12

CHAPTER

Artificial Intelligence Applied to Antimicrobial Stewardship Programs: Step-by-step Guidance and Real-Life Experience.

CHAPTER 12 ARTIFICIAL INTELLIGENCE APPLIED TO ANTIMICROBIAL STEWARDSHIP PROGRAMS: STEP-BY-STEP GUIDANCE AND REAL-LIFE EXPERIENCE.

Artificial intelligence (AI) is increasingly becoming a pivotal element in clinical decision-making, particularly in the realm of infectious diseases and antimicrobial stewardship. Recent advancements have allowed for the integration of machine learning algorithms into molecular diagnostic workflows, enabling clinicians to make more accurate, timely, and personalized therapeutic decisions. This multidisciplinary approach—bridging medicine, molecular biology, and data science—requires leadership not only in clinical practice but also in institutional transformation and health informatics. Professionals engaged in this field often hold dual expertise in both clinical microbiology and systems-level management, reflecting the complex ecosystem needed to implement AI at scale. The role of academic and private institutions, such as medical schools and diagnostic laboratories, is fundamental in supporting this transformation. These centers serve as both data-generating environments and implementation hubs for AI-based stewardship tools. Thus, the convergence of clinical, academic, and entrepreneurial efforts forms the cornerstone of next-generation infectious disease management.

Artificial intelligence is not introduced into a vacuum; it must be aligned with global health strategies to have meaningful clinical and epidemiological impact. The World Health Organization (WHO) has laid out a comprehensive blueprint to combat antimicrobial resistance (AMR), which provides an essential context for the implementation of AI in clinical practice. Among the key pillars outlined are the strengthening of surveillance systems such as GLASS (Global Antimicrobial Resistance Surveillance System), the optimization of antimicrobial use through national stewardship programs, and the promotion of infection prevention and control (IPC) protocols. Furthermore, the WHO emphasizes the importance of enforcing appropriate regulation and access to antimicrobials, raising public and professional awareness, investing in research and innovation, and adopting a One Health approach that integrates human, animal, and environmental health. AI systems, particularly those embedded within stewardship frameworks, can serve as accelerators

across all these domains by transforming raw microbiological and clinical data into actionable insights at both the patient and policy levels. **(WHO, 2025)**

The optimization of antimicrobial use, as outlined by the WHO, is not only a policy objective but also a clinically imperative one. Central to this strategy is the application of standardized tools such as the AWaRe classification, which organizes antibiotics into Access, Watch, and Reserve categories to promote responsible prescribing. AI systems have the potential to operationalize these guidelines in real time, integrating them into clinical decision-making at the bedside. The WHO's AWaRe Antibiotic Book becomes more than a reference—it becomes a programmable logic embedded into algorithms that guide therapy according to drug class, local susceptibility patterns, and patient-specific variables such as route of administration, dosage, and duration. Education and capacity building among healthcare providers are vital to ensure adoption, but AI can also serve as an educational tool, reinforcing correct prescribing behaviors through dynamic feedback. The integration of these platforms with surveillance systems such as GLASS enables the continuous monitoring of antimicrobial use and resistance trends, thereby allowing stewardship programs to be both data-driven and adaptable. Moreover, AI-driven interventions are strengthened through partnerships with international networks such as GARDP, the AMR Action Fund, and CARB-X, and are increasingly guided by a One Health perspective that unites diverse disciplines under a shared goal of preserving antimicrobial efficacy. **(WHO, 2025)**

The roots of artificial intelligence trace back to 1956, when the Dartmouth Conference formally introduced the concept as a new scientific field. Early AI research was driven by bold optimism, with the first generation of neural networks—the Perceptron—attempting to simulate human learning through simple pattern recognition models. Despite the technological limitations of the time, these foundational ideas laid the groundwork for the current revolution in computational medicine. AI was envisioned not just as a technical tool, but as a philosophical shift toward machines capable of learning, reasoning, and adapting—qualities now being harnessed to confront modern challenges in infectious disease management. **(Diana & Tecnología, 2025)**

AI in medicine relies on its core capabilities: learning from data and adapting to evolving environments. These systems differ fundamentally from traditional software in that they do not follow static rules; instead, they analyze patterns, generate probabilistic predictions, and update themselves as new data becomes available. This learning capacity is especially critical in microbiology, where pathogen resistance profiles and treatment responses vary across time, geography, and patient populations. Adaptability ensures that AI tools remain clinically relevant even in the face of rapidly changing resistance trends, emerging pathogens, and new antimicrobial agents.

Machine learning, a subset of AI, provides the framework for predictive modeling in clinical microbiology. It operates by training algorithms on large datasets to identify patterns and infer associations. In antimicrobial stewardship, these models are fed with data from susceptibility tests, demographic variables, clinical outcomes, and treatment regimens. Predictive models then determine the likelihood of treatment success or resistance emergence under various conditions. High-performing models can generalize these predictions to new patients, offering tailored guidance even in the absence of full clinical information. The clinical utility of such models depends on their ability to process both structured (e.g., MIC values, lab results) and unstructured (e.g., physician notes, historical behaviors) data.

There are three principal types of machine learning algorithms utilized in clinical settings. Supervised learning, the most commonly applied, involves using labeled datasets to train models to classify or predict outcomes such as resistance status or treatment response. Unsupervised learning, by contrast, finds hidden patterns in data without pre-existing labels, useful for identifying unknown clusters of resistance or transmission networks. Reinforcement learning, inspired by behavioral psychology, involves models learning through trial and error—receiving feedback after each decision to improve over time. In a stewardship context, reinforcement learning could optimize dosing strategies or escalation protocols in ICU environments.

The selection of machine learning methodology must consider both strengths and limitations. Supervised learning is powerful and interpretable when large, high-quality labeled datasets are available. However, it is less robust in novel or poorly labeled situations. Unsupervised learning offers greater flexibility in data exploration but may yield results that are difficult to validate clinically. Hybrid models that incorporate elements of both approaches, and are refined with human feedback, represent a promising path forward. These models must be tested rigorously to ensure reproducibility, interpretability, and patient safety.

The technical structure of AI models in medicine often involves deep neural networks—multi-layered architectures designed to process complex relationships within high-dimensional data. A typical deep network includes an input layer that receives clinical variables, multiple hidden layers that perform weighted computations, and an output layer that generates predictions, such as antimicrobial recommendations. These layers are trained through iterative backpropagation to minimize prediction errors. Such architectures are particularly suited for integrating diverse datasets (e.g., genetics, microbiology, demographics), producing insights that may elude conventional analytical methods.

The clinical decision-making environment is inundated with multidimensional data. Each patient interaction involves dozens of factors: pathogen species, resistance genes, an-

timicrobial mechanisms, prior allergies, site of infection, diagnostic code, age, and sex. When combined with treatment variables—such as drug interactions, guideline constraints, and FDA alerts—the number of potential permutations exceeds seventeen tetrakaquadragintillion. No human can process this complexity in real time, but AI can. These systems synthesize disparate variables to generate clinically actionable recommendations tailored to the unique microbial and patient profile, making AI a necessary tool rather than a luxury.

Arkstone's Machine Learning platform incorporates a Human-in-the-Loop (HTL) architecture to balance algorithmic precision with clinical judgment. Rather than removing the clinician from the process, HTL models actively solicit and integrate physician input, particularly in ambiguous or high-risk cases. The model assigns weighted relevance to each variable based on clinical context—prioritizing pathogen-resistance matches, alerting on potential drug interactions, and accounting for patient-specific limitations. This results in a customized, context-sensitive antimicrobial regimen that remains aligned with evidence-based guidelines while accommodating real-world complexity. **(Figure 1)**

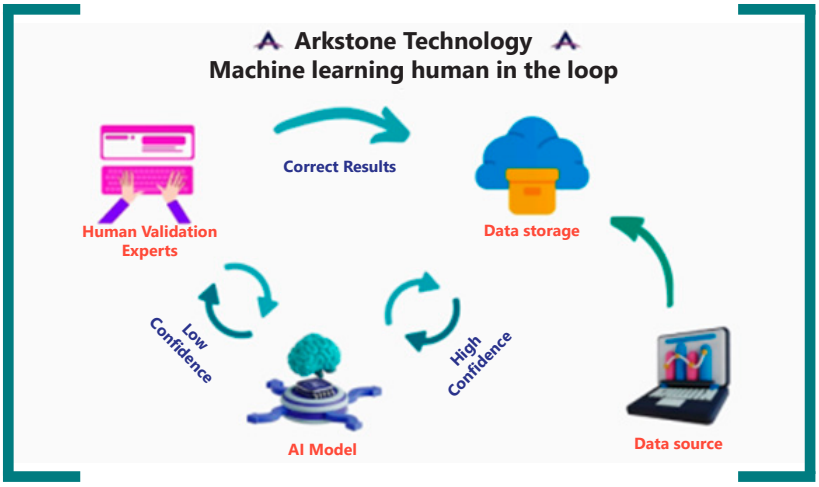

Figure 1. Schematic representation of the Arkstone model

Implementing an artificial intelligence model in medicine follows a multistep process that begins with dataset acquisition and preprocessing. High-quality, annotated data are first split into training and test subsets, often through multiple iterations to assess model stability and prevent overfitting. One common technique, cross-validation, systematically partitions the dataset to ensure robustness and generalizability. In repeated random train-test splits, data are randomly divided without fixed logic, whereas K-fold cross-validation divides data into K segments, rotating through them to validate performance across subsets. More refined methods like stratified K-fold maintain class distribution across folds, while leave-one-out cross-validation (LOOCV) tests the model on every individual

case iteratively. External validation, where the model is applied to datasets from different institutions, is essential to assess portability. Prospective validation examines model performance in real-time clinical workflows, while retrospective validation compares model predictions to past clinical decisions. These layers of validation ensure that an AI tool is not just technically sound but clinically meaningful across diverse settings. More detailed explanation you can see in the chapter before. (Chapter X)

The steps to implement a model in medicine can be seen in **Figure 2**.

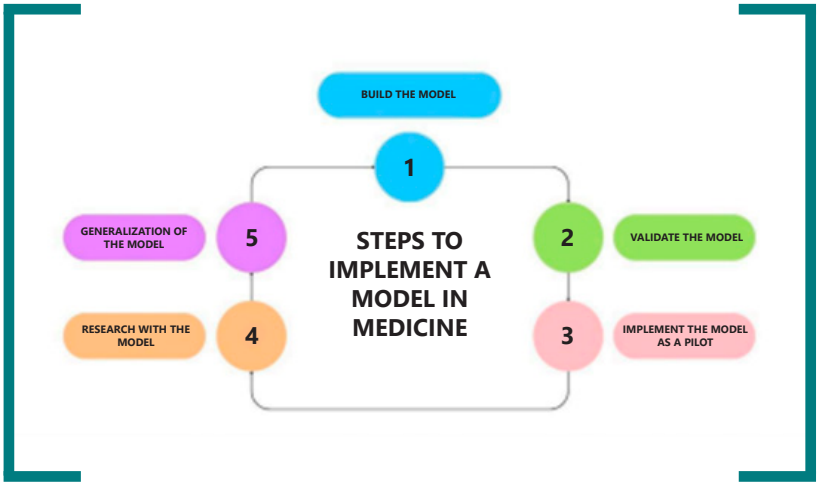

Figure 2. Steps to implement a model based on AI in medicine

A practical application of AI in infectious disease care is demonstrated through Arkstone's integration of molecular test results into therapeutic recommendations. By analyzing the genetic profiles of pathogens and aligning them with antimicrobial susceptibility patterns, the system produces targeted treatment options, often hours or days before traditional phenotypic methods. This shift in temporal advantage not only accelerates appropriate therapy but also curbs the use of broad-spectrum antibiotics, reducing resistance pressure. Visual graphics accompanying the tool illustrate the simplified flow of decision-making: molecular result input, algorithmic processing, and output in the form of a personalized regimen. The system can account for molecular markers such as CTX-M, NDM, or VIM, guiding clinicians in real time toward optimal therapies. A graphic representation of this integration can see in **Figure 3**.

Artificial intelligence and machine learning must function not in isolation but within the framework of established clinical guidelines. In infectious diseases, the IDSA (Infectious Diseases Society of America) guidelines provide evidence-based standards for managing infections such as bacteremia, pneumonia, and urinary tract infections. AI platforms trained on molecular (OneChoice®) and phenotypic data (OneChoice Fusion®) incorpo-

Figure 3. Data transformation through the model created by Arkstone, from a molecular blood culture result to report generation

rate these protocols into their decision logic, ensuring that recommendations align with gold-standard care. Furthermore, integration with local antibiograms, drug availability lists, and formulary restrictions allows models to adapt to institutional realities. Thus, AI acts as a digital steward, transforming static guidelines into dynamic, individualized treatment decisions.

The implementation of AI in medicine requires a multi-tiered strategy for validation, refinement, and integration. As detailed previously, data partitioning techniques such as K-fold cross-validation and LOOCV help ensure robustness. Yet beyond technical accuracy, validation must include clinical sensibility. Models are evaluated not just for their predictive strength but for their concordance with expert infectious disease recommendations. For example, was the pathogen correctly identified as the cause of infection? Did the AI recommend an effective drug at the correct dose and duration? Were relevant pathogens overlooked? These questions frame the core of internal validation studies. The reference by Frenkel *et al.* **(Frenkel A, *et al.* 2025)** presents a structured framework to evaluate the trainability and clinical fidelity of an AI-powered decision support system. The model's ability to recall simple and complex patterns, as well as to replicate sound therapeutic choices, is subjected to rigorous scrutiny across hundreds of cases.

The practical utility of the AI system is further exemplified by a head-to-head comparison between OneChoice Molecular Result (AOCHMR) (based solely on molecular data) vs Onechoice Fusion Results (AOCHFR) (which incorporates both molecular and phenotypic data) in Bacteremic patients. **(Gomez de la Torre, J.C, *et al.* 2025)**. Time-motion graphics display how AOCHMR accelerates decision-making by up to two days compared to conventional workflows **(Figure 4)**.

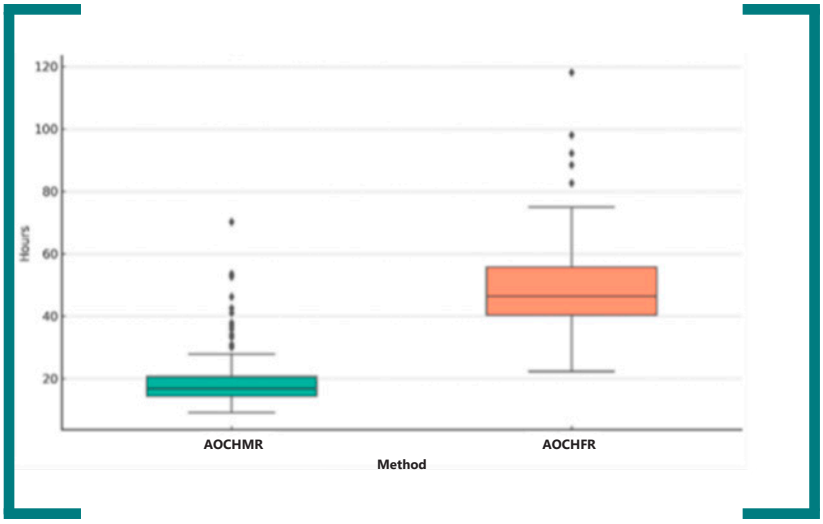

Figure 4. Time difference between AOCHMR and AOCHFR.

This temporal advantage translates into earlier optimization of therapy as described in **Figure 5**. and potentially reducing mortality, length of stay, and costs.

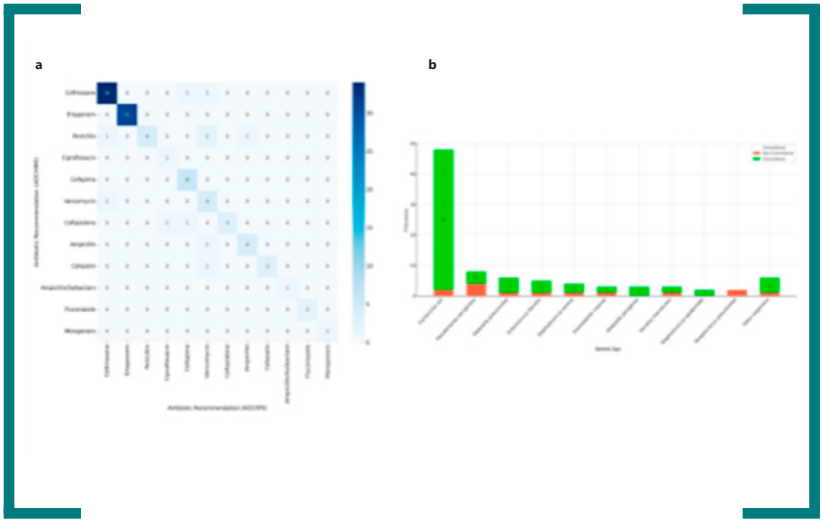

Figure 5. (a) Confusion matrix for antibiotic recommendations between AOCHMR and AOCHFR; (b) Concordance in therapeutic recommendations between AOCHMR

In a real-world cohort, the AI system's recommendations were compared to final prescriptions in patients with confirmed infections. This analysis was conducted across multiple metrics, including mortality at 28 days, duration of hospitalization, adverse effects, and drug interactions. The AI system demonstrated a capacity to reduce unnecessary broad-spectrum use, thereby minimizing toxicities and resistance propagation. The da-

taset also allowed subgroup analyses by infection site, pathogen, and resistance mechanism. These metrics offer compelling evidence for the system's clinical value, not only in terms of accuracy but also in improving meaningful patient outcomes.

This ongoing retro-prospective study seeks to evaluate the OneChoice® AI report in a Peruvian private hospital. Beginning in January 2021 and extending through March 2026, this study involves over 200 patients hospitalized with microbiologically confirmed infections. Molecular diagnostic data are processed through the AI engine, and the clinical team evaluates the concordance between AI suggestions and actual therapeutic outcomes. Oversight is provided by the INSN institutional ethics committee, ensuring methodological rigor. Although the study is in progress, preliminary results have demonstrated promising alignment between AI-driven and physician-selected therapies. **(Gomez de la Torre, JC *et al.*, 2025)**.

One of the clearest examples of the value of AI emerges in the management of urinary tract infections. Traditional workflows rely on urine microscopy and culture, with turnaround times of 48–72 hours. During this window, clinicians must prescribe empirically, often without knowing the causative organism or resistance profile. Failure rates in empirical therapy can reach 20–30%. By contrast, if we use the OneChoice® model using machine learning predictions based on over 185,000 historical isolates to propose optimal therapy from the outset, we may be able to reduce these risks of failure by predicting the microorganism, antimicrobial resistance thanks to the model. Variables such as age, gender, prior infection history, and comorbidities are factored in, generating predictive models for pathogen presence and resistance. These predictions reduce empirical guesswork and may be allowed clinicians to tailor treatment even before final culture results are available.

The implementation of the AI tool at the biggest public hospital in Perú further illustrates how clinical models must be adapted to local conditions. The hospital's antimicrobial formulary lacks several advanced agents such as cefiderocol, daptomycin, and ceftolozane-tazobactam. The AI engine is thus customized to exclude unavailable options while adjusting recommendations based on drug accessibility and institutional restrictions. For example, linezolid and levofloxacin are tightly restricted due to high tuberculosis prevalence. Surveillance practices, such as rectal swabs for CRE carriers, must also be integrated into AI outputs. These nuances underscore the necessity of contextual adaptation for successful AI deployment in diverse health systems.

On a global scale **(Han, Ryan *et al.*, 2024)**, AI is being tested across numerous clinical domains through randomized controlled trials (RCTs). A scoping review published in The Lancet Digital Health mapped the distribution of AI RCTs across over 20 countries and multiple specialties, with European nations leading in implementation. Infectious diseases remain

underrepresented, highlighting a critical opportunity for growth. As this field matures, the application of AI to antimicrobial stewardship may shift from isolated pilots to multicenter, multicountry implementations driven by standardized validation protocols.

The application of molecular technologies and machine learning in the diagnosis and management of infectious diseases is transforming the clinical landscape. Rapid molecular identification of pathogens accelerates the time to appropriate therapy, while AI systems ensure that treatment decisions are not only fast but evidence-based and individualized. The generalization of AI models across institutions enables scaling of these benefits, reducing hospital length of stay and healthcare costs. The future of antimicrobial stewardship will rely increasingly on AI to navigate the growing complexity of microbiological data, drug resistance trends, and therapeutic options. Ongoing research must continue to refine these tools, ensure equitable access, and validate their effectiveness in diverse populations and care settings.

BIBLIOGRAPHY

1. **World Health Organization.** (2023). Antimicrobial resistance. https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance
2. http://www.dianaytecnologia.com/historia-de-la-inteligencia-artificial. Take at 07/09/2025
3. **Frenkel A, Rendon A, Chavez-Lencinas C, Gomez De la Torre JC, MacDermott J, Gross C, Allman S, Lundblad S, Zavala I, Gross D, *et al.*** Internal Validation of a Machine Learning-Based CDSS for Antimicrobial Stewardship. Life. 2025; 15(7):1123. https://doi.org/10.3390/life15071123
4. **Gomez de la Torre, J. C., Frenkel, A., Chavez-Lencinas, C., *et al.*** (2025). AI-based treatment recommendations enhance speed and accuracy in bacteremia management: A comparative study of molecular and phenotypic data. Life, 15(6), 864. https://doi.org/10.3390/life15060864
5. **Gómez de la Torre, JC; Frenkel A; Hueda M, Chavez L C**; Clinical utility of a clinical decisión support system: Onechoice® report on the choice of rational antibiotic therapy in hospitalized patients with positive molecular microbiology results in a Private Clinic in Lima, in the period 2021-2024"(In progress)
6. **Han, R., *et al.*** (2024). Randomised controlled trials evaluating artificial intelligence in clinical practice: a scoping review;The Lancet Digital