applied microbiology

MDPI

Article

A Reinforcement Learning Approach for Modeling Organic Compound-Induced Antimicrobial Resistance Dynamics

Hayden D. Hedman [1*]

[*1]School of Computer Science, Georgia Institute of Technology, Atlanta, GA, USA; hhedman3@gatech.edu

Abstract

This study investigates the use of reinforcement learning (RL) to model antimicrobial resistance (AMR) dynamics driven by copper exposure. In a simulated environment, three agent strategies were evaluated: (1) random, (2) rule-based, and (3) Q-learning. These agents were compared to assess how varying levels of adaptability affect bacterial resistance to chloramphenicol and polymyxin B. The Q-learning agent consistently achieved lower resistance levels and copper burden compared to both the rule-based and random agents, demonstrating its potential for optimizing AMR management. The rule-based agent showed intermediate performance, while the random agent exhibited higher variability in resistance trajectories. This study introduces an exploratory RL framework for simulating AMR control strategies, especially in scenarios where traditional diagnostic tools are unavailable. The results demonstrate that RL can provide valuable insights into optimizing AMR management, particularly in settings with limited access to traditional diagnostic tools and resources.

Keywords: reinforcement learning; copper exposure; antibiotic resistance dynamics; Q-learning; agent-based modeling; microbial resistance control

Academic Editor: Firstname Lastname

Received: date
Revised: date
Accepted: date
Published: date

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Antimicrobial resistance (AMR) continues to be one of the most critical global health challenges *(1–3)*, with resistance to conventional antibiotics growing at an alarming rate. As the effectiveness of many synthetic antibiotics continues to decline *(4–6)*, alternative organic compounds represent a complementary avenue for exploring strategies to manage resistance *(7–9)*. Copper is a well-studied antimicrobial agent with broad activity across microbial species and a documented ability to inhibit microbial growth *(10–12)* However, its application in clinical and agricultural settings remains poorly understood, particularly regarding its interaction with microbial resistance over time. Rather than serving as a replacement for conventional therapies, copper provides a biologically grounded model compound for examining how organic antimicrobial pressure shapes resistance dynamics over time, motivating the need for systematic and adaptive experimentation.

The evolution of antimicrobial resistance can be naturally framed as a sequential decision-making and feedback-control problem. Recent foundational work has demonstrated that reinforcement learning (RL) can learn adaptive intervention strategies capable of influencing resistance trajectories in complex evolutionary systems defined by antibiotic fitness landscapes, even under noisy or delayed feedback conditions *(13,14)*. These

studies establish the feasibility of RL-based approaches for AMR control but focus primarily on synthetic antibiotics and highly structured treatment environments. As a result, comparatively little work has examined how adaptive decision-making frameworks behave when applied to organic antimicrobial compounds or simplified settings that more closely reflect constrained clinical or infrastructural capacities.

Building on this foundation, the present study applies an RL-based simulation framework using copper as a representative organic antimicrobial stressor to explore how different decision-making strategies respond to antimicrobial pressure across variable and constrained environments. The goal is not to evaluate the effectiveness of copper itself, but to examine how varying levels of adaptability influence resistance dynamics under simplified assumptions. To this end, three agents are considered: (1) a random agent representing uncontrolled exposure, (2) a rule-based agent implementing fixed decision criteria, and (3) a Q-learning agent employing an adaptive administration strategy. By comparing these agents, this work provides an exploratory framework for understanding how adaptive control strategies may perform in settings where diagnostic feedback, infrastructure, or precise treatment control is limited.

2. Materials and Methods

2.1. Overview of Study and Framework

This study implemented a reinforcement learning–based simulation framework to model antibiotic resistance dynamics under varying copper exposure conditions. The experimental setup was rooted in copper concentration estimates reported in the foundational study by Li et al. (2021) *(15)*, with parameters adapted to simulate copper exposure and antibiotic resistance evolution in *Escherichia coli K12 and Staphylococcus aureus*. The RL environment and agents were implemented in Python (Gymnasium, NumPy, and Matplotlib) for simulation, computation, and figure generation. Simulations modeled bacterial populations exposed to varying copper concentrations in combination with antibiotics chloramphenicol and polymyxin B *(15)*. Three agents were evaluated: (1) a random policy agent, (2) a rule-based agent, and (3) a Q-learning agent. These models were designed to simulate control strategies for bacterial growth and resistance under copper exposure.

2.2. Reinforcement Learning Agent Design

The RL agents interacted with a simulation environment designed to capture the key features of copper-driven antimicrobial pressure and resistance dynamics over discrete cycles *(15)*. The environment represents bacterial populations exposed to variable copper concentrations in the presence of antibiotics, without explicitly modeling underlying molecular or genetic mechanisms. The state space included copper concentration, minimum inhibitory concentration (MIC) values for chloramphenicol and polymyxin B, a growth inhibition proxy, and cycle count. These state variables were selected to reflect commonly reported microbiological indicators of resistance and treatment response, including MIC and growth inhibition *(16,17)*.

The action space allowed agents to adjust copper exposure by decreasing, maintaining, or increasing concentration at each cycle. Copper concentration ranges were grounded in sublethal exposure estimates reported by Li and colleagues *(15)*, with values spanning 10–100 mg/L. Other biological dynamics, including changes in MIC values and growth inhibition, were modeled using stochastic update rules intended to represent resistance progression and partial reversibility under reduced antimicrobial pressure, rather than explicit mechanistic mutation processes. To promote interpretability and transparency, continuous variables were discretized into bins prior to agent learning. This design

choice emphasized conceptual biological relevance and clarity over fine-grained mechanistic detail, consistent with the exploratory objectives of this study.

2.3. Random Policy Agent

The random policy agent served as a baseline, selecting actions at random without adapting to bacterial resistance dynamics or environmental feedback. This agent illustrated the progression of resistance in the absence of strategic control and provided a reference point for evaluating the impact of structured and adaptive decision-making. Its performance reflected baseline resistance evolution under non-optimized copper exposure.

2.4. Rule-Based Agent

The rule-based agent followed predefined decision rules to adjust copper exposure based on observed antibiotic resistance, as measured by MIC values. Copper levels were reduced when resistance increased, maintained when MIC values remained stable, and increased when resistance decreased. This agent reflects a protocol-driven decision strategy that could be applied in settings with limited computational or diagnostic resources. Its performance was compared with that of the Q-learning agent to evaluate the potential benefits of adaptive strategies for managing resistance dynamics.

2.5. Q-Learning Agent

The Q-learning agent employed a model-free reinforcement learning algorithm to optimize copper exposure through iterative interaction with the environment. Continuous state variables, including copper concentration, MIC values, and cycle count, were discretized into bins and mapped to a Q-table. The agent updated the Q-table based on trial-and-error learning using observed rewards. The reward function encouraged suppression of bacterial growth and resistance while penalizing excessive copper exposure. Performance of the Q-learning agent was compared with that of the random and rule-based agents to assess the efficiency of adaptive decision-making in managing copper exposure and antibiotic resistance.

2.6. Analytical Methods

Performance metrics included the Area Under the Curve (AUC) of MIC values, copper burden, and mean growth inhibition, calculated for each agent across simulation cycles. These metrics were used to quantify resistance dynamics and control efficiency under different decision-making strategies. Statistical analyses were conducted using generalized linear mixed models (GLMMs) *(18)* in R (packages: lme4, lmerTest, readr, and emmeans). For each outcome (MIC for chloramphenicol and polymyxin B), models included cycle and agent type as fixed effects, with episode specified as a random effect to account for variability across experimental runs. Pairwise contrasts were evaluated using 95% confidence intervals (CIs) to assess meaningful differences between agents.

2.6. Reproducibility and Generative AI Statement

All code used for experiments, data analysis, and figure generation is provided in the supplementary material (SM1). Scripts are organized with clear instructions for setup and execution. Raw data and generated figures are saved in CSV and image formats, respectively, allowing validation and extension of the framework. Generative AI tools were used for coding assistance, troubleshooting, and initial idea generation; all outputs were reviewed prior to inclusion in the final manuscript.

3. Results

3.1. Agent-level resistance dynamics based on AUC metrics

Across both antibiotics, the Q-learning agent exhibited lower cumulative resistance exposure than the random and rule-based agents, as measured by the area under the curve (AUC) of minimum inhibitory concentration (MIC) values (Table 1; Table 2; Figures 1–4). Bootstrap contrasts indicated lower AUC MIC values for the Q-learning agent relative to the random agent for both chloramphenicol and polymyxin B. Differences between the Q-learning and rule-based agents were smaller, consistent with intermediate performance under rule-based decision-making (Table 1).

These patterns were reflected in aggregate performance metrics across experimental runs. The Q-learning agent achieved lower mean AUC MIC values for both antibiotics and lower copper burden compared with the random and rule-based agents. In contrast, the random agent exhibited higher cumulative MIC exposure and copper burden, with greater variability across runs, while the rule-based agent demonstrated intermediate values across metrics. Mean growth inhibition remained high across agents throughout experimental cycles. This suggests that differences in AUC MIC primarily reflected divergence in resistance trajectories over time rather than short-term differences in growth suppression.

Table 1. Bootstrap AUC summary for agent comparisons

Contrast	Metric	Median Difference	
Q-Learner vs. Random	AUC MIC (Chloro)	−114.818	−1
Q-Learner vs. Random	AUC MIC (PolyB)	−102.130	−1
Q-Learner vs. Rule-Based	AUC MIC (Chloro)	−41.116	−5
Q-Learner vs. Rule-Based	AUC MIC (PolyB)	−50.967	−5
Rule-Based vs. Random	AUC MIC (Chloro)	−58.424	−1
Rule-Based vs. Random	AUC MIC (PolyB)data	−44.177	−1

Table 2. Summary metrics for agent performance (Mean and Standard Deviation) for Chloramphenicol (Chloro) and Polymyxin B (PolyB).

Agent	AUC MIC (Chloro) Mean (SD)	AUC MIC (PolyB) Mean (SD)	Copper Burden Mean (SD)	Mean Growth Inhibition Mean (SD)
Q-Learner	10.131 (8.917)	22.229 (12.035)	15.050 (9.087)	0.999 (0.000)
Random	183.340 (175.391)	149.517 (113.335)	564.200 (394.399)	0.992 (0.007)
Rule-Based	69.739 (52.945)	76.042 (32.698)	142.200 (91.390)	0.996 (0.002)

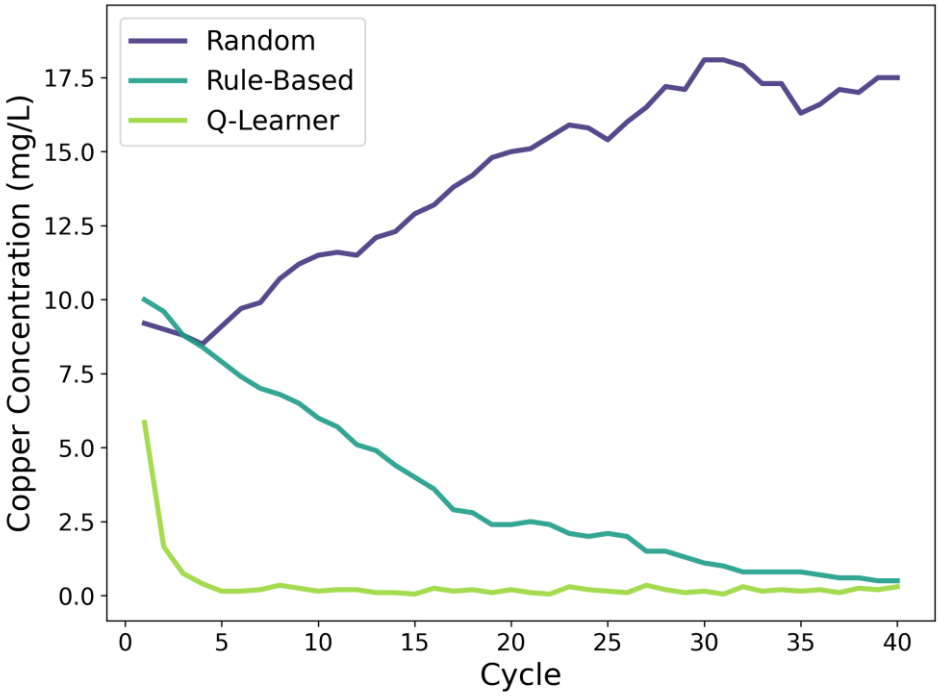

Figure 1. Copper concentration (mg/L) across experimental cycles for each agent. The figure shows how copper exposure was adjusted over cycles by the random, rule-based, and Q-learning agents.

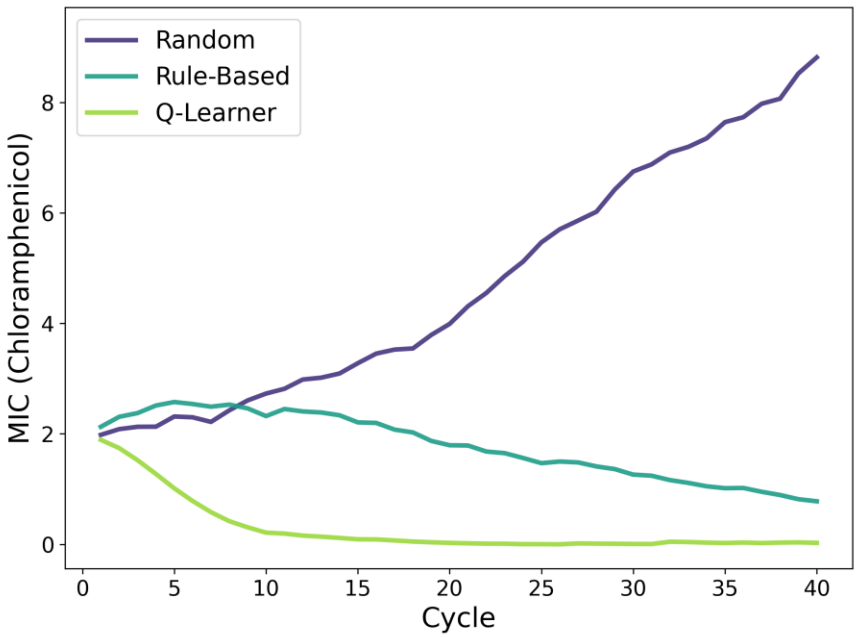

Figure 2. Chloramphenicol MIC trajectories across experimental cycles, illustrating differences in agent-driven control of resistance.

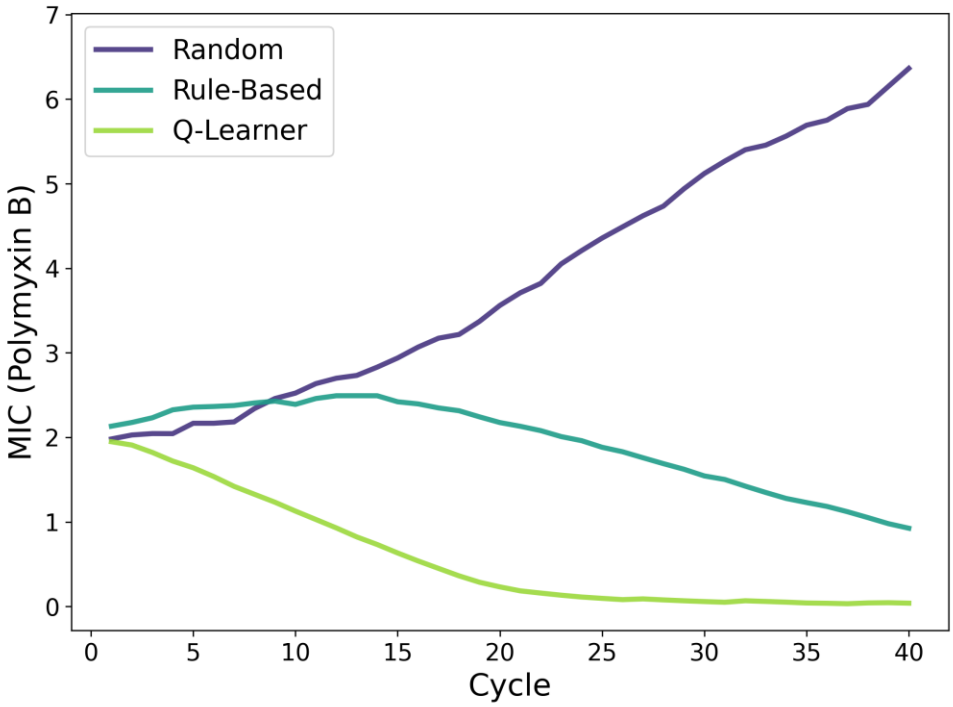

Figure 3. Polymyxin B minimum inhibitory concentration (MIC) across experimental cycles for the random, rule-based, and Q-learning agents.

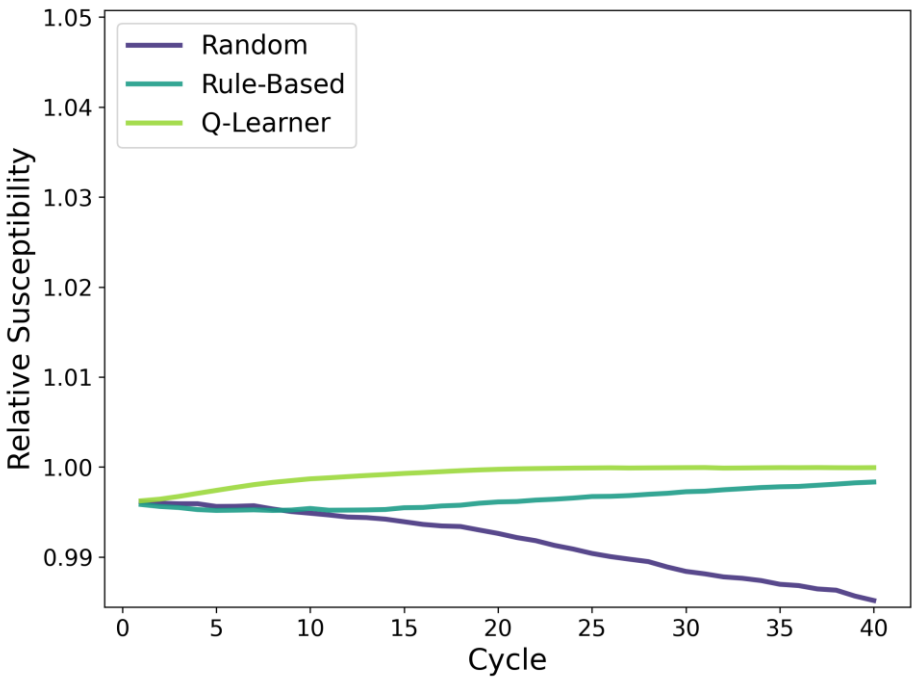

Figure 4. Growth inhibition across experimental cycles for the random, rule-based, and Q-learning agents.

3.2. GLMM analysis of resistance trajectories

Results from generalized linear mixed models (GLMMs) aligned with the AUC-based summaries, providing model-based estimates of differences in resistance trajectories across agents. Relative to the Q-learning agent, the random and rule-based agents were associated with higher MIC values over cycles for both antibiotics, with effect sizes varying by agent and antibiotic. Interaction terms between cycle and agent indicated differential temporal patterns in MIC trajectories. The random agent showed increasing MIC values over cycles relative to the Q-learning agent, while the rule-based agent exhibited smaller interaction effects, consistent with intermediate resistance trajectories. Across models, estimated effects for polymyxin B were generally smaller in magnitude than those observed for chloramphenicol, although the direction of effects was consistent across antibiotics.

Table 3. Changes in antibiotic minimum inhibitory concentration (MIC) across simulation cycles by agent type, with 95% confidence intervals.

Effect	Chloramphenicol Estimate (95% CI)	Polymyxin B Estimate (95% CI)
Cycle	−0.031 (−0.038 – −0.024)	−0.051 (−0.055 – −0.046)
Random	0.062 (−0.229 – 0.352)	−0.222 (−0.412 – −0.033)
Rule-Based	1.879 (1.589 – 2.169)	1.113 (0.924 – 1.302)
Cycle × Random	0.214 (0.202 – 0.226)	0.170 (0.162 – 0.178)
Cycle × Rule-Based	−0.018 (−0.030 – −0.006)	0.012 (0.004 – 0.020)

4. Discussion

This study explored the application of reinforcement learning (RL) to model and optimize strategies for controlling antimicrobial resistance (AMR). The Q-learning agent consistently outperformed both the rule-based and random agents in reducing resistance levels, highlighting the potential of RL as a tool for simulating and refining AMR control strategies. These findings suggest that RL, particularly Q-learning, could serve as an effective method for dynamic decision-making in environments where AMR is a growing concern *(14)*. However, while the results demonstrate clear agent-level differences, further refinement and validation of the model are necessary to more accurately capture the complexities of AMR dynamics in natural environments *(19,20)*.

This study establishes an exploratory framework for applying reinforcement learning (RL) to the use of organic compounds in mitigating antimicrobial resistance (AMR) across various bacterial species and environmental conditions. While the model provides a novel framework, it serves as a preliminary baseline for RL applications in AMR, offering a foundation for further development. Future studies could build upon this framework by incorporating greater complexity in AMR evolution, such as microbial interactions, horizontal gene transfer, and host immunity *(21–23)* . This study also sets the stage for future work evaluating other organic compounds as potential AMR interventions *(24,25)*. This study focuses on establishing foundational principles, setting the stage for future advancements, rather than attempting to provide a comprehensive solution to

AMR. Ultimately, it serves as a starting point for future research to refine and expand the model.

From a broader systems perspective, this study highlights the potential of reinforcement learning (RL) to provide a dynamic and adaptive framework for studying AMR. By simulating various agent-driven interventions, RL enables the exploration of complex feedback loops between microbial resistance, drug exposure, and environmental factors. This approach offers flexibility in developing more targeted and effective AMR interventions, especially in resource-limited settings *(26,27)*. While this study focused on copper as an organic compound, it lays the groundwork for future research into RL applications for other organic antimicrobials. Incorporating RL into AMR research could lead to novel strategies for managing AMR in both clinical and agricultural settings, offering insights that could significantly advance global efforts to combat this urgent health crisis.

Supplementary Materials: All relevant code, data, and figures are available in SM1.

References

1. Oldenkamp, R.; Schultsz, C.; Mancini, E.; Cappuccio, A. Filling the gaps in the global prevalence map of clinical antimicrobial resistance. *Proceedings of the National Academy of Sciences* **2021**, *118*(1), e2013515118. doi:10.1073/pnas.2013515118.

2. Vikesland, P.; Garner, E.; Gupta, S.; Kang, S.; Maile-Moskowitz, A.; Zhu, N. Differential Drivers of Antimicrobial Resistance across the World. *Accounts of Chemical Research* **2019**, *52*(4), 916–924. doi:10.1021/acs.accounts.8b00643.

3. Bertagnolio, S.; Dobreva, Z.; Centner, C. M.; Olaru, I. D.; Donà, D.; Burzo, S.; et al. WHO global research priorities for antimicrobial resistance in human health. *The Lancet Microbe* **2024**, *5*(11). doi:10.1016/S2666-5247(24)00134-4.

4. Muteeb, G.; Rehman, M. T.; Shahwan, M.; Aatif, M. Origin of Antibiotics and Antibiotic Resistance, and Their Impacts on Drug Development: A Narrative Review. *Pharmaceuticals* **2023**, *16*(11), 1615. doi:10.3390/ph16111615.

5. Chiș, A. A.; Rus, L. L.; Morgovan, C.; Arseniu, A. M.; Frum, A.; Vonica-Țincu, A. L.; et al. Microbial Resistance to Antibiotics and Effective Antibiotherapy. *Biomedicines* **2022**, *10*(5), 1121. doi:10.3390/biomedicines10051121.

6. Marston, H. D.; Dixon, D. M.; Knisely, J. M.; Palmore, T. N.; Fauci, A. S. Antimicrobial Resistance. *JAMA* **2016**, *316*(11), 1193–1204. doi:10.1001/jama.2016.11764.

7. Elshobary, M. E.; Badawy, N. K.; Ashraf, Y.; Zatioun, A. A.; Masriya, H. H.; Ammar, M. M.; et al. Combating Antibiotic Resistance: Mechanisms, Multidrug-Resistant Pathogens, and Novel Therapeutic Approaches: An Updated Review. *Pharmaceuticals* **2025**, *18*(3), 402. doi:10.3390/ph18030402.

8. Berger, I.; Loewy, Z. G. Antimicrobial Resistance and Novel Alternative Approaches to Conventional Antibiotics. *Bacteria* **2024**, *3*(3), 171–182. doi:10.3390/bacteria3030012.

9. Elfadadny, A.; Ragab, R. F.; AlHarbi, M.; Badshah, F.; Ibáñez-Arancibia, E.; Farag, A.; et al. Antimicrobial resistance of Pseudomonas aeruginosa: navigating clinical impacts, current resistance trends, and innovations in breaking therapies. *Frontiers in Microbiology* **2024**, *15*. doi:10.3389/fmicb.2024.1374466.

10. Ivanova, I. A.; Daskalova, D. S.; Yordanova, L. P.; Pavlova, E. L. Copper and Copper Nanoparticles Applications and Their Role against Infections: A Minireview. *Processes* **2024**, *12*(2), 352. doi:10.3390/pr12020352.

11. Yu, J.; Huang, X.; Ren, F.; Cao, H.; Yuan, M.; Ye, T.; et al. Application of antimicrobial properties of copper. *Applied Organometallic Chemistry* **2024**, *38*(7), e7506. doi:10.1002/aoc.7506.

12. Wang, Y.; Wen, T.; Mao, F.; Yang, S.; Zhang, Q.; Fu, X.; et al. Engineering copper and copper-based materials for a post-antibiotic era. *Frontiers in Bioengineering and Biotechnology* **2025**, *13*. doi:10.3389/fbioe.2025.1644362.

13. Tomurcuk, A. E.; Bozcali, E.; Baytemur, F. D.; Sivri, C.; Soyak, E. G. A Reinforcement Learning-inspired Estimation of Antibiotic Resistance in Poultry Environments. In *Engineering and Technology Management in Challenging Times*. Çebi, F., Tekin, A. T., Çorum, A., Cömert, G. D., Bolatan, G. I. S., Eds.; Springer Nature Switzerland: Cham, 2024; pp 91–99. doi:10.1007/978-3-031-70935-7_7.

14. Weaver, D. T.; King, E. S.; Maltas, J.; Scott, J. G. Reinforcement learning informs optimal treatment strategies to limit antibiotic resistance. *Proceedings of the National Academy of Sciences* **2024**, *121*(16), e2303165121. doi:10.1073/pnas.2303165121.

15. Li, J.; Phulpoto, I. A.; Zhang, G.; Yu, Z. Acceleration of emergence of E. coli antibiotic resistance in a simulated sublethal concentration of copper and tetracycline co-contaminated environment. *AMB Express* **2021**, *11*(1), 14. doi:10.1186/s13568-020-01173-6.

16. Smaill, F. Antibiotic Susceptibility and Resistance Testing: An Overview. *Canadian Journal of Gastroenterology and Hepatology* **2000**, *14*(10), 382415. doi:10.1155/2000/382415.

17. Kadeřábková, N.; Mahmood, A. J. S.; Mavridou, D. A. I. Antibiotic susceptibility testing using minimum inhibitory concentration (MIC) assays. *npj Antimicrobials and Resistance* **2024**, *2*(1), 37. doi:10.1038/s44259-024-00051-6.

18. Bolker, B. M.; Brooks, M. E.; Clark, C. J.; Geange, S. W.; Poulsen, J. R.; Stevens, M. H. H.; et al. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution* **2009**, *24*(3), 127–135. doi:10.1016/j.tree.2008.10.008.

19. Frei, A.; Verderosa, A. D.; Elliott, A. G.; Zuegg, J.; Blaskovich, M. A. T. Metals to combat antimicrobial resistance. *Nature Reviews Chemistry* **2023**, *7*(3), 202–224. doi:10.1038/s41570-023-00463-4.

20. Denk-Lobnig, M. K.; Wood, K. B. Spatial population dynamics of bacterial colonies with social antibiotic resistance. *Proceedings of the National Academy of Sciences* **2025**, *122*(7), e2417065122. doi:10.1073/pnas.2417065122.

21. Abel, K.; Agnew, E.; Amos, J.; Armstrong, N.; Armstrong-James, D.; Ashfield, T.; et al. System-wide approaches to antimicrobial therapy and antimicrobial resistance in the UK: the AMR-X framework. *The Lancet Microbe* **2024**, *5*(5), e500–e507. doi:10.1016/S2666-5247(24)00003-X.

22. Woods, L. C.; Gorrell, R. J.; Taylor, F.; Connallon, T.; Kwok, T.; McDonald, M. J. Horizontal gene transfer potentiates adaptation by reducing selective constraints on the spread of genetic variation. *Proceedings of the National Academy of Sciences* **2020**, *117*(43), 26868–26875. doi:10.1073/pnas.2005331117.

23. Wit, G. D.; Svet, L.; Lories, B.; Steenackers, H. P. Microbial Interspecies Interactions and Their Impact on the Emergence and Spread of Antimicrobial Resistance. *Annual Review of Microbiology* **2022**, *76*(Volume 76, 2022), 179–192. doi:10.1146/annurev-micro-041320-031627.

24. Guedes, B. N.; Krambeck, K.; Durazzo, A.; Lucarini, M.; Santini, A.; Oliveira, M. B. P. P.; et al. Natural antibiotics against antimicrobial resistance: sources and bioinspired delivery systems. *Brazilian Journal of Microbiology* **2024**, *55*(3), 2753–2766. doi:10.1007/s42770-024-01410-1.

25. Alaoui Mdarhri, H.; Benmessaoud, R.; Yacoubi, H.; Seffar, L.; Guennouni Assimi, H.; Hamam, M.; et al. Alternatives Therapeutic Approaches to Conventional Antibiotics: Advantages, Limitations and Potential Application in Medicine. *Antibiotics* **2022**, *11*(12), 1826. doi:10.3390/antibiotics11121826.

26. DiazGranados, C. A.; Cardo, D. M.; McGowan, J. E. Antimicrobial resistance: international control strategies, with a focus on limited-resource settings. *International Journal of Antimicrobial Agents* **2008**, *32*(1), 1–9. doi:10.1016/j.ijantimicag.2008.03.002.

27. Hijazi, K.; Joshi, C.; Gould, I. M. Challenges and opportunities for antimicrobial stewardship in resource-rich and resource-limited countries. *Expert Review of Anti-infective Therapy* **2019**, *17*(8), 621–634. doi:10.1080/14787210.2019.1640602.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.