Harnessing Machine learning models for Enhanced Antimicrobial Resistance Surveillance and Intervention in African Countries

Kamini Govender[1,2*], Ali Mahmoud Youssouf[3], Wendy Dogbegah[4], Dieudonne Nyamifofe[5], and Joseph Igbokwe[6]

[1]University of the Free State (UFS), South Africa.
[2]SynBio Africa, Entebbe, Uganda.
[3]Amoud University, Somalia
[4]University of Ghana, Ghana
5 Universite Officielle de Bukavu, Democratic Republic of the Congo
[6]Obafemi Awolowo University, Nigeria
*Team Lead

ABSTRACT

Antimicrobial resistance (AMR) susceptibility testing is crucial for assessing the efficacy of antimicrobial agents against pathogens, ensuring effective treatment, and preventing the spread of resistant infections. Despite its importance, Africa faces significant challenges, with only a small number of laboratories equipped for essential bacteriology testing, many conducting fewer than 1,000 susceptibility tests annually. To address this, we propose integrating AMR susceptibility testing with Minimum Inhibitory Concentration (MIC) data into an AI-driven analysis framework, aiming to identify emerging resistance patterns, improve antibiotic prescribing practices, and reduce the spread of resistant pathogens. Using data from the ATLAS dataset from Vivli, WHO GLASS, CLSI M100 Performance Standards, and the WHO Bacteria Priority List 2024, we developed AI models for Methicillin-resistant *Staphylococcus aureus*. The integration of AI techniques with AMR testing offers real-time insights into resistance dynamics, enabling public health officials to implement timely interventions, allocate resources efficiently, and design targeted awareness campaigns. In Objective 1, we aimed to determine the MIC for Tier 1 antibiotic, where the root mean square error (RMSE) of predictions from four methods ranged from 1.2 to 2.4. Among these methods, Partial Least Squares Regression (PLSR) was identified as the most accurate predictor of MIC for Erythromycin. In Objective 2, focusing on Tier 2 antibiotic, RMSE values varied from 0.9 to 1.9, with XGBoost emerging as the best predictor for Penicillin. For Objective 3, which targeted Tier 3 antibiotics, RMSE values ranged from 1.8 to 2.3, with XGBoost again proving to be the top predictor, this time for Cefazolin. Although these RMSE values are relatively high, this may be attributed to the small sample size of 3,943 for *S. aureus* from the Africa region. While this research provides tools and strategies tailored to Africa's unique healthcare landscape, it also highlights challenges such as limited data availability, laboratory underrepresentation, variability in data collection practices, and infrastructure constraints. Addressing these issues is essential to fully realize the potential of AI in combating AMR across the continent.

INTRODUCTION

AMR susceptibility testing is a critical laboratory procedure used to evaluate the effectiveness of antimicrobial agents, such as antibiotics, against specific pathogens (1). By determining whether microorganisms are resistant or susceptible to certain drugs, these tests play a vital role in ensuring effective treatment, controlling the spread of resistant infections, and supporting public health efforts to combat AMR. However, the state of AMR susceptibility testing in Africa presents considerable challenges. Recent reports reveal that only a small

percentage of laboratories across the continent are equipped to perform the necessary bacteriology testing. In 14 African Union Member States, a mere 1.3% of the 50,000 medical laboratories conduct such tests, with many performing fewer than 1,000 antimicrobial susceptibility tests annually (2).

To address these challenges, we propose integrating AMR susceptibility testing, MIC data with AI-driven analysis to identify emerging resistance patterns. This approach will optimize antibiotic prescribing practices and reduce the spread of resistant pathogens. The insights gained from this analysis will enable public health officials to implement timely interventions, allocate resources more effectively, and design targeted awareness campaigns. By combining AI techniques with AMR susceptibility testing and data analysis, we aim to drive effective interventions and support global efforts to combat antimicrobial resistance. Our goal is to develop continent-wide AI models for AMR susceptibility testing, using information sourced from ATLAS dataset from Vivli (3), WHO GLASS (4), CLSI M100 Performance Standards for Antimicrobial Susceptibility Testing (5), and the WHO Bacteria Priority List, 2024 (6).

OBJECTIVES

Objective 1: Develop an AI Model for AMR Susceptibility Testing (MIC) for Tier 1 antibiotics.
The objective is to develop an AI model that predicts antibiotic susceptibility in antimicrobial resistance (AMR) testing. The model will be trained and evaluated using information from Tier 1 antibiotics, as outlined in Table 1C of the CLSI M100 Performance Standards for Antimicrobial Susceptibility Testing (5). The focus will be on training and testing the model for Erythromycin susceptibility in *S. aureus.*

Objective 2: Develop an AI Model for AMR Susceptibility Testing (MIC) for Tier2 antibiotics.
The objective is to develop an AI model that predicts susceptibility across various antibiotics with comparable interpretive categories (susceptible, intermediate, or resistant) and similar clinical efficacy. The model will be trained and evaluated using Penicillin data from Tier 2 (5), Table 1C.

Objective 3: Develop an AI Model for AMR Susceptibility Testing MIC) for Tier 3 antibiotic.
The objective is to develop an AI model grounded in the principles of selective and cascade reporting for AMR susceptibility testing. This model will be trained and evaluated using data from Cefazolin in Tier 3, as specified in Table 1C (5), explained in Objectives 1 and 2.

METHODS

In selecting priority pathogens, the WHO Bacteria Priority List 2024 (BPPL 2024 [6]) was used, with Methicillin-resistant *S. aureus* (MRSA) identified as a high-priority, gram-positive bacterium due to its prevalence and high MRSA rates in South Africa (7). This choice was informed by demographic data visualized on an interactive GIS dashboard, created from the ATLAS dataset covering all African countries (8), and presented in (9). The selection of antibiotics (Erythromycin, Penicillin, and Cefazoline) was based on information from WHO GLASS (4) and the Performance Standards for Antimicrobial Susceptibility Testing, Table 1C of the referenced article (5). Wildfire et al (10) found that that *S. aureus* MIC distribution by

gender, age-group and infection site displayed a clear trend. We will therefore use these three variables as our predictor variables.

Machine Learning Modeling
MIC values were modelled using R version 4.4.0 (11), the machine learning methods used was XGBoost, Random Forest, and Partial Least Square Regression (PLSR). Root mean square errors (RMSE) of prediction was used for performance evaluations. The AI methods were trained on 75% (2,957) of the data and tested on 25% (986) of the data. The R scripts are available on GitHub (https://github.com/so13839/SynBio-Team2).

RESULTS
In Objective 1, the aim was to determine the MIC for Tier 1 (5). The root mean square error (RMSE) of predictions from the three methods, as shown in Table 1, ranged from 1.2 to 2.4. Among these methods, PLSR was identified as the most accurate predictor of MIC for Erythromycin. Moving to Objective 2, the focus was on determining the MIC for Tier 2, where RMSE values varied from 0.9 to 1.9, and XGBoost was found to be the best predictor for Penicillin. In Objective 3, the goal was to determine the MIC for Tier 3. In this case, RMSE values ranged from 1.8 to 2.3, with XGBoost again emerging as the top predictor, this time for Cefazolin. While these RMSE values are relatively high, this may be due to the small sample size of 3,943 for *S. aureus* from the Africa region, in contrast to the larger datasets used in studies by Wildfire and colleagues (10).

THE IMPACT OF THE STUDY

For Researchers in Africa
For African researchers working on AMR developing AI-driven models that specifically analyze AMR data from African nations, the research provides tools that are finely tuned to the continent's unique healthcare landscape. African researchers now have access to machine learning models that can predict MIC for antibiotics commonly used across the continent, such as Erythromycin, Penicillin, and Cefazoline for *S. aureus*. These models, which consider local variables like age, gender, and infection sites, enable researchers to understand regional resistance patterns more deeply. The study not only adds to the global scientific community but also empowers African researchers to lead in the fight against AMR by using data that reflects the specific challenges and dynamics of their own regions.

For Medical Personnel in Africa
The outcomes of this study are particularly relevant for healthcare providers across Africa, where antibiotic resistance is a growing concern. The AI models developed can significantly enhance the decision-making process in clinical settings by providing accurate predictions of antibiotic effectiveness. For instance, a clinician in South Africa or Kenya could use these models to determine the most appropriate antibiotic for treating an infection caused by Methicillin-resistant *S. aureus* (MRSA), a high-priority pathogen in the region. This targeted approach not only improves patient outcomes but also helps in conserving limited healthcare resources by reducing the misuse of antibiotics, which is a key driver of resistance. By implementing selective and cascade reporting based on local resistance patterns, African medical personnel can prioritize treatments that are most likely to succeed, ultimately improving public health outcomes on the continent.

For Policymakers in Africa

For African policymakers, the findings of this study provide crucial insights into how AI can be harnessed to combat AMR more effectively. With the ability to predict resistance trends, policymakers can proactively develop strategies that are tailored to the specific needs of their countries. For example, in regions where certain antibiotics are losing their effectiveness, governments can implement targeted awareness campaigns and adjust national treatment guidelines accordingly. The study also highlights the importance of investing in laboratory infrastructure, as only a small percentage of medical laboratories in Africa currently perform AMR susceptibility testing. By prioritizing such investments, African nations can improve their capacity to monitor and respond to AMR, ensuring that healthcare systems are better equipped to handle this growing threat. Ultimately, integrating AI into public health policy will not only strengthen Africa's health systems but also position the continent as a leader in the global fight against antimicrobial resistance.

CONCLUSIONS AND LIMITATION OF THE STUDY
This study explores the integration of AI with AMR susceptibility testing to tackle antibiotic resistance in Africa. The research demonstrates that AI models can predict MIC for antibiotics, offering valuable tools for improving antibiotic prescribing, infection control, and public health interventions. The study shows that AI can provide real-time insights into resistance patterns, enhancing patient outcomes and supporting evidence-based policy decisions.

However, the research faces significant challenges, including the limited availability of comprehensive AMR data in Africa. The small sample sizes, such as the 3,943 samples for *S. aureus*, may not fully capture the diversity of resistance patterns, affecting the accuracy of the AI models. Additionally, inconsistent data collection and reporting practices across the continent hinder the models' reliability and generalizability. The study emphasizes the need for standardized data collection, expanded laboratory capabilities, and increased investment in healthcare infrastructure to maximize the benefits of AI in combating AMR in Africa.

TABLES

Table 1. Root Mean Square Error (RMSE) of AI Methods (Random Forest, Partial Least Squares Regression, and XGBoost) in Predicting MIC for Tier 1 Erythromycin on *S. aureus*. The AI methods were trained on 75% of the data (2,957 samples) and tested on the remaining 25% (986 samples) that were not included in the training set. The table below shows the RMSE value for each method.

No	AI method	RMSE
1	Random Forest	2.4
2	Partial Least Square Regression	1.2
3	XGBoost	1.7

Table 2. Root Mean Square Error (RMSE) of AI Methods in Predicting MIC for Tier 2 Penicillin on *S. aureus*. The AI methods were trained on 75% of the data and tested on the remaining 25% that were not included in the training set. The table below shows the RMSE value for each method.

No	AI method	RMSE
1	Random Forest	1.9
2	Partial Least Square Regression	1.3
3	XGBoost	0.9

Table 3. Root Mean Square Error (RMSE) of AI Methods in Predicting MIC for Tier 3 Cefazolin on *S. aureus*. The AI methods were trained on 75% of the data and tested on the remaining 25% that were not included in the training set. The table below shows the RMSE value for each method.

No	AI method	RMSE
1	Random Forest	2.3
2	Partial Least Square Regression	2.0
3	XGBoost	1.8

REFERENCES

1. Gajic I, Kabic J, Kekic D, Jovicevic M, Milenkovic M, Mitic Culafic D, Trudic A, Ranin L, Opavski N. Antimicrobial Susceptibility Testing: A Comprehensive Review of Currently Used Methods. Antibiotics (Basel). 2022 Mar 23;11(4):427. doi: 10.3390/antibiotics11040427. PMID: 35453179; PMCID: PMC9024665.
2. https://healthpolicy-watch.news/less-than-2-of-laboratories-in-14-african-countries-are-unable-to-test-for-antimicrobial-resistance/
3. https://amr.vivli.org
4. https://www.who.int/initiatives/glass
5. Clinical and Laboratory Standards Institute (CLSI). *Performance Standards for Antimicrobial Susceptibility Testing. 34th ed.* CLSI supplement M100 (ISBN 978-1-68440-220-5 [Print]); ISBN 978-1-68440-222-2 [Electronic]). Clinical and Laboratory Standards Institute, USA, 2024.
6. https://www.who.int/publications/i/item/9789240093461
7. Perovic, Olga, et al. "Prevalence and Trends of Staphylococcus aureus Bacteraemia in Hospitalized Patients in South Africa, 2010 to 2012: Laboratory-Based Surveillance Mapping of Antimicrobial Resistance and Molecular Epidemiology." PLOS ONE, vol. 10, no. 12, 2015, e0145429. https://doi.org/10.1371/journal.pone.0145429
8. Opiyo SO, Nalunkuma R, Nanyonga SM et al. Empowering Global AMR Research Community: Interactive GIS dashboards for AMR data analysis and informed decision-making [version 1; peer review: 1 approved with reservations]. Wellcome Open Res 2024, 9:234
9. Li R, D'Arcy CV, Hermsen ED et al. Creative usage of Vivli's Antimicrobial Resistance Surveillance Open Data through the Re-Use Data Challenge [version 1; peer review: not peer reviewed]. Wellcome Open Res 2024, 9:235
10. Wildfire J, Waterlow NR, Clements A et al. MIC distribution analysis identifies differences in AMR between population sub-groups [version 1; peer review: 2 approved]. Wellcome Open Res 2024, 9:244 (https://doi.org/10.12688/wellcomeopenres.21269.1)
11. R Core Team (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.URL https://www.R-project.org/.