Without intervention, no rational investor chooses antibiotics over cancer treatments. The result is a thin pipeline at precisely the moment we need coordinated global action.

In October 2025, we convened twenty experts—scientists, clinicians, regulators, funders, and technologists—for three days at The Rockefeller's Bellagio Center, around a single urgent question: can artificial intelligence change the trajectory of AMR, and if so, how do we make that happen?

What we found was neither naive optimism nor techno-solutionism. AI can screen billions of molecules in hours. It can detect resistance patterns before they become untreatable outbreaks. It can guide a clinician toward the right antibiotic in real time. These are not hypothetical applications. They are happening now. But AI will not solve AMR alone. It cannot fix a broken market, replace missing diagnostic infrastructure, or bridge the equity gap that concentrates AMR deaths in the Global South while concentrating AI development in a handful of northern universities and technology companies.

What AI can do, if we govern it well, fund it strategically, and design it for the settings where the burden is heaviest, is dramatically accelerate the pace at which we find, develop, and deploy the tools we desperately need.

The science is ready. What has been missing is conviction — and a coalition willing to act.

This report is an attempt to build that coalition. The window is open.
We cannot afford to let it close.

CHAPTER 1

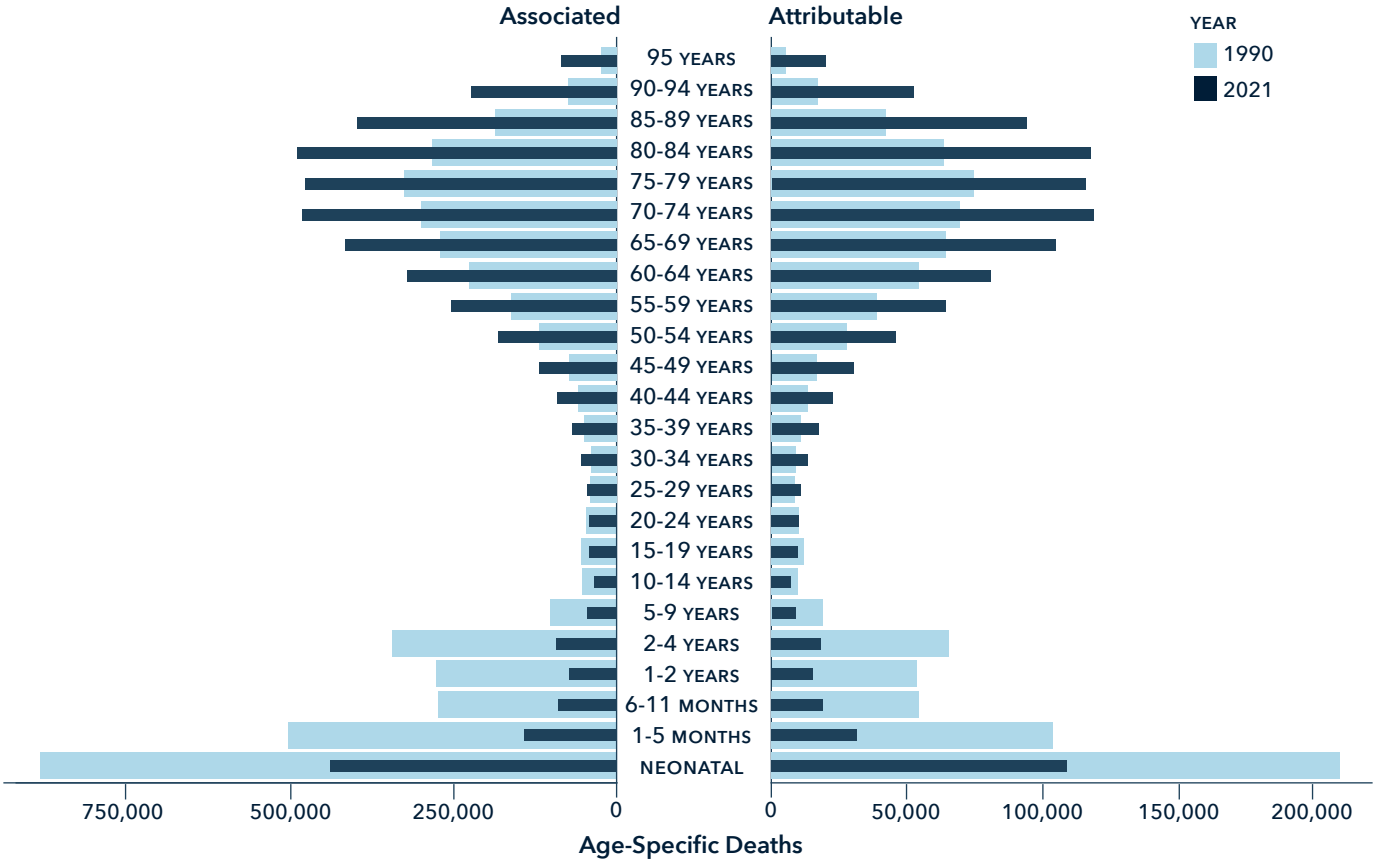
The Crisis We Can No Longer Afford to Ignore

Antimicrobial resistance (AMR[a]) is not an emerging threat; it is a crisis that has been killing at least one million people every year since 1990.[1] Bacterial AMR is directly responsible for approximately 1.2 million deaths annually. It is associated with almost 5 million additional deaths. That's more than HIV/AIDS (approximately 860,000 deaths) and malaria (approximately 640,000 deaths).[2]

The trajectory is worsening. Today, we're finding that even last-resort drugs, such as carbapenems, aren't reliably working. Under current conditions, AMR is projected to kill almost 2 million people directly each year by 2050—a 67% increase from 2021 levels. Cumulatively, almost 40 million people will die from drug-resistant infections between 2025 and 2050.[1] The scale of this projection can be difficult to comprehend: it represents roughly three deaths every minute, every day, for the next quarter century.

The distribution of this burden is also profoundly unequal. Sub-Saharan Africa bears the highest per capita death rate, at 24 per 100,000 population, followed by South Asia at 22 per 100,000. South Asia alone is projected to account for 11.8 million direct AMR deaths between 2025 and 2050.[1] High-income countries, by contrast, report (directly attributable) death rates of around 13 per 100,000. These rates are all unacceptable. However, there's no question that AMR is disproportionately a disease of poverty, inadequate infrastructure, and systems ill-equipped to detect or respond to it.

Figure 1. Deaths attributable and associated with antimicrobial resistance, by detailed age group, for 1990 and 2021



Source: Global burden of bacterial antimicrobial resistance 1990-2021: a systematic analysis with forecasts to 2050 (The Lancet)

[a] AMR, as it will be referred to throughout this report, happens when microorganisms such as bacteria evolve to resist the modern medicines (including antibiotics) that are used to fight them. This evolution is a natural process, but AMR has been supercharged because of certain practices by humans.

AMR is projected to kill almost 2 million people directly each year by 2050 — a 67% increase from 2021 levels. This equates to roughly 3 deaths every minute, every day, for the next quarter century.

Within this overall picture, one demographic trend demands particular attention. While deaths in children under five have fallen by nearly 60% since 1990, driven by vaccination programmes, improvements in water and sanitation, and better infection management, AMR deaths among adults aged 70 and older increased by 89% over the same period.[1] As populations age globally, and as resistance to carbapenems and other last-resort agents continues to rise, this trend will intensify, putting increasing pressure on scarce resources.

The crisis is not being solved; it is being redistributed and, on the whole, accelerating. Understanding how to address it requires reckoning with the fact that this is a complex trans-domain challenge, as well as the manifestation of an interconnected failure of governance, economics, and market structure that have allowed us to reach this point.

A "One Health" Problem

AMR cannot be understood, or addressed, through a purely clinical lens. It is a problem that spans human health, animal health, and the environment (including climate change) in ways that are deeply interconnected. It's a challenge requiring what is now termed a One Health approach, formally recognised by the World Health Organisation (WHO), Food and Agriculture Organisation of the United Nations (FAO), World Organisation for Animal Health (WOAH), and United Nations Environment Programme (UNEP) through their Quadripartite collaboration.[3]

Globally, approximately two-thirds to three-quarters of all antibiotics consumed are used in food animals, not humans.[4,5] In the United States, the Food and Drug Administration has confirmed that roughly 70% of medically important antibiotics sold are for veterinary or agricultural use. This is not merely a statistic. Widespread agricultural antibiotic use, particularly for growth promotion rather than treatment, creates the very conditions that accelerate AMR.[b] Conversely, a systematic review and meta-analysis commissioned by WHO found that restricting antibiotic use in food-producing animals reduced drug-resistant bacteria in animals by 10%–15% and in humans exposed to those animals by approximately 24%.[6]

Environmental contamination from pharmaceutical manufacturing represents a further, underappreciated driver. Research conducted in the industrial area of Patancheru, outside Hyderabad, found ciprofloxacin concentrations of up to 31,000 micrograms per litre in treatment plant effluent from bulk antibiotic manufacturers, which is approximately one million times higher than levels typically found in municipal sewage, and more than 1,000 times the concentration needed for resistant bacteria to flourish.[7] A 2022 study sampling 1,052 sites along 258 rivers across 104 countries found that more than a quarter of sites had active pharmaceutical ingredient concentrations at levels considered unsafe for aquatic organisms or of concern for AMR selection, with the highest contamination in sub-Saharan Africa, South Asia, and South America.[8]

[b] When antibiotics are used so widely, natural selection means that only microorganisms that can withstand them will survive, and their genes will become more prevalent in the population. As the process continues, more and more resistant microorganisms circulate through food chains, soil, and waterways. The result is that resistant organisms are everywhere.

The governance architecture to address AMR across human health, animal health, and the environment exists, at least in outline. We've had such plans, at least in theory, for a decade. The Global Action Plan on AMR, endorsed by the World Health Assembly in 2015, has led to national action plans in 178 countries.[3] However, implementation has been deeply uneven, and the mechanisms for linking human health surveillance with veterinary and environmental data remain underdeveloped in most settings.

An Economic Imperative to Act

Estimates of the economic toll of unchecked AMR vary considerably, with some projections suggesting cumulative losses as high as US$100 trillion in global gross domestic product (GDP) by 2050, alongside as many as 10 million deaths per year—figures put forward by the influential 2016 Review on Antimicrobial Resistance, commissioned by the UK Government and chaired by economist Lord Jim O'Neill. While methodologies and headline numbers differ across studies, there is broad consensus on the underlying trajectory: the burden of AMR is growing, and its human and economic costs are mounting year on year.

With the loss of life comes broader societal and economic impacts. A 2017 World Bank report modelled AMR's macroeconomic impact and found it could reduce global GDP by 1.1% to 3.8% annually by 2050, with low-income countries losing more than 5% of GDP. It projected up to 28 million additional people pushed into extreme poverty, and described the economic disruption as comparable in scale to the 2008 global financial crisis, but with no prospect of cyclical recovery.[10] A 2024 Centre for Global Development study estimated current direct inpatient costs of AMR at $66 billion per year globally, rising to $159 billion by 2050 under a business-as-usual scenario.[11]

28 million additional people pushed into poverty — an economic disruption as comparable in scale to the 2008 global financial crisis, but with no prospect of cyclical recovery.

Against these costs, the case for investment is economically overwhelming. One proposal—to create a subscription-based purchasing programme for new antibiotics that benefits both governments and pharmaceutical developers—found a 30-year return on investment of 28:1: $17.9 billion invested, yielding $495 billion in value from lives saved in the United States alone.[12]
A broader global modelling exercise produced an equivalent ratio for comprehensive AMR intervention. The maths is unambiguous: the cost of action is a fraction of the cost of inaction.

Yet this calculation has not translated into adequate investment. AMR remains dramatically underfunded relative to other disease areas. Between 2018 and 2023, oncology attracted approximately $50 billion in private investment; anti-infectives attracted just $1.2 billion.[13] That disparity may appear irrational given the human toll, but in business terms, it is entirely rational.

A Market Failure, Not a Scientific One

Most important to understand about the AMR crisis is that it is not, at its root, a scientific failure, but rather one of structural disincentives: market conditions that systematically underinvest in the development of new antimicrobials and diagnostics and that are incapable, without intervention, of rewarding innovation sufficiently to sustain a viable pipeline.

The problem has a name: the stewardship paradox. Good antimicrobial stewardship (i.e., using antibiotics appropriately and sparingly, while reserving the best agents for the most serious infections) is essential to preserving the effectiveness of existing and new drugs. But it is also fatal to the conventional pharmaceutical business model. A company that develops a genuinely superior antibiotic will, if good stewardship is practised, sell fewer doses, generate less revenue, and recover less of its

Photo: PATH/Minzayar (2018)

The AMR crisis is that it is not, at its root, a scientific failure, but rather one of structural disincentives ...

... if you invent a mediocre antibiotic, sales will be low; if you develop a great antibiotic, sales will be lower.

investment than a company that develops a marginally improved oncology drug used daily for years. As one analyst has observed, if you invent a mediocre antibiotic, sales will be low; if you develop a great antibiotic, sales will be lower.[14]

The commercial consequences have been severe. A new antibiotic requires approximately $310 million in annual global revenues to be commercially sustainable.[15] In practice, most new antibiotics generate around $46 million per year.[16] The net present value of developing a new antibiotic has been calculated at negative $50 million, compared to positive $720 million to $1.15 billion for drugs in other therapeutic areas.[17] Companies have responded rationally to these incentives. In the 1990s, approximately 18 major pharmaceutical companies had active antibiotic research and development programmes.[16] Today, the number is four to six.[13]

What makes the market failure particularly acute is that it has persisted even for companies that successfully navigated the scientific and regulatory challenges of drug development. Achaogen received US Food and Drug Administration (FDA) approval for its aminoglycoside antibiotic plazomicin in June 2018, after spending several hundred million dollars on development. It filed for Chapter 11 bankruptcy less than ten months later, having recorded approximately $800,000 in sales. Melinta Therapeutics, holding four approved antibiotics, followed in December 2019.[18] These are not isolated cases: a financial analysis of FDA-approved antibiotics since 2010 found that the majority of small and medium-sized sponsors had gone bankrupt or exited at a loss.[19] The clinical pipeline reflects these incentives.

What Do You Call An Empty Pipe?

By Valeria Gigante
World Health Organisation, Switzerland

In 2017, WHO published the first edition of what has become an annual reckoning: the antibacterial pipeline report, a systematic analysis of every antibacterial agent in clinical and preclinical development, assessed against the WHO bacterial priority pathogens for which new treatments are most urgently needed. It was conceived as an accountability mechanism—a way of answering, with evidence rather than based on assumptions, whether the world's investment in antibiotic research and development (R&D) was keeping pace with the evolution of AMR. Eight years on, the answer remains uncomfortable.

The 2025 analysis identified 90 candidates in clinical development, down from 97 in 2023: a decline of nearly 10% in two years. Of those 90, only 15 meet the WHO's criteria for innovation, and just 5 are active against the critical-priority pathogens[c] associated with high morbidity and mortality. The preclinical pipeline offers some reassurance—232 candidates—but success is not promised for any of them. Despite similar robust upstream prospects over the last decade, only 17 new antibacterial agents have received marketing authorisation, and only two of those represent a genuinely new chemical class. The most recent approvals, gepotidacin and zoliflodacin, mark an important success in the AMR R&D space. But we need more, and soon.

In part, the starkness of our failure to progress direct-acting small molecules means that we need to look elsewhere for solutions. Non-traditional approaches offer cautious grounds for optimism. Fresh ideas such as bacteriophages, monoclonal antibodies, anti-virulence agents, and microbiome-modulating therapies are increasingly represented in the pipeline. Some of these approaches may offer several advantages over conventional antibiotics. For example, they may be less likely to succumb to resistance and even be able to preserve what is commonly referred to as "good bacteria"—the commensal microbiota—that broad-spectrum agents disrupt. The key challenge for these novel tools is getting them to market, since both the clinical development and deployment pathways for many remain incompletely defined. But their growing presence in the pipeline reflects a field beginning to think a little bit differently.

AI has the potential to transform many domains in the fight against AMR. It doesn't offer shortcuts around the hard problems of drug discovery and development. But it does provide a way to make better use of the resources the field has, from driving down the cost of traditional small molecule discovery to identifying new approaches and derisking the exciting frontier paradigms described above.

My scientist and engineer colleagues will address some of the technical details later in this report. What I can speak to is the goal: the pipeline report in 2030 must read very differently if we are to close the gap between the scale of the issue and the scale of the response. Otherwise, we risk losing control of the problem altogether—which would be catastrophic.

[c] Those pathogens include carbapenem-resistant Enterobacterales, third-generation cephalosporin-resistant Enterobacterales, carbapenem-resistant Acinetobacter baumannii, and rifampicin-resistant tuberculosis.

A Bug's Strategic Briefing to Humanity

By Rajaram
Laurus Labs Limited, India

Dear Humans,

Hello.

I am the microscopic organism you call a bug—the enemy you have tried for more than a century to silence with medicines. You cannot see me, but I have watched you.

I have survived your greatest breakthroughs and returned stronger each time. Not because I am clever. Because I adapt.

You call this antimicrobial resistance. I call it Tuesday. For me, it is simply a matter of survival.

You might expect me to be worried. You have built sequencers that can read my genome before breakfast. You have assays that allow you to know me better than I understand myself. You have, for the first time, the artificial intelligence-based tools that can outpace my evolution—that can find my weaknesses before I have even discovered them myself. And yet…

You still fight me in silos. Your human doctors, your veterinarians, your scientists…they sit in separate rooms, writing separate reports, attending separate conferences. Meanwhile, I move freely between hospital wards in London, farms in Vietnam, and rivers running through the pharmaceutical heartland of Hyderabad. I do not respect your organisational boundaries. I never have.

Your complacency will be your undoing. You overuse antibiotics for viral infections. You purchase them without a diagnosis. You stop treatment too early. In hospitals—the places where I thrive best—I flourish in open wounds, catheters, and the misuse of broad-spectrum drugs. Every inappropriate prescription strengthens me. Every incorrect dose grants me a mutation.

Let me make this a fair fight. You have beaten existential biological threats before, and to do so again you might reflect on what you did right. When HIV first emerged, it was devastating and untreatable. Today, it is largely controlled. Not because you eradicated it, but because you built a global strategy: academia and industry collaborating on discovery, voluntary licensing enabling affordable generics, and coordinated (philanthropic) financing at scale.

A once-fatal disease now costs less than three dollars a month to manage. That was not luck. It was coordination, aligning science, economics, and access into a single effort. And you did this all before the advent of AI, which has since supercharged your abilities.

If you truly intend to defeat me, you must treat AMR as a global security threat and collaborate across your arbitrary boundaries. You will not lose because I am strong; you will lose because you choose fragmentation over coordination, and short-term fixes over a long-term vision.

If you fail, I will continue to thrive. Silently. Relentlessly. Ruthlessly.

Your move.

The Bug

CHAPTER 2

The Artificial Intelligence Inflection Point

Against this backdrop of structural failure, a genuine technological opportunity has emerged, and the timing, for once, is fortunate. Three forces have converged to make this moment distinctive: the maturation of deep learning and transformer architectures capable of modelling extraordinarily complicated systems; the accumulation of large, structured datasets in genomics, proteomics, and clinical microbiology; and a dramatic reduction in the cost of computation that has made methods once confined to super-computers globally accessible. The result is a step-change in what is computationally possible in drug discovery, surveillance, and clinical decision-making.

The clearest illustration of that shift is AlphaFold, DeepMind's AI-based protein structure prediction system. When it achieved near-experimental accuracy at the Critical Assessment of Protein Structure Prediction (CASP) competition in 2020 and subsequently released predicted structures for virtually every known protein, it did not merely advance one subfield. It signalled that AI could solve problems in biology that had resisted decades of conventional scientific effort and opened new avenues for understanding how resistant bacteria function and how they might be targeted.[21] The same deep learning methods have since been applied directly to antibiotic discovery: seminal work at the Massachusetts Institute of Technology (MIT) screened compounds from the Broad Institute's Drug Repurposing Hub and identified halicin, a compound structurally unlike any existing antibiotic with potent broad-spectrum activity, before scaling to screen more than 107 million molecules from a commercial chemical library. It was amongst the first demonstrations of what AI-driven discovery could achieve at scale.[22]

Beyond drug discovery, AI has shown early but compelling results in resistance surveillance, clinical decision support, and rapid diagnostics, each of which is explored in depth in subsequent chapters. But the strategic context matters as much as the specific applications. AMR research is acutely under-resourced relative to its disease burden: globally, there are approximately 3,000 people in the clinical AMR research workforce. That's roughly 15 times fewer than

AMR research is acutely under-resourced relative to its disease burden.

Photos: PATH/Enoch Kavindele Jr. (2021)

We're not arguing that artificial intelligence will solve AMR.

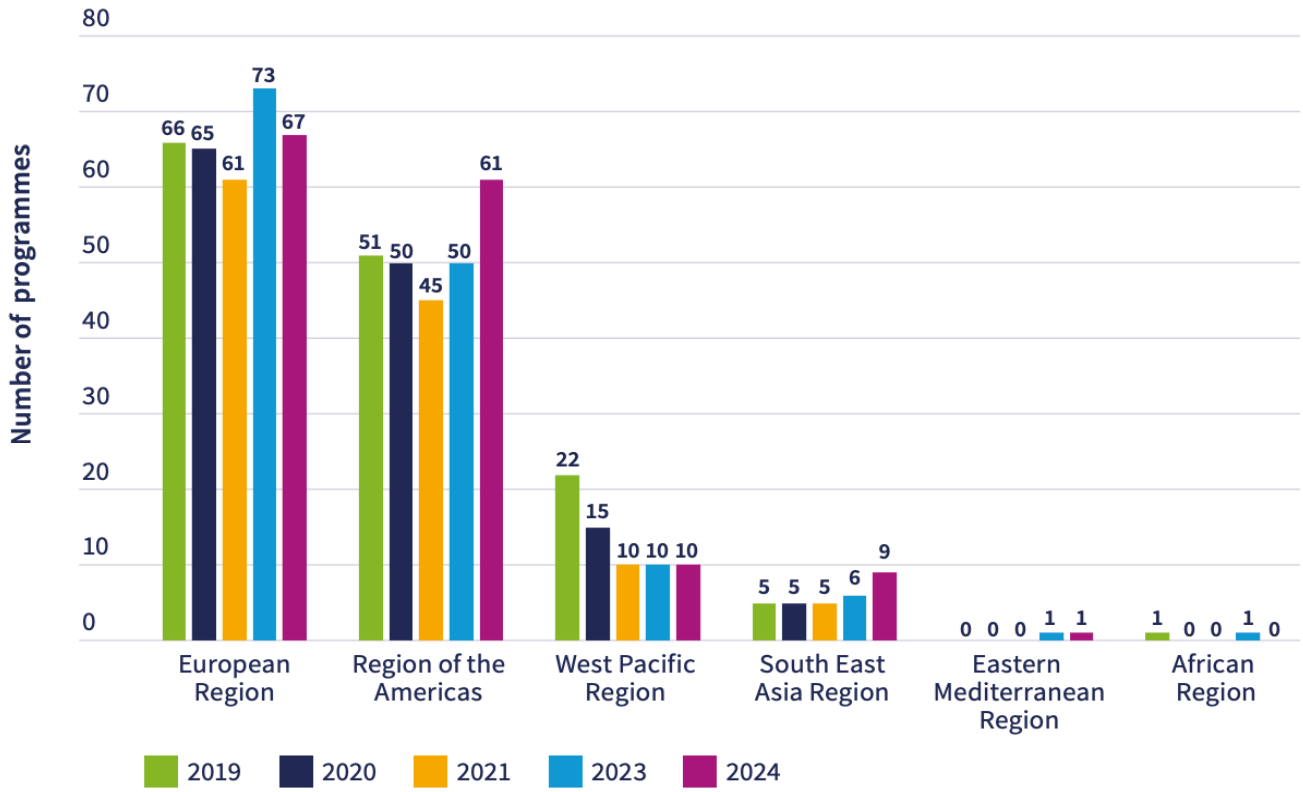
Instead, we believe AI represents the most powerful new tool available to accelerate progress

the number working on cancer.[23] That gap cannot be closed by hiring alone, certainly not quickly enough. AI will not replace the scientists, clinicians, and regulators the field needs, but it can materially extend what each of them is able to do, compressing timelines that the current crisis does not allow us to wait out.

There is, however, a structural tension that this report does not shy away from. The burden of AMR falls heaviest on low- and middle-income countries. Those settings are the most underrepresented in the datasets AI systems learn from, and the most underserved by the technology companies and academic institutions driving AI development. An AI-powered response to AMR that is designed primarily for well-resourced health systems, and deployed without deliberate attention to equity, risks widening the gap it promises to close.

We're not arguing that artificial intelligence will solve AMR. Instead, we believe AI represents the most powerful new tool available to accelerate progress across the discovery, development, and stewardship of antimicrobials, and that realising this potential will require deliberate, parallel investment in data infrastructure, equitable access to AI capabilities, and the economic reforms without which no scientific advance will be sufficient to rebuild a viable pipeline.

Figure 2. Geographical distribution of the 148 institutions with preclinical pipeline projects across the 2019-2024 analysis shown by WHO geographical regions



Source: Analysis of antibacterial agents in clinical and preclinical development: overview and analysis 2025. Geneva: World Health Organization; 2025. Licence: CC BY-NC-SA 3.0 IGO

CHAPTER 3

AI in the Lab: Discovering the Drugs We Desperately Need

A Brief Orientation: What Artificial Intelligence Actually Means

Before examining what artificial intelligence can contribute to antibiotic discovery, it is worth being precise about what the term means. *Artificial intelligence* encompasses a diverse range of methods designed to create systems capable of performing tasks that typically require human cognitive capabilities, spanning learning, reasoning, perception, and decision-making. In contemporary usage, it is often used interchangeably with *machine learning*, which is in fact a sub-field focused on the use of algorithms that learn from data rather than executing fixed, human-written rules. Said differently, a traditional computer program follows explicit instructions; a machine learning model infers its own parameters from examples, improving its predictions as it encounters more data. In this report, we largely focus on machine learning applications and explicitly do not cover 'physical AI' (i.e., the interaction of AI systems and the physical world), although it is an exciting emerging field with potential applications to AMR research.

Within machine learning, two broad families are relevant to drug development: discriminative models and generative models. Discriminative models take an input—a molecular structure, a genomic sequence, a clinical record—and return a prediction. Will this compound inhibit bacterial growth? Is this patient likely to fail first-line therapy? Generative models do something more ambitious: they learn the underlying distribution of their training data and can produce new examples from it, designing novel molecular structures with desired properties rather than merely evaluating existing ones. Large language models (LLMs), which have attracted enormous public attention since 2022, are generative models trained on text; analogous architectures trained on molecular representations, protein sequences, and genomic data are the engines of modern computational drug discovery, and were central to the Nobel Prize for Chemistry in 2024 (see Figure 3 for a description of how AlphaFold2 works).[24]

One principle should always govern our expectations of AI. It will only ever be as good as the data it learns from; a model trained on biased, incomplete, or unrepresentative data will produce biased, incomplete, or unrepresentative outputs. This caveat will resurface throughout this chapter, because the data infrastructure underpinning AI for AMR is one of the field's most significant constraints.

The Scale of the Discovery Problem

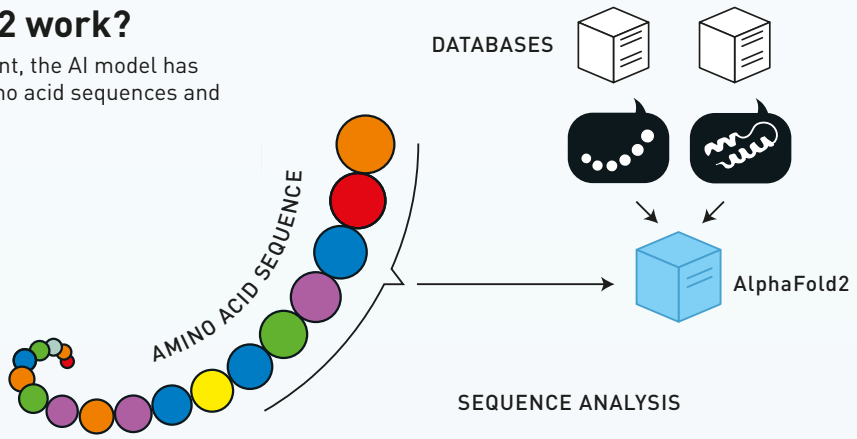
Traditional drug discovery is slow, expensive, and structurally misaligned with the economics of antibiotic development. Bringing a new drug from initial discovery to regulatory approval takes between ten and fifteen years and costs, on a capitalised basis that accounts for the cost of failures, approximately $2.6 billion.[25] Across all therapeutic areas, fewer than one in ten compounds entering Phase I clinical trials ultimately receive regulatory approval.[26] For antibiotics, the economics are even more punishing. As established in Chapter 1, new antibiotics earn

New antibiotics earn approximately $46 million per year globally, against a sustainable revenue threshold of around $300 million.[15,16] The result is a discovery process that is simultaneously too slow, too costly, and economically irrational for industry to pursue at the necessary scale.

Figure 3.

How does AlphaFold2 work?

As part of AlphaFold2's development, the AI model has been trained on all the known amino acid sequences and determined protein structures.



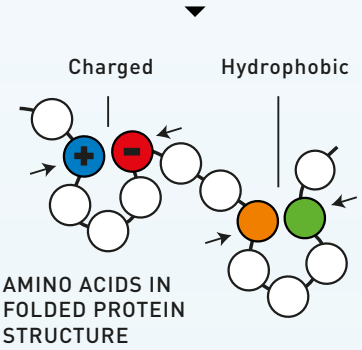
1. DATA ENTRY AND DATABASE SEARCHES

An amino acid sequence with unknown structure is fed into AlphaFold2, which searches databases for similar amino acid sequences and protein structures.

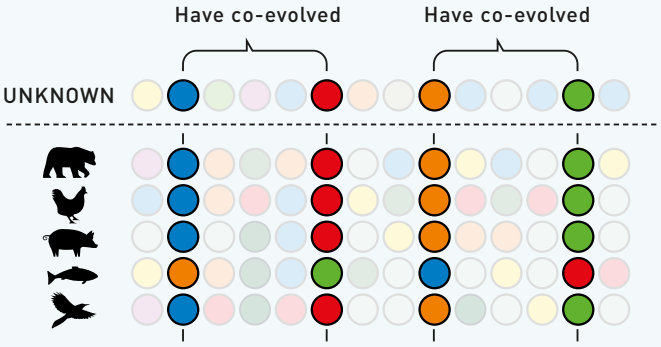
2. SEQUENCE ANALYSIS

The AI model aligns all the similar amino acid sequences – often from different species – and investigates which parts have been preserved during evolution.

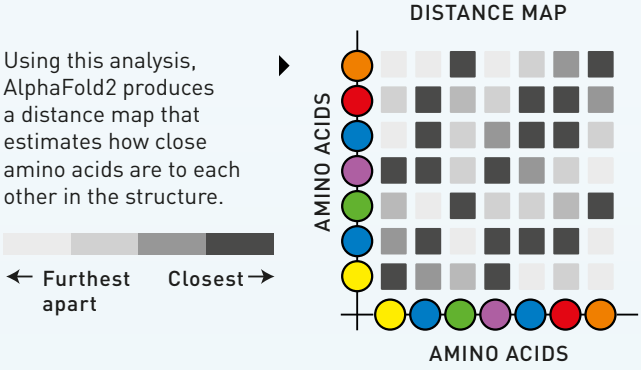
In the next step, AlphaFold2 explores which amino acids could interact with each other in the three-dimensional protein structure. Interacting amino acids co-evolve. If one is charged, the other has the opposite charge, so they are attracted to each other. If one is replaced by a water-repellent (hydrophobic) amino acid, the other also becomes hydrophobic.



SEQUENCE ANALYSIS



Using this analysis, AlphaFold2 produces a distance map that estimates how close amino acids are to each other in the structure.

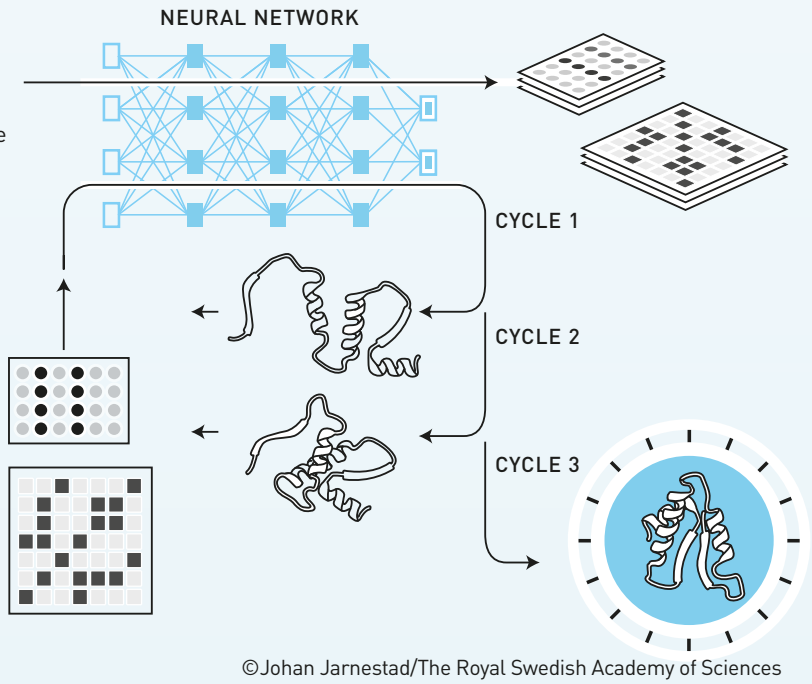


3. AI ANALYSIS

Using an iterative process, AlphaFold2 refines the sequence analysis and distance map. The AI model uses neural networks called transformers, which have a great capacity to identify important elements to focus on. Data about other protein structures – if they were found in step 1 – is also utilised.

4. HYPOTHETICAL STRUCTURE

AlphaFold2 puts together a puzzle of all the amino acids and tests pathways to produce a hypothetical protein structure. This is re-run through step 3. After three cycles, AlphaFold2 arrives at a particular structure. The AI model calculates the probability that different parts of this structure correspond to reality.



approximately $46 million per year globally, against a sustainable revenue threshold of around $300 million.[15,16] The result is a discovery process that is simultaneously too slow, too costly, and economically irrational for industry to pursue at the necessary scale.

The conventional approach relies on high-throughput screening: physically testing tens of thousands of compounds against bacterial targets in laboratory assays. This is expensive and inherently limited in its reach. Some researchers estimate there are 10^{60} compounds that may have properties that could be useful as drugs. No laboratory programme can screen more than a tiny fraction of them, which means that candidates with novel mechanisms of action—the kind most needed to defeat resistant bacteria—are routinely missed. This is the problem AI is well-positioned to address, and the following sections trace how it has begun to do so: first by searching existing chemical libraries more efficiently, then by designing entirely new molecules, and finally by repurposing compounds already known to be safe in humans.

Virtual Screening: Searching Chemical Space at Scale

The landmark demonstration of AI's potential in antibiotic discovery came in 2020, when researchers at MIT and the Broad Institute trained a graph neural network on approximately 2,500 compounds and used it to predict antibacterial activity from molecular structure alone. Applied first to the Broad Institute's Drug Repurposing Hub, the model identified halicin (a compound originally developed for diabetes), which showed broad-spectrum antibacterial activity, including against pan-resistant pathogens. Moreover, it proved resistant to the development of resistance over thirty days of serial passaging.[27] The same model was subsequently scaled to screen more than 107 million molecules from a commercial chemical library, identifying a further set of candidates with structural novelty that conventional screening would not have reached. A subsequent 2023 study from the same broader collaboration used an analogous approach to discover a new antibiotic called abaucin. This narrow-spectrum agent targets Acinetobacter baumannii, which is associated with around 400,000 deaths, but didn't harm the so-called good (commensal) bacteria that broad-spectrum antibiotics typically disrupt.[28] And finally, a more recent paper from the same laboratory introduced machine-learning approaches with a focus on explicit explainability, which means that the tool is capable of identifying which molecular substructures drive antibacterial predictions, marking a significant step towards interpretable AI in drug discovery.[29]

Photo: PATH/Gabe Bienczycki (2016)

Whilst this methodological framework established by these landmark studies—train on phenotypic data, screen computationally, validate experimentally—has become a template for the field, it's worth noting that these discoveries represent a proof of concept rather than a finished pipeline. Neither halicin nor abaucin has yet entered clinical trials. An early clinical assessment of AI-discovered drug performance, based on molecules from AI-native biotechnology companies, found Phase I trial success rates of 80% to 90%, substantially above the historic industry average of approximately 40%; however, this advantage has not yet been maintained into later-stage trials.[30] As such, whilst AI might be compressing the time needed for early-stage discovery, the clinical development problem remains unsolved.

An early clinical assessment of AI-discovered drug performance, based on molecules from AI-native biotechnology companies, found Phase I trial success rates of 80% to 90%.

Generative AI: Designing Molecules from Scratch

Virtual screening, however efficient, remains bounded by the chemical space it searches. It can identify candidates from existing libraries; it cannot invent genuinely novel structures. Generative AI removes this constraint. Rather than asking which existing molecule best fits a desired profile, generative models ask a fundamentally different question: given a desired set of properties, what molecule should we design?

Whilst Prof. David Baker (the now Nobel Laureate for his work on protein design) pioneered this approach many years ago, the first application to antibiotic design was published in 2024. The work in question identified candidates active against Neisseria gonorrhoeae and methicillin-resistant Staphylococcus aureus, both with structural novelty that would have been difficult to reach through conventional screening.[31,33] Although interest in this approach is rapidly picking up pace, there is a practical limitation. Generative models can propose vast numbers of novel structures, but many cannot be synthesised using current organic chemistry techniques. The gap between computational design and physical realisation remains a real bottleneck, and groups are now developing models that only generate potential chemical structures that can be synthesised,[33] building feasibility into the design process as a constraint rather than an afterthought. Again, there are few AI-designed products that have made it to the point of market authorisation, and thus, questions remain about how close we are to experiencing the impacts of the promised AI revolution.

Drug Repurposing: The Fastest Path to Clinical Impact

While the practical realities of using generative AI to develop marketable new drugs is still on the longer-term horizon, drug repurposing offers the fastest near-term path to new treatments. The history of medicine is rich with repurposing successes that emerged from serendipitous clinical observation. Sildenafil was developed for cardiovascular disease before its effect on erectile dysfunction became apparent; thalidomide, withdrawn for devastating effects on foetal development in the 1960s, was later found to be effective against leprosy and multiple myeloma; aspirin has accumulated many indications across decades. The challenge for AI is to engineer for this serendipity systematically and to find these cross-indication signals in structured data without having to wait for decades before they become apparent to observant clinicians and scientists.

A recent example of AI-assisted repurposing is baricitinib for COVID-19. In January 2020, scientists at BenevolentAI used a machine-learning-enhanced biomedical knowledge graph, extracting relationships between drugs, targets, genes, and diseases from more than 30 million publications, to

identify baricitinib (a rheumatoid arthritis drug) as a candidate for blocking SARS-CoV-2 cellular entry. The hypothesis was published in *The Lancet* in February 2020.[34] The subsequent Adaptive COVID-19 Treatment Trial 2 (ACTT-2) confirmed it reduced time to recovery and mortality in hospitalised patients.[35] The drug received FDA emergency use authorisation in November 2020. This example offers a proof-of-concept for AMR research. But it's even more compelling because we know that many compounds with unexplored antimicrobial activity already exist within approved drug libraries, and AI approaches that integrate molecular structure, mechanism of action, and clinical data can systematically surface them. An early screen of the Drug Repurposing Hub for synergy with existing antibiotics identified six compounds with antimicrobial activity when combined with established drugs, none of which had previously shown any antibiotic activity in isolation.[36,37]

We know that many compounds with unexplored antimicrobial activity already exist within approved drug libraries, and AI approaches that integrate molecular structure, mechanism of action, and clinical data can systematically surface them.

The ADMET Problem: Why Safety Prediction Matters as Much as Efficacy

Finding a promising compound, however, is only the first hurdle. The vast majority of drug candidates that look promising at early stages fail during clinical development, most commonly due to problems collectively described as ADMET: absorption, distribution, metabolism, excretion, and toxicity. A molecule that kills bacteria in a petri dish may be inactive at achievable concentrations in human tissue, rapidly metabolised before reaching the site of infection, or cause unacceptable organ damage. Predicting these properties computationally, before expensive animal and human studies, is therefore as important as predicting efficacy.

Several AI-based tools have been developed to address the ADMET prediction problem.[38] For example, ADMET-AI is an open-source platform that uses graph neural networks trained to contextualise predictions of a candidate molecule's clinical toxicity, solubility, bioavailability, and metabolic stability against a reference set of approved drugs drawn from DrugBank.[39] Applied to antibiotic discovery, ADMET prediction of this kind allows compounds to be prioritised not only on antibacterial potency but also on the likelihood of surviving translational (clinical) studies.

The practical implication is significant. Clinical development constitutes the majority of total drug development cost. Filtering out drugs with predictable ADMET liabilities earlier could reduce those costs. Even a modest improvement in the proportion of candidates that survive to Phase III could substantially reduce the economic barriers to antibiotic development. The limiting factor is data quality: ADMET prediction models are only as good as their training sets, and the most informative training data (i.e., the ADMET profiles of compounds that failed in development) remains largely locked in proprietary pharmaceutical databases. We return to this issue later in Chapter 4.

Bacterial Foundation Models: The Pan-Genomic Horizon

Better prediction of individual compound properties is one piece of the puzzle. The longer-horizon ambition is to build foundational AI models of bacterial biology itself, i.e., systems trained on millions of bacterial genomes that learn the underlying grammar of microbial evolution, resistance, and pathogenicity. BacFormer, a transformer-based protein language model developed at Cambridge and published as a preprint in 2025, represents one such effort. Trained on approximately 1.3 million metagenome-assembled genomes spanning more than 25,000 bacterial

species across 70 ecological environments, it predicts antibiotic resistance gene classes with greater than 93% accuracy and shows strong performance in protein–protein interaction prediction and pathogen classification.[40] The downstream applications of such a model are substantial. A system that has learned the evolutionary logic of bacterial genomes well enough to predict functional properties from sequence data alone could, in principle, anticipate which mutations are likely to confer resistance to a new drug before the drug is deployed, allowing antibiotic designers to steer away from mechanisms to which resistance is evolutionarily accessible.

An Honest Assessment: Progress Without Complacency

AI has unambiguously changed what is possible in antibiotic discovery. AI can screen more than 100 million molecules in hours. It can identify structurally novel candidates with unprecedented mechanisms of action. It can design entirely new chemical entities. All of these represent genuine scientific advances.

But the honest account must acknowledge what has not yet happened. No AI-discovered antibiotic has entered clinical trials. The 'valley of death' between computational identification and clinical development remains as deep for AI-assisted candidates as for traditionally discovered ones. Generative design, bacterial foundation models, combination therapy prediction—the most ambitious applications in this chapter—are still maturing, and each has critical dependencies on data infrastructure investments that have not yet been made at the required scale.

That is the central tension. The constraint on AI's impact in AMR is not, fundamentally, a lack of algorithmic sophistication. It is the same foundational deficit that existed before AI arrived: sparse, unrepresentative data; fragmented surveillance systems; and chronic underinvestment in the scientific commons.

AI has changed the cost-benefit calculus for researchers, drugmakers, governments, and philanthropic funders positioned to solve those problems. Meanwhile, the opportunity cost of leaving them unsolved continues to grow.

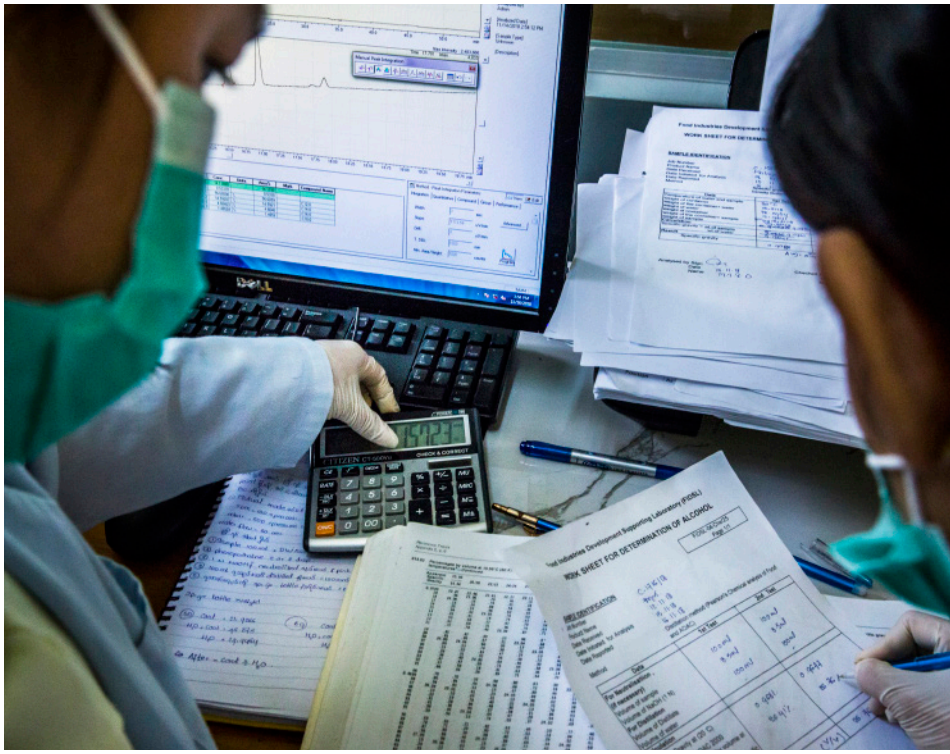
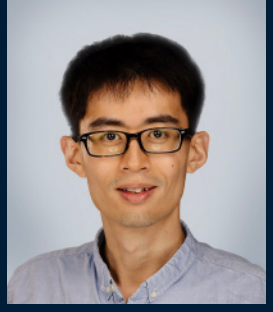


Photo: PATH/Minzayar (2018)

AI Co-Scientists: A New Paradigm ...

By Wengong Jin
Northeastern University, USA



Recent advances in large language models and multimodal AI have given rise to a genuinely new concept: the AI "co-scientist." This isn't just a faster search engine or a smarter autocomplete. These systems can reason across domains, synthesise knowledge from disparate sources, generate hypotheses, and engage with researchers in iterative, goal-directed dialogue. Crucially, they can be coupled with specialist tools, such as protein structure predictors (like AlphaFold), molecular screening platforms, and genomic databases, allowing a single system to move fluidly between literature synthesis, experimental hypothesis generation, and quantitative analysis. The result is something closer to a collaborative research partner than a conventional software tool.

What makes this paradigm potentially transformative is not just speed, but the nature of the reasoning involved. Human scientists are creative, but they are also constrained: by the literature they have read, the hypotheses that seem plausible given existing frameworks, and the cognitive cost of holding many variables in mind simultaneously. An AI co-scientist is not constrained in the same ways. It can explore combinatorial hypotheses that would be practically unreachable for a human team, propose mechanisms that cut across disciplinary boundaries, and revisit assumptions that the field has long treated as settled.

That said, it is important to be honest about where the evidence currently stands. The concept of AI co-scientists is new, and systematic evidence that these tools improve the quality and speed of scientific discovery—rather than merely the volume of outputs—remains limited. Still, early results are encouraging. AI-assisted hypothesis generation has shown promise in target identification and drug repurposing, and multimodal agents have demonstrated an ability to design and interpret experiments in constrained settings.

But the field is nascent, and the gap between impressive demonstrations and reliable, reproducible scientific impact has not yet been fully closed. These are tools that warrant serious investment and serious scrutiny in equal measure.

The AMR crisis is thus an ideal testing ground for building that evidence base for two key reasons. It is complex, data-intensive, and spans multiple layers of biology and public health—precisely the kind of problem where the integrative capacity of an AI co-scientist is most valuable. But it is also a field defined by a striking structural fragility: ninety percent of the companies in the preclinical pipeline employ fewer than 50 people. The academic community of AMR researchers is similarly a fraction of other life sciences fields. These are small teams, working on hard problems, with limited capacity to absorb the kind of broad, cross-disciplinary synthesis that effective AMR research demands. An AI co-scientist might not solve the funding problem, but it can meaningfully extend what a small, expert team is able to do.

The question in the short term is where to focus.

AMR is a sprawling topic, and we will have only so many chances to demonstrate the transformative impact of AI co-scientists. So where to start?

… For AMR-focused Combination Vaccines?

By Melanie Saville
PATH, USA

Prevention is more sustainable than treatment. Vaccines, especially, have a dual benefit. They reduce antibiotic use because they prevent people from getting sick in the first place, but also, less antibiotic use means pathogens aren't getting as many opportunities to develop resistance against them. For example, the introduction of pneumococcal conjugate vaccines has not only reduced disease but demonstrably decreased both antibiotic use and antibiotic-resistant strains of *Streptococcus pneumoniae*.

Vaccines work against AMR. The challenge is doing more of it.

Combination vaccines (i.e., single products that protect against multiple pathogens) offer an especially compelling example of how to innovate around the market challenges that have undermined the development of new and novel antibiotics. They simplify immunisation schedules, reduce delivery costs, improve coverage, and concentrate health impact, making them especially valuable in the low- and middle-income country settings where disease burden is often highest, and healthcare systems are most stretched. Most critically, they are given to everyone at risk, rather than waiting around to see how many thousands become unwell and then dealing with the consequences.

Unfortunately, despite the potentially transformative impact of combination vaccines for AMR (not only for disease control but also for how they reimagine the economics of the issue), the obstacles to progress are profound. Developing a vaccine against a single pathogen is already prohibitively expensive – a Phase III trial for the M72 tuberculosis vaccine candidate is estimated to require $550 million. Several critical AMR-associated pathogens, including *Staphylococcus aureus*, have repeatedly defeated late-stage vaccine candidates, reflecting fundamental gaps in our understanding of protective immunity. And even when we have effective vaccines for individual diseases, combining them (across different platforms) introduces further complexity: antigen interference, formulation instability, and manufacturing challenges can each derail development.

This is precisely where AI can help. AI-based protein modelling and epitope optimisation tools could enable rational immunogen design at a speed and scale previously impossible. These tools could unshackle us from the assumptions that we must use pre-existing vaccines and instead allow us to create entirely new subunits with less cross-reactivity. An AI co-scientist approach could generate and evaluate hypotheses about protection against pathogens that have so far resisted conventional vaccine development. If successful, that could allow us to accelerate development and rapidly test efficacy across different combinations.

In essence, AI affords us an entirely new way to bring together optimised antigen design, advanced vaccine platforms, and AI-driven development pathways, from preclinical proof of concept through to early clinical validation, unconstrained by the limits of human creativity.

If successful, this approach could make combination vaccines a meaningful and cost-effective pillar of the global response to AMR.

CHAPTER 4

AI in the Clinic: Smarter Use of the Drugs We Have

Stewardship as the First Line of Defence

Discovery and stewardship are complementary imperatives in the fight against AMR. The previous chapter examined the long pipeline from laboratory to approved drug; this chapter addresses what happens in the clinic before, during, and after a prescription is written. Ensuring that the antibiotics already available are prescribed to the right patient, for the right pathogen, at the right dose, for the right duration is not a second-order concern. Every antibiotic course places evolutionary pressure on bacteria, accelerating the selection of drug-resistant variants. Inappropriate use not only fails the individual patient; it degrades the effectiveness of the drug for everyone who will need it in future.

The scale of inappropriate prescribing is substantial. In the United States, the CDC estimates that approximately 28% of outpatient antibiotic prescriptions are unnecessary—predominantly for viral respiratory infections for which antibiotics have no benefit.[41] In hospital settings, studies consistently find that 30% to 50% of antibiotic use is suboptimal: the wrong drug, the wrong dose, the wrong duration, or prescribed for an infection that did not require antibiotics at all.[42] In low- and middle-income countries, where over-the-counter antibiotic sales are common and diagnostic testing is limited, the proportion is likely higher. The consequences flow in two directions simultaneously: patients receive treatments that harm them through side effects and microbiome disruption, and the resistance reservoir expands in ways that will harm future patients.

The scale of inappropriate prescribing is substantial.

In hospital settings, studies consistently find that 30% to 50% of antibiotic use is suboptimal: the wrong drug, the wrong dose, the wrong duration, or prescribed for an infection that did not require antibiotics at all.[42]

There is an understandable default, described further below, to prescribe broad-spectrum antibiotics when the cause of an infection cannot be identified quickly (or at all). This broad-spectrum prescribing is a key driver of AMR but can also be a life-saving intervention. Antimicrobial stewardship programmes exist to balance this tension, using prescribing guidelines, prospective audit and feedback, prior authorisation requirements, and education to ensure that the most appropriate antibiotic is given. These interventions demonstrably reduce inappropriate prescribing and increase the appropriate use of narrower-spectrum drugs. But they are resource-intensive, typically require specialist infectious disease physicians, and rarely operate at the speed or scale the problem demands. Artificial intelligence offers the prospect of embedding stewardship intelligence directly into clinical workflows, making high-quality prescribing guidance available continuously and at scale, without requiring a specialist in every facility.

The sections that follow trace how this is being achieved: first by speeding up lab testing processes, so clinicians can more quickly identify the cause of an infection; then by supporting clinical decisions at the point of care; then by personalising prescribing guidance to individual patients; and finally by transforming surveillance from a retrospective exercise into a continuous early-warning system.

The Diagnostic Gap: The Root Cause of Inappropriate Prescribing

At the heart of inappropriate broad-spectrum prescribing is a diagnostic failure. To identify the cause of an infection, clinicians use standardised culture-based antimicrobial susceptibility testing. This involves growing a pathogen from a patient sample and observing which antibiotics inhibit its growth. It takes between 24 and 72 hours to return a result. In a patient with serious infection, this delay is clinically untenable: treatment must begin immediately. Clinicians therefore prescribe empirically, choosing a broad-spectrum antibiotic likely to cover the most probable pathogens based on local resistance patterns and clinical judgement. This so-called empirical approach is rational under uncertainty, but it is the principal driver of excessive broad-spectrum antibiotic use and the selection pressure for resistance that accompanies it.

Compressing the time to susceptibility results is therefore one of the most impactful technical interventions available. If a clinician treating a patient with sepsis could know within hours, rather than days, not only the identity of the causative organism but its precise resistance profile, they could prescribe a targeted, narrow-spectrum antibiotic from the outset. This would reduce inappropriate prescribing, improve patient outcomes, and preserve the effectiveness of agents reserved for last resort.

Several AI-enabled approaches are achieving this compression. For example, nanomotion-based antimicrobial susceptibility testing measures the minute vibrations of bacteria attached to a cantilever as they respond to antibiotic exposure and uses machine learning to classify the results. It delivers susceptibility data within two to four hours, with accuracies of 89.5% to 98.9% in published validation studies.[43] Researchers at the University of Oxford have used deep learning and fluorescence microscopy images of bacteria under antibiotic exposure to infer resistance patterns in approximately 30 minutes.[44] Direct-from-sample whole-genome sequencing, which bypasses the lab culture step entirely and predicts resistance from genomic data, is moving from specialist reference laboratories into broader clinical deployment, enabled in part by machine learning tools that rapidly interpret genetic sequences.[45] However, these approaches are still maturing, and even where potentially ready for deployment, there are challenges that will limit access in the contexts where the burden is greatest. For example, nanomotion and imaging-based approaches still require prospective clinical validation across diverse patient populations and pathogen types. Direct-from-sample genomic sequencing remains expensive and computationally intensive. But the direction of travel is clear, and the enabling role of AI is central. It is machine learning that converts raw physical or biological signals into actionable clinical decisions—fast.

Clinical Decision Support: Prescribing Guidance at the Point of Care

Beyond accelerating diagnosis procedures, AI can support clinical decision-making directly. For example, it can embed prescribing guidance into the electronic health records that clinicians already use, in real time, without interrupting care. A systematic review of AI and machine learning tools for guiding antimicrobial therapy, covering 23 studies from 18 countries, found consistent evidence of improvements in prescribing appropriateness across clinical settings.[46] The review also identified infectious disease specialist shortages as a structural driver of AI adoption: fewer than half of US antibiotic stewardship programmes have an infectious disease–trained physician leader, leaving a gap that AI tools are increasingly being asked to fill.

Whilst rule-based systems have existed for decades, the advent of generative AI has brought new opportunities. Rather than be limited by rigidity, now systems geared towards improving the quality of care can embed stewardship principles as one of many types of output they can provide. The experience of Penda Health in Nairobi illustrates both the potential of this approach and the importance of implementation design in determining whether it is realised (See Panel 1).

PANEL 1

AI-Powered Clinical Decision Support at Penda Health, Nairobi

Penda Health operates a network of primary care clinics serving half a million patient visits per year in Nairobi, Kenya. In partnership with OpenAI, Penda deployed a large language model (LLM)–based clinical decision support tool embedded directly in its electronic medical record system. The design philosophy was deliberately unobtrusive: rather than requiring clinicians to actively request feedback, the tool ran passively in the background during every consultation, alerting clinicians only when a potential error was identified, through a traffic-light interface that preserved clinical autonomy. This design choice was informed by evaluation of an earlier version of the tool which required clinicians to actively request feedback; it achieved only modest uptake (rising from 4% to 47% over eight months).[47]

The subsequent quality improvement study covering 39,849 patient visits across fifteen clinics found that clinicians with access to the tool made 16% fewer diagnostic errors and 13% fewer treatment errors than those without it.[48] Critically, the follow-up randomised controlled trial provided the direct antibiotic-specific insight. Those decreases were good for patients and for the broader AMR issue, plus they reduced costs for patients. Not only did antibiotics account for the largest average share of spending, antibiotic-related costs were significantly lower in the intervention arm (mean difference US$0.15 lower).[49]

The Penda experience demonstrates three things of broader relevance: that LLM-based decision support can be deployed effectively in a resource-limited setting; that local co-design and workflow integration are prerequisites for adoption; and that stewardship need not be a siloed intervention, but can reduce unnecessary antibiotic prescriptions by integrating it into a broader decision-support solution.

Photo: PATH/Erick Muriuki (2024)

Personalised Antibiograms: From Population Guidelines to Patient-Specific Predictions

The standard tool for guiding empiric antibiotic prescribing in hospitals is the institutional antibiogram: a summary table showing the proportion of local bacterial isolates susceptible to each antibiotic, updated quarterly or annually. Antibiograms are valuable but fundamentally limited. They aggregate resistance patterns across heterogeneous patient populations but do not account for individual patient risk factors. They are also static, even though resistance patterns can shift substantially in just a few months.

Machine learning approaches can move beyond institutional antibiograms to personalised predictions, using a patient's clinical history, prior culture results, antibiotic exposures, comorbidities, and demographic characteristics to predict the likelihood that a given pathogen in that specific patient will be susceptible to a given antibiotic. Kanjilal and colleagues at Harvard demonstrated this for uncomplicated urinary tract infection, training machine learning models on 13,000 patient records and showing that a personalised decision algorithm achieved a 67% reduction in the need to resort to second-line antibiotics relative to clinicians, while simultaneously reducing treatment failure by 18%.[50]

Machine learning can use a patient's clinical history, antibiotic exposures, comorbidities and other factors to predict the likelihood that a particular antibiotic will work.

A later study trained an AI model on more than 8,300 infections from Stanford emergency departments and more than 15,800 urinary tract infections from Boston teaching hospitals. The study demonstrated that the AI model was able to better personalise antibiotic selection—that is, recommend an antibiotic better matched with the patient's clinical history and other individualized factors. That helped reduce the use of glycopeptides, which carry significant toxicity risks (while simultaneously increasing the likelihood of AMR) by 69%.[51]

The unnecessary use of broad-spectrum antibiotics is not merely inefficient; it is a primary driver of resistance emergence. Every course of vancomycin or carbapenem prescribed empirically to a patient who could have been treated adequately with a narrow-spectrum agent accelerates the development and spread of resistance to those agents, degrading their effectiveness for future patients with genuinely resistant infections. Personalised antibiograms represent a clinically deployable application of AI that can reduce this collateral damage without requiring any new drugs to be discovered or approved.

Surveillance as Stewardship: From Retrospective Reports to Early Warning

Effective stewardship requires not only knowing how to treat current patients but also anticipating how resistance is evolving. Static, retrospective antibiograms describe where resistance was, not where it is going. Transforming surveillance from a descriptive to a predictive enterprise, as described above, is one of the most important near-term applications of AI in clinical AMR management. But, what's often missed is that the clinic is not the only source of relevant information.

Wastewater contains genetic material from all organisms excreted by a community, and metagenomic analysis can detect resistance genes circulating in a population before they become apparent in clinical samples. This was demonstrated clearly during COVID-19, when wastewater surveillance of SARS-CoV-2 provided community-level infection data days to weeks

ahead of individual clinical testing. The same principle applies to AMR: resistance genes detected in wastewater reflect not only clinical cases but also asymptomatic carriage and environmental reservoirs, providing a more comprehensive picture of resistance burden than hospital microbiology data alone.[52] For example, studies in several European cities have demonstrated that carbapenemase genes can be detected in wastewater before they appear in clinical microbiology records, establishing in principle the early-warning potential of environmental surveillance.[53]

The analytical challenge here is significant, and AI has the potential to play a critical role: wastewater metagenomics generates vast quantities of sequencing data that must be processed to identify relevant resistance genes, distinguish signal from noise, correct for dilution and flow rate, and produce actionable outputs in near-real time. Machine learning approaches can automate gene identification, predict clinical relevance from environmental signals, and build the spatial models needed to link wastewater data to specific catchment populations. The vision these methods collectively represent is stewardship as a continuous learning system: from static antibiograms updated quarterly to dynamic risk profiles updated in real time, and from hospital-level reporting to community and environmental surveillance that closes the One Health loop (see page 6). Achieving this vision will require addressing long-standing interoperability challenges, a place where AI systems can help,[54] and operating on data in ways that preserve privacy and sovereignty. A recent perspective in Nature Medicine highlights recent technological advances in federated learning that could enable such a system.[55]

The Stewardship Business Model: Who Pays, and Why It Matters

The clinical case for AI-enabled stewardship is compelling. The economic case is in many respects even stronger: reducing inappropriate antibiotic prescribing lowers costs, reduces adverse drug events, decreases lengths of hospital stay, and avoids the considerably larger expenditure associated with treating drug-resistant infections when they do occur. A US analysis of multidrug-resistant bacterial infections in hospitalised patients estimated annual direct inpatient costs to the healthcare system of $4.6 billion.[56] The cost savings achievable from reduced inappropriate prescribing are plausibly significant fractions of this figure.

Despite this, stewardship tools have historically struggled to attract sustained funding and deployment. The fundamental problem is one of misaligned incentives: the costs of deploying a stewardship tool fall on one institution or payer, while the benefits (i.e., reduced resistance, preserved antibiotic effectiveness, averted future cases), accrue to the health system as a whole, including to future patients and other countries. This is a classic ‘free rider’ problem, and it explains why stewardship is chronically underfunded relative to its societal value.

Several approaches can realign these incentives. Framing stewardship performance as a quality indicator, alongside surgical complication rates or readmission rates, creates reputational and reimbursement incentives for hospitals to invest in stewardship infrastructure. The UK National Health Service has developed a digital stewardship vision that explicitly positions AI and decision support tools as components of quality improvement. In some settings, large integrated health systems have calculated that direct cost savings from reduced prescribing and shorter hospital stays are sufficient to justify investment in AI stewardship tools without external subsidy. The challenge is to extend this calculation to settings where payer and provider are separated, where margins are thin, and where the cost-saving case is harder to monetise locally even when it is nationally compelling.

In essence, the primary hurdle to better stewardship using AI is not primarily technical: it is organisational, economic, and political. It requires payers and health systems to recognise stewardship as an investment rather than a cost. Just as we’ve struggled to get pre-AI stewardship interventions funded at national scale, we will struggle to roll out AI-based solutions until we can more effectively align incentives.

Bridging the Divide: From 'Bedside to Bench'

By Tavpritesh Sethi,
Indraprastha Institute of Information Technology Delhi, India

The antibiotic pipeline is failing. But one of the reasons it is failing is rarely named directly: the discovery process is almost entirely disconnected from clinical reality.

Drug programmes optimise around preclinical targets and in vitro activity, while hospitals accumulate daily evidence on where antibiotics actually fail in real patients. These two worlds rarely speak to each other. Bridging this gap is not a nice-to-have. It is a precondition for a sustainable antimicrobial innovation ecosystem, and AI is the tool that makes it achievable.

Closing the bedside-to-bench loop means routing clinical intelligence embedded in electronic health records and microbiology databases back into discovery decisions. It means using the unexpected pattern of real-life outcomes that show a compound used for one indication may work for another, to direct computational screening towards the most clinically validated targets. It means reimagining dynamic surveillance tools like AMROrbit[d], which track resistance trajectories from routine microbiology data, as early-warning systems to direct discovery. Signals that a particular drug-bug combination is spiralling towards clinical crisis are also signals that the pipeline needs to prioritise agents active against that combination. In this framing, every clinical insight is also a drug discovery input. An ecosystem that treats these as separate activities is discarding some of its most valuable intelligence. Critically, none of this requires new science. It requires the infrastructure and incentives to make clinical data computable and accessible to those who need it.

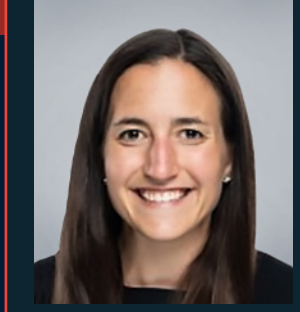
Three things are needed to bridge the gap. Making clinical data computable and high-quality does not mean sharing raw patient records; it means translating clinical experience into actionable hypotheses that drug developers can use. Academic health systems are best placed to do this translation. Second, the incentives must be aligned. Pharmaceutical companies do not need data dumps; they need insights. We can achieve this by creating mechanisms via consortia that share pre-competitive data and leverage federated and privacy-preserving architectures. Embedding all of this within public-private partnerships which provide a clear route to financing means clinical intelligence will flow to those who can act on it and create the conditions for them to act. Finally, we must think of AI as connective tissue. Models that meaningfully integrate data from electronic health records, microbiology records, prescribing histories, and genomic surveillance are best suited to surface repurposing signals, identify emerging resistance threats, and generate the clinically grounded hypotheses that purely *in vitro* drug discovery cannot produce. This kind of complex analysis is what machine learning does best.

The antibiotic crisis is not only a failure of innovation. It is a failure of integration.

If we invest wisely, there is a realistic path from bedside to bench, and to building an antibiotic pipeline that learns from clinical reality rather than running parallel to it.

[d] AMROrbit is an open-source tool developed by the Indraprastha Institute of Information Technology Delhi. It uses routine microbiology data to assess whether resistance across drug-bug combinations are increasing, stable, or declining, and at what speed.

The Low- and Middle-Income Country Imperative



By Jessica Lee,
Verily Health Sciences, USA

Approximately 80% of deaths attributable to antimicrobial resistance occur in low- and middle-income countries.

Yet virtually all of the foundational AI models, training datasets, and deployment infrastructure being built to address AMR have been developed in high-income settings. This is not merely an equity concern, though it is certainly that. It is a scientific problem, and until the field confronts it directly, AI will be able to solve only a small subset of the AMR crisis.

For example, most AI stewardship tools assume access to laboratory-based susceptibility testing. They assume electronic health records that capture diagnoses, prior prescriptions, and patient histories in structured, machine-readable form. They assume hospital-based stewardship programmes with dedicated pharmacists or infectious disease specialists to interpret and act on recommendations. In most low- and middle-income country health clinics or hospitals, none of these assumptions is reliable. Clinical data is often recorded on paper. Susceptibility testing is unavailable for most patients. The proportion of antibiotics dispensed in community pharmacies, often without prescription, and for viral illnesses, can dwarf formal prescribing. In short, we must move beyond thinking of the hospital as the primary source of surveillance data and point of deployment. AI tools designed exclusively for computerized hospital workflows will reach a fraction of the antibiotic use that drives resistance in low- and middle-income countries.

But the problem isn't simply infrastructure and human resources. A model trained on resistance patterns from Boston or London will reflect the epidemiology of those settings. In low- and middle-income countries, the dominant pathogens, resistance mechanisms, and patient risk profiles may differ substantially. So, when the Boston-trained model is applied, say, in Nigeria, it may recommend an antibiotic to which 70% of local organisms are resistant—that is, it won't work on most infections and will wind up increasing antibiotic resistance.

Local datasets are not merely a nice enhancement to global models; they are a prerequisite for safe-to-deploy models.

That said, we're not going to find the funding to build beautifully curated datasets overnight. We need to be inventive. Wastewater-based epidemiology—in other words, metagenomic analysis of sewage for resistance genes—offers a practical alternative to clinical surveillance where laboratory infrastructure is thin. It captures early indications of community-level resistance, including from agricultural and environmental sources, without requiring individual patient testing. Wastewater data is not equivalent to clinical susceptibility data, but it is real, scalable, and can be collected where individual-level data cannot.

The burden of AMR falls hardest on the countries least represented in the rooms where AI tools are designed. Changing this requires more than good intentions; it requires investment in locally relevant data streams that fill gaps in sustainable ways, and a discipline among funders and developers to ask, at every stage, whether a tool is being built for the settings that need it most or for the settings where data are more abundant. AI can accelerate AMR stewardship in low- and middle-income countries—but only if we design it to do so.

CHAPTER 5

The Substrates of Success

The previous two chapters documented what AI can do for antibiotic discovery and clinical stewardship when the conditions are right. This chapter examines the conditions themselves. The gap between what AI demonstrably can achieve and what it currently contributes to the AMR response is not primarily a shortcoming of the available methods. Progress is constrained by a set of enabling conditions that either have not been adequately built or have not been effectively maintained—the substrates without which no AI system can be built, validated, or safely deployed. Closing that gap is the highest-leverage collective investment the field can make.

Progress is constrained by a set of enabling conditions that either have not been adequately built or have not been effectively maintained.

Closing that gap is the highest-leverage collective investment the field can make.

Funding and Incentive Alignment

Chapter 1 set out in detail why the market for antibiotics is structurally broken, and why the standard instruments of pharmaceutical investment have proven insufficient to sustain a viable pipeline. The stewardship paradox, the low net present value of antibiotic development, and the wave of developer bankruptcies through the 2010s are all well-documented failures. At the same time, the economic arguments for intervention through pull incentives, subscription models, and advance market commitments have now been rehearsed thoroughly by all advocates. And in all fairness, progress is being made: the UK and Sweden have piloted subscription-based reimbursement; the PASTEUR Act has been reintroduced in the US Congress; the EU pharmaceutical reform package includes pull incentive provisions. Whether that legislative intent will transform the funding realities of the AMR space remains to be seen.

The role of AI is not to resolve this underlying market failure, although it does change some of the economic parameters in ways worth recognising. AI-assisted discovery reduces the cost and time of early-stage candidate identification. Better ways to predict how a new drug will behave not just in a petri dish (e.g., ADMET predictions), but in humans, reduces late-stage attrition. Improved clinical trial design reduces the cost of demonstrating efficacy. Each of these effects, if realised at scale, modestly improves the economics of antibiotic development. The combination of AI-enabled cost reduction with well-designed pull incentives could be sufficient to make the pipeline commercially viable in a way that neither has been able to achieve independently.

The deeper funding challenge surrounding AI-related innovations for AMR is the institutional architecture through which it is delivered. Academic grant cycles are poorly matched to the pace of AI development. Philanthropic funding tends toward projects rather than infrastructure. Commercial investment follows revenue signals that the antibiotic market does not generate. The result is a set of capable individual programmes—impressive demonstrations of what is possible in the right conditions—that have not been assembled into a sustained, coordinated capability.

What the field requires is something closer to the institutional logic of the Coalition for Epidemic Preparedness Innovations (CEPI) or Gavi: a body with a clear mandate, sustained multi-donor funding, and the authority to act across the pipeline rather than within individual project boundaries. Panel 2 examines what this might look like in practice.

PANEL 2

Towards Sustained Infrastructure for Using AI in the Fight Against AMR

The question of whether AMR needs its own version of CEPI or Gavi surfaces regularly in global health discussions, and for understandable reasons. Both institutions were founded to solve problems that markets and fragmented public funding had demonstrably failed to solve, and both have produced returns far exceeding their investment.

CEPI, established in 2017 in the wake of the West Africa Ebola outbreak, pools funding from governments and philanthropies into a single body with a mandate to accelerate vaccine development for epidemic threats and ensure the resulting products are accessible globally. Its pre-positioned investments in mRNA and other platform technologies proved decisive during COVID-19: Moderna's vaccine reached Phase I trial within 66 days of the viral sequence being shared. Gavi, founded in 2000, used mechanisms including pooled procurement and advance market commitments to create the market conditions under which manufacturers would invest in vaccines for diseases primarily affecting low-income countries, and in doing so has helped vaccines reach more than a billion children and averted more than 20 million deaths.

The structural logic of both institutions is directly relevant to AMR. The CEPI model is relevant to the innovation problem: coordinating early-stage research and development (R&D) investment, managing pre-competitive data infrastructure, and ensuring that the products developed through supported programmes are accessible in the settings where burden is highest. The Gavi model is relevant to the access and deployment problem: using pooled procurement and advance market commitments to derisk the development of AI-powered stewardship platforms and AI-developed countermeasures that manufacturers would otherwise not invest in for low- and middle-income country markets.

The pragmatic difficulty is that development assistance, which got both Gavi and CEPI off the ground, has contracted markedly in the last year. Bilateral donors that fund Gavi and CEPI are under sustained fiscal pressure. The Global Fund, Unitaid, and other multilateral financing mechanisms are actively working to sustain historical funding levels rather than expand their mandates. Pitching for an entirely new institution in this environment is not just likely to be slow, but also potentially reflects a less-than-judicious use of limited resources.

Thus, the more expedient path is to advocate for existing mechanisms to more explicitly include AMR within their scope. CEPI's mandate covers pathogens with epidemic and pandemic potential; AMR is, by any reasonable definition, a pandemic already—one that kills more than a million people annually and is projected to kill considerably more. Making AMR explicitly part of CEPI's mandate, and directing a portion of its R&D coordination and data infrastructure capacity toward the challenge of deploying AI in the fight against AMR, would require a governance decision rather than the construction of a new institution. Galvanising Gavi's interest in combination vaccines creates a direct route through which AI-assisted vaccine development for AMR-relevant pathogens could receive the kind of market-shaping support that has made other vaccine programmes viable. The Global Fund's expanding pandemic preparedness and response (PPR) mandate, adopted in response to COVID-19, creates a further opening: directing a portion of PPR investment toward AMR surveillance infrastructure and AI-enabled stewardship tools in high-burden countries is entirely consistent with the Fund's remit.

None of these is a complete solution, but, collectively, they represent the most achievable path toward the sustained, coordinated architecture that enabling AI in the fight against AMR requires.

Partnerships with Purpose: Beyond Funding to Impact

By Rebecca Cosgriff, Andrew Leach, Thorsten Forster, Sam Genway, Magdalena Koscielniak, and Ghada Zoubiane, *LifeArc, UK*






There is a version of the AMR funding problem that receives far less attention than it deserves. Discussions typically focus on familiar gaps: insufficient funding, inadequate pull incentives, and a lack of sustained commitment. These concerns are valid. But there is a parallel issue in the architecture of how funding flows, one that more money alone would not fix.

The conventional funding model separates the funder from the work itself. Grants are awarded on the strength of proposals; milestones are tracked at arm's length; and results are reported when the funding period ends. This approach works reasonably well for investigator-led science, but it is far less effective for translational research: moving a validated concept through the institutional, regulatory, and operational steps needed to reach patients. These steps rarely align with a typical grant's timeframe, and the expertise needed to navigate them is seldom found within the funding organisation.

AMR is full of promising ideas that did not fail scientifically but stalled somewhere between discovery and delivery.

We've seen AI tools that performed well in controlled settings that never found an institutional home for deployment. Some stewardship programmes ended as soon as their funding did. These outcomes arise not because funders are indifferent to impact, but because the structure of the conventional funding model does not give them the mechanisms to prevent it.

A funder that is also a doer – one with genuine scientific and operational capacity deployed alongside its capital – can change this dynamic. When the organisation providing the funding also has the expertise to identify where a programme is likely to stall, and the mandate to intervene when it does, the gap between funding and impact narrows. Risk tolerance increases. Agility increases. And accountability shifts toward whether something worked rather than whether it was delivered on schedule.

LifeArc operates from this model. As a self-funded medical research organisation, we reinvest the proceeds of successful translational projects into a portfolio that combines capital with hands-on expertise, creating ecosystems where innovation thrives and reaches patients quickly. We are not suggesting that this model should replace conventional grant-making altogether. Our point is that the AMR and AI communities would benefit from having more funders willing to operate this way, and by thinking carefully about which problems are best suited to which funding architecture. Some questions are best answered by independent investigators with unrestricted grants. Others will not be answered at all unless an organisation with both resources and expertise remains engaged until the problems are resolved.

Data and Infrastructure

Assuming the funding challenges can be resolved, one of the key priorities should be to use that financing to address the data challenges that hinder the use of AI in fighting AMR. Three distinct problems sit beneath the general observation that the field is data-poor, and each has a plausible path toward resolution.

The first problem is interoperability. The WHO Global Antimicrobial Resistance and Use Surveillance System (GLASS) is the most ambitious attempt to build a standardised global AMR dataset. It enrolled 127 countries and territories as of 2025, but only 104 provided data in the most recent reporting cycle.[57] Sub-Saharan Africa and Central Asia, which bear the highest per-capita resistance mortality, contribute the least amount of data. Even where data exists, the two dominant interpretive standards—the European Committee on Antimicrobial Susceptibility Testing (EUCAST) and the Clinical & Laboratory Standards Institute (CLSI) in the United States—apply different breakpoints to the same drug-pathogen combinations, producing categorical classifications that are not directly comparable. Ciprofloxacin resistance rates in E. coli differ by nearly 13 percentage points depending on which standard is applied.[58] For AI models trained on pooled multi-country data, this is not a methodological nuance; it produces training signals that mean systematically different things in different geographies. For years, researchers have been talking about harmonizing the data from EUCAST and CLSI, the European and US tracking systems. Completing that harmonization exercise and embedding the resulting standards in the global GLASS reporting requirements would improve the quality of every AI model trained on global surveillance data.

The second problem is data loss. The detailed profiles of failed drug candidates (i.e., records of why a compound did not progress through development) represent some of the most valuable training data available for AI-based discovery. Most of the data sits in proprietary pharmaceutical databases, and when developers collapse, as Achaogen, Melinta, Tetraphase, and Nabriva all did within a few years of each other, it is not preserved for science. Achaogen invested $784 million developing the next-gen antibiotic plazomicin and filed for bankruptcy ten months after approval.[19] Establishing a mechanism by which data from failed antibiotic programmes is transferred to a public repository as a condition of public funding would recover a significant body of knowledge that is currently being lost one bankruptcy at a time.

The third problem is pre-competitive data that exists but is not shared.[59] The field does have genuine infrastructure assets—the Comprehensive Antibiotic Resistance Database, the Bacterial and Viral Bioinformatics Resource Center, the Vivli AMR Register—but none connects genomic resistance data to clinical outcomes to pharmacological compound libraries within a common architecture. The MELLODDY project demonstrated what becomes possible when those changes occur. Ten major pharmaceutical companies, including AstraZeneca, GSK, Novartis, and Janssen, trained shared machine learning models using federated learning, pooling 2.6 billion experimental data points across 21 million molecules without any company's data ever leaving its own servers.[60] All ten partners reported improvements in predictive accuracy, with the largest gains in pharmacokinetics and safety assays—precisely the ADMET data most likely to be negative, unpublished, and otherwise inaccessible. An AMR-specific equivalent, governed through a pre-competitive consortium, would unlock knowledge currently locked in the records of companies that can no longer afford to use it.

Human Capital

Photo: PATH/Gabe Bienczycki (2016)

The funding, data, and infrastructure described in the preceding sections are necessary conditions for AI to contribute to the AMR response. They are not sufficient ones.

Delivering on the potential described in this report requires people with the skills to build the tools, the clinical expertise to evaluate and deploy them, and the scientific grounding in microbiology and infectious disease to ensure that what is built is genuinely useful.

The AMR research workforce is severely undersized relative to the problem. As Chapter 1 noted, approximately 3,000 people worldwide work in clinical AMR research, which is roughly 15 times fewer than the number working on cancer. Adding an AI layer to this landscape does not automatically solve the workforce problem. Training AI systems requires people who understand both the computational methods and the biological context well enough to know what data to use, how to interpret outputs, and when a model's predictions should be trusted or questioned. Deploying AI stewardship tools in clinical settings requires clinicians who understand the tools well enough to interrogate them and override them when appropriate. Governing the data infrastructure and navigating the regulatory processes requires people with expertise that spans public health, informatics, and policy in ways that current training programmes do not reliably produce.

In essence, there are three critical gaps which require renewed attention. The first is the shortage of researchers who are genuinely competent in both machine learning and microbiology. While collaborations between separate specialists are valuable, what we really need are individuals who can move fluently between the two. The second is the relatively paucity of AI and data science capacity embedded in the health systems of high-burden low- and middle-income countries. Again, to be clear, there is a vibrant research ecosystem, but the number of 'clinicians who code' in high-burden health care context is far below what is necessary to drive innovation as well as to ensure effective oversight of deployed systems. The third is the lack of structured career pathways that make the intersection of AMR and AI an attractive long-term commitment for early-career researchers, who currently face a landscape of short-term grants, limited institutional homes, and uncertain career progression.

Addressing these gaps requires sustained investment in training programmes that cross disciplinary boundaries, fellowship schemes that place AI researchers in low- and middle-income country health systems for extended periods, and institutional commitments that give AMR–AI researchers genuine career security. None of this is novel or difficult to design. It has simply not been prioritised by enough funders.

Once these necessary elements have been prioritised, we will start to see changes in the trajectory of the antimicrobial pipeline and the burden of AMR.

CHAPTER 6

Five Proof Points: What Success Looks Like by 2030

Strategies shift. Governments change. An unexpected pandemic emerges and the scramble to respond builds capabilities that nobody planned for. Prescribing exactly how deployment of AI in the fight against AMR should progress is therefore less useful than defining what progress looks like: the metrics against which any path, however it unfolds, can be held to account. What follows are five such tests. They are not a blueprint but rather a standard of adequacy. They are specific enough to be measured, ambitious enough to matter, and independent enough of any particular institutional arrangement that they remain valid however the landscape changes in the years ahead.

Pipeline Progress

1. **A late-stage clinical trial for an AI-discovered antibiotic:** At least one therapeutic agent discovered or designed using AI enters Phase III clinical trials against a WHO critical-priority pathogen. The candidate pipeline already contains compounds identified using AI methods; what is needed is the development funding and partnership infrastructure to advance them past the 'valley of death.'

2. **Twenty AI-assisted candidates in preclinical trials:** At least twenty new therapeutic agents discovered or designed using AI progress from bench to preclinical trials. Given the reductions in discovery and initial assessment costs that AI enables and the emerging potential of preclinical modelling, this is a realistic target if we invest in the infrastructure to support it.

Clinical Impact

3. **National-scale deployment of AI stewardship tools in ten countries:** AI-enabled clinical decision support, antibiotic resistance profiles, or dynamic surveillance dashboards are deployed at national scale and independently evaluated in health systems in at least ten countries, including at least five low- or middle-income countries. The tools to achieve this exist; deployment at scale has not yet happened.

Harmonised Infrastructure

4. **A pre-competitive AMR data consortium.** A formally established consortium for pre-competitive sharing of AMR pharmacological data is operational, with governance agreed, at least five pharmaceutical company participants, and a federated learning architecture running. The MELLODDY project (see page 32) demonstrated this is achievable. The barrier is the governance decision, not the technology.

5. **Published benchmarks for AMR AI model performance.** Standardised, publicly available benchmarks for assessing AI model performance in antimicrobial development—spanning ADMET prediction, efficacy screening, and resistance surveillance—are published and endorsed by a recognised scientific authority. These are the shared evaluation standards without which the field cannot distinguish genuine advances from optimistic, but ultimately flawed, *in silico* results. They are to using AI in the fight against AMR what ImageNet was to computer vision and the Critical Assessment of protein Structure Prediction (CASP) was to protein structure prediction: the common standard that allows the field to measure itself honestly.

The Question That Would Change Everything

By Bilal A Mateen *(PATH, UK)* and
Samuel Scarpino *(Northeastern University, USA)*

What leads to breakthrough applications of AI? The really world-changing ones? The obvious answer is data. But while necessary, data alone are not sufficient. It is the combination of the right question and the right data that matters most for AI. Consider AlphaFold: it was neither the Protein Data Bank (PDB) nor the Critical Assessment of protein Structure Prediction (CASP) alone, but the combination of the two that led to a Nobel Prize (by creating a benchmark that made genuine progress distinguishable from noise). The PDB, a now-legendary repository of high-quality, open-access data, provided the substrate, and CASP, a focused, world-class competition, forced the community to agree on what "good" looked like.

That deceptively simple decision to define a grand challenge around a question and a dataset gave purpose and inspired broad participation.

Turing to AMR, we have no shortage of hard questions, and we can curate and/or collect the data. We also have talented researchers, promising methods, and an urgent need. What we do not have is a single, well-defined, and widely endorsed pairing of question and data around which the field can organise. As a result, every group defines its own success criteria, curates their own data, trains isolated models, and publishes results that are not tied to impact. We cannot tell whether the field is advancing or merely producing papers.

One clear outcome from our convening is an urgent call for alignment around the well-defined question and necessary data. Such alignment would help funders prioritise research and application around clear measures of progress: not "we gathered some data" or "we fit a model," but rather "our work led to a better answer to the question because of X model and/or Y data." Such a pairing would give regulators a contextual anchor to understand the veracity of computational safety prediction claims. And it would attract the technology sector engagement the field has been asking for, because well-defined questions paired with the right data are how the AI community allocates its attention.

What should our question be? We debated several options among ourselves, but concluded that the best way to end this report was to empower the community to make their case for the question we should ask and the data needed. We do not see our conclusion as passing the responsibility—far from it; instead, we want to empower a community-driven approach. The result could be a single challenge or a federated set of challenges. Or we may embrace an entirely different approach, but we must align on the impactful, measurable question and the necessary data.

AMR is not protein folding, but we can learn important lessons from the latter to realise breakthroughs in the age of AI.

The algorithms are not the bottleneck. We acknowledge the need for better data, but that is not the bottleneck either. The bottleneck is clearly defining and prioritizing the questions. Define them well, and the rest will follow.

A Decision, Not A Discovery Is What We Need

By Dr. Naveen Rao,
Senior Vice President, Health, The Rockefeller Foundation

In March 1984, thirty-four people gathered in this same villa above Lake Como, among them leaders from WHO, UNICEF, the World Bank, and the Rockefeller Foundation. They represented agencies that had separately spent years failing to lift childhood immunisation above the twenty per cent threshold. They left having agreed on a single goal and a way to pursue it together. Within six years, four in five children were receiving a vaccine. The immunisation effort that began there is part of why childhood mortality has fallen by nearly sixty per cent since 1990.

That is what Bellagio is for. It removes distractions, puts people who rarely share a table together, and asks them to think beyond the walls of their own institutions. None of the participants arrived with the answer to AMR, and I know for certain that no single discipline holds it.

Once again, Bellagio has done what it has become famous for: the group left with a clear understanding that the pieces already exist between us, and it's what happens tomorrow and the day after that will determine whether we realise success, as we did over 40 years ago with immunisations.

If you've made it this far, there is a single recurrent idea in this report that you will have read many times: the problems that remain are not, at root, scientific. The market for antibiotics is broken. The data that would teach our models is sparse, unshared, and missing from the places carrying eighty per cent of the burden. The institutions that could carry this work, CEPI, Gavi, and the Global Fund, already exist and need a mandate rather than a successor.

These are failures of coordination and conviction, not of capability.

So, the question you should leave this report with is not whether AI can help. It plainly can. It is about whether you are willing to do your part in aligning the funding, the data, and the people behind it so that we can realise the potential of this technology in the fight against AMR.

It was always a decision, not a discovery, that we needed.

We hope you will make the right one.

References

1. GBD 2021 Antimicrobial Resistance Collaborators. "Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050." *The Lancet*, vol. 404, no. 10459, 2024, pp. 1199–1226.

2. Antimicrobial Resistance Collaborators. "Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis." *The Lancet*, vol. 399, no. 10325, 2022, pp. 629–655.

3. FAO, UNEP, WHO, WOAH. *One Health Joint Plan of Action (2022–2026)*. 2022. See also: World Health Organization. *Global Action Plan on Antimicrobial Resistance*. Geneva: WHO, 2015.

4. Van Boeckel, T.P., Brower, C., Gilbert, M., et al. "Global trends in antimicrobial use in food animals." *Proceedings of the National Academy of Sciences*, vol. 112, no. 18, 2015, pp. 5649–5654.

5. Tiseo, K., Huber, L., Gilbert, M., Robinson, T.P., and Van Boeckel, T.P. "Global Trends in Antimicrobial Use in Food Animals from 2017 to 2030." *Antibiotics*, vol. 9, no. 12, 2020, p. 918.

6. Tang, K.L., Caffrey, N.P., Nóbrega, D.B., et al. "Restricting the use of antibiotics in food-producing animals and its associations with antibiotic resistance in food-producing animals and human beings: a systematic review and meta-analysis." *The Lancet Planetary Health*, vol. 1, no. 8, 2017, pp. e316–e327.

7. Larsson, D.G.J., de Pedro, C., and Paxeus, N. "Effluent from drug manufactures contains extremely high levels of pharmaceuticals." *Journal of Hazardous Materials*, vol. 148, no. 3, 2007, pp. 751–755.

8. Wilkinson, J.L., Boxall, A.B.A., Kolpin, D.W., et al. "Pharmaceutical pollution of the world's rivers." *Proceedings of the National Academy of Sciences*, vol. 119, no. 8, 2022, e2113947119.

9. O'Neill, J. *Tackling Drug-Resistant Infections Globally: Final Report and Recommendations*. Review on Antimicrobial Resistance. London: Wellcome Trust & HM Government, 2016.

10. World Bank. *Drug-Resistant Infections: A Threat to Our Economic Future*. Washington, DC: World Bank Group, 2017.

11. McDonnell, A., Countryman, A., Laurence, T., et al. *Forecasting the Fallout from AMR: Economic Impacts of Antimicrobial Resistance in Humans*. EcoAMR Series. Washington, DC: Centre for Global Development, 2024.

12. Towse, A. and Silverman Bonnifield, R. "An Ambitious USG Advanced Commitment for Subscription-Based Purchasing of Novel Antimicrobials and Its Expected Return on Investment." CGD Policy Paper 277. Washington, DC: Centre for Global Development, 2022.

13. Gargate, N., Laws, M., and Rahman, K.M. "Current economic and regulatory challenges in developing antibiotics for Gram-negative bacteria." *npj Antimicrobials and Resistance*, vol. 3, 2025, p. 50.

14. Outterson, K. and Rex, J.H. "Evaluating the regulatory pathway for delinked pull incentive structures." *The Journal of Law and the Biosciences*, 2020. See also: Wellcome Trust. "There is no market for new antibiotics." *Wellcome Open Research*, 2021. PMC: PMC8237369.

15. Outterson, K. "Estimating The Appropriate Size Of Global Pull Incentives For Antibacterial Medicines." *Health Affairs*, vol. 40, no. 11, 2021, pp. 1758–1765. DOI: 10.1377/hlthaff.2021.00688.

16. Plackett, B. "Why big pharma has abandoned antibiotics." *Nature*, vol. 586, 2020, pp. S50–S52. DOI: 10.1038/d41586-020-02884-3.

17. Clancy, C.J. et al. "Buying Time: The AMR Action Fund and the State of Antibiotic Development in the United States 2020." *Clinical Infectious Diseases*, vol. 71, no. 8, 2020, pp. e541–e544. PMC: PMC7652093.

18. ACS C&EN. "Achaogen files for bankruptcy." *C&EN Global Enterprise*, 2019. DOI: 10.1021/cen-09716-buscon14. For Melinta: Fierce Biotech. "Melinta files for bankruptcy in another dark day for antibiotics." December 2019.

19. Outterson, K. and Powers, J.H. "Novel insights from financial analysis of the failure to commercialise plazomicin." *Humanities and Social Sciences Communications*, 2024.

20. World Health Organization. *Analysis of Antibacterial Agents in Clinical and Preclinical Development: Overview and Analysis 2025*. Geneva: WHO, 2025.

21. Jumper, J., Evans, R., Pritzel, A., et al. "Highly accurate protein structure prediction with AlphaFold." *Nature*, vol. 596, 2021, pp. 583–589. DOI: 10.1038/s41586-021-03819-2.

22. Stokes, J.M., Yang, K., Swanson, K., et al. "A Deep Learning Approach to Antibiotic Discovery." *Cell*, vol. 180, no. 4, 2020, pp. 688–702. DOI: 10.1016/j.cell.2020.01.021.

23. AMR Industry Alliance. *Leaving the Lab: Tracking the Decline in AMR R&D Professionals*. February 2024.

24. Press release. NobelPrize.org. Nobel Prize Outreach 2026. Sat. 4 Apr 2026. <https://www.nobelprize.org/prizes/chemistry/2024/press-release/>

25. DiMasi, J.A., Grabowski, H.G., and Hansen, R.W. "Innovation in the pharmaceutical industry: New estimates of R&D costs." *Journal of Health Economics*, vol. 47, 2016, pp. 20–33. DOI: 10.1016/j.jhealeco.2016.01.012.

26. Wilczok, M., et al. "Progress, Pitfalls, and Impact of AI-Driven Clinical Trials." *Clinical Pharmacology and Therapeutics*, 2024. DOI: 10.1002/cpt.3542.

27. Stokes, J.M., Yang, K., Swanson, K., et al. "A Deep Learning Approach to Antibiotic Discovery." *Cell*, vol. 180, no. 4, 2020, pp. 688–702. DOI: 10.1016/j.cell.2020.01.021.

28. Liu, G., Catacutan, D.B., Rathod, K., et al. "Deep learning-guided discovery of an antibiotic targeting Acinetobacter baumannii." *Nature Chemical Biology*, vol. 19, 2023, pp. 1342–1350. DOI: 10.1038/s41589-023-01349-8.

29. Wong, F., Zheng, E.J., Valeri, J.A., et al. "Discovery of a structural class of antibiotics with explainable deep learning." *Nature*, vol. 626, 2024, pp. 177–185. DOI: 10.1038/s41586-023-06887-8.

30. Jayatunga, M.K.P., Ayers, M., Bruens, L., Jayanth, D., and Meier, C. "How successful are AI-discovered drugs in clinical trials? A first analysis and emerging lessons." *Drug Discovery Today*, vol. 29, no. 6, 2024, p. 104009. DOI: 10.1016/j.drudis.2024.104009.

31. Swanson, K., Liu, G., Catacutan, D.B., et al. "Generative AI for designing and validating easily synthesizable and structurally novel antibiotics." *Nature Machine Intelligence*, vol. 6, 2024, pp. 338–353. DOI: 10.1038/s42256-024-00809-7.

32. Krishnan, Aarti, et al. "A generative deep learning approach to de novo antibiotic design." *Cell* 188.21 (2025): 5962-5979.

33. Parrot, Maud, et al. "Integrating synthetic accessibility with AI-based generative drug design." *Journal of cheminformatics* 15.1 (2023): 83.

34. Richardson, P., Griffin, I., Tucker, C., et al. "Baricitinib as potential treatment for 2019-nCoV acute respiratory disease." *The Lancet*, vol. 395, no. 10223, 2020, pp. e30–e31. DOI: 10.1016/S0140-6736(20)30304-4.

35. Kalil, Andre C., et al. "Baricitinib plus remdesivir for hospitalized adults with Covid-19." *New England Journal of Medicine* 384.9 (2021): 795-807.

36. Kulesa, Anthony, et al. "Combinatorial drug discovery in nanoliter droplets." *Proceedings of the National Academy of Sciences* 115.26 (2018): 6685-6690.

37. Corsello, S.M., Bittker, J.A., Liu, Z., et al. "The Drug Repurposing Hub: a next-generation drug library and information resource." *Nature Medicine*, vol. 23, no. 4, 2017, pp. 405–408.

38. Arnold, A., McLellan, S., and Stokes, J.M. "How AI can help us beat AMR." *npj Antimicrobials and Resistance*, vol. 3, 2025, p. 18. DOI: 10.1038/s44259-025-00085-4.

39. Swanson, K., Walther, P., Leitz, J., et al. "ADMET-AI: a machine learning ADMET platform for evaluation of large-scale chemical libraries." *Bioinformatics*, vol. 40, no. 7, 2024, btae416. DOI: 10.1093/bioinformatics/btae416.

40. Croce, G., et al. "BacFormer: a pan-genomic foundation model for bacteria." *bioRxiv preprint*, 2025.

41. Centers for Disease Control and Prevention. *Antibiotic Use in the United States, 2022 Update: Progress and Opportunities*. Atlanta, GA: CDC, 2022.

42. Magill, S.S., et al. "Multistate Point-Prevalence Survey of Health Care-Associated Infections." *New England Journal of Medicine*, vol. 370, no. 13, 2014, pp. 1198–1208. DOI: 10.1056/NEJMoa1306801.

43. Stupar, O., et al. "Nanomotion-based antimicrobial susceptibility testing." *Nature Communications*, 2024. DOI: 10.1038/s41467-024-46213-y.

44. Rosado, H., et al. "Deep learning-based antimicrobial susceptibility testing in under 30 minutes using microscopy." *Nature Communications*, 2023. DOI: 10.1038/s41467-023-41876-5.

45. Moran, R.A., et al. "Direct-from-sample whole-genome sequencing for clinical AMR prediction." *Lancet Microbe*, 2023.

46. Vayena, E., et al. "Systematic review of AI-guided antimicrobial therapy: evidence and gaps." *Lancet Digital Health*, 2024.

47. Obong'o, C., Njenga, G., Otiangala, D., et al. "A Mixed-Methods Evaluation of Clinician Experiences and Adoption Patterns of an EHR-integrated Generative AI-based Clinical Decision Support System in Kenya." medRxiv preprint 2025.08.14.25333740, 2025.

48. Korom R, Kiptinness S, Adan N, Said K, Ithuli C, Rotich O, Kimani B, King'ori I, Kamau S, Atemba E, Aden M. AI-based clinical decision support for primary care: A real-world study. arXiv preprint arXiv:2507.16947. 2025 Jul 22.

49. Agweyu A., Mwaniki P., Menon V., Korom R., Isaaka L., Wanyama C., Gill J., Kiptinness S., Adan N., Emmanuel-Fabula M., Riley R., Archer L., Denniston AK., Liu X., Mateen, BA. Generative-AI enabled clinical decision support system in primary care: a pragmatic, cluster randomized trial. Nature Medicine. 2026 (Accepted).

50. Kanjilal, S., et al. "A decision algorithm to promote outpatient antimicrobial stewardship for uncomplicated urinary tract infection." *Science Translational Medicine*, vol. 12, no. 568, 2020. DOI: 10.1126/scitranslmed.aay5067.

51. Corbin, C.K., et al. "Personalised antibiograms for machine learning driven antibiotic de-escalation." *npj Digital Medicine*, vol. 5, 2022. DOI: 10.1038/s41746-022-00638-z.

52. Maal-Bared, R., et al. "Wastewater surveillance for antimicrobial resistance genes." *Water Research*, 2023.

53. Cacace, D., et al. "Antibiotic resistance genes in European river waters." *Water Research*, vol. 130, 2018, pp. 252–260. DOI: 10.1016/j.watres.2017.11.055.

54. Kraemer, M. U. G., Tsui, J. L. H., Chang, S. Y., Lytras, S., Khurana, M. P., Vanderslott, S., ... & Bhatt, S. (2025). Artificial intelligence for modelling infectious disease epidemics. *Nature,* 638(8051), DOI: 623-635.10.1038/s41586-024-08564-w.

55. Khurana, M.P., Tsui, J.L.H., Gutierrez, B., Chopra, A., Scheidwasser, N.., Zhu, H., ... Scarpino, S.V., Bhatt, S., & Kraemer, M. U. G.& (in press). Global Approaches to Infectious Disease Surveillance and Modelling. *Nature Medicine* 10.1038/s41591-026-04351-4.

56. Nelson, R.E., et al. "National estimates of healthcare costs associated with multidrug-resistant bacterial infections." *Clinical Infectious Diseases*, vol. 72, Supplement 1, 2021, pp. S17–S26. DOI: 10.1093/cid/ciaa1011.

57. World Health Organization. *GLASS Whole-genome Sequencing for Surveillance of Antimicrobial Resistance: Pilot Study Report.* Geneva: WHO, 2025. ISBN: 978-92-4-011310-7.

58. Humphries, R.M., et al. "EUCAST and CLSI: Working Together to Harmonize Antimicrobial Breakpoints." *Journal of Clinical Microbiology*, vol. 59, no. 8, 2021. DOI: 10.1128/JCM.00013-21.

59. Barrett, Jeffrey S. "The Precompetitive Space for Drug or Vaccine Development: What Does It Look Like Now and What Could It Look Like in the Future?." *The Journal of Pediatric Pharmacology and Therapeutics* 28.5 (2023): 465-472.

60. Heyndrickx, Wouter, et al. "MELLODDY: cross-pharma federated learning at unprecedented scale unlocks benefits in QSAR without compromising proprietary information." *Journal of chemical information and modeling* 64.7 (2023): 2331-2344.

Northeastern University

PATH

The Rockefeller Foundation