AMR Surveillance Open Data Re-use Challenge

Proposal Title: Analysis of variations in minimum inhibitory concentration distributions by patient groups

Date of Submission (dd-mmm-yy): 27-Jul-2023

Research Team Members details *(put the Lead Applicant 1st in the table)*:

Team Member Name	Role in the Data Challenge	Affiliation	Email	Country
Jacob Wildfire	Lead Applicant	LSHTM/SGUL	Jacob.wildfire1@lshtm.ac.uk	UK
Gwen Knight	Data analysis	LSHTM	Gwen.knight@lshtm.ac.uk	UK
Naomi Waterlow	Data Analysis	LSHTM	Naomi.waterlow1@lshtm.ac.uk	UK
Naomi Fuller	Data Analysis	LSHTM	Naomi.Fuller@lshtm.ac.uk	UK
Alastair Clements	Data Analysis	LSHTM	Alastair.Clements@lshtm.ac.uk	UK

Datasets included in the analysis (Tick all those that apply):

/	GSK – SOAR 201818
/	Johnson & Johnson – Bedaquiline Drug Resistance Assessment in MDR-TB (DREAM)
/	Paratek - KEYSTONE
/	Pfizer – ATLAS_Antibiotics
	Pfizer – ATLAS_Antifungals
/	Shionogi – SIDERO-WT
/	Venatorx – GEARS
	Other data (please provide details): _____

Minimum Inhibitory Concentration Analysis by Group (MICAG) Data Challenge entry
Title: Analysis of variations in minimum inhibitory concentration distributions by patient groups

Abstract: Phenotypic data, such as the minimum inhibitory concentrations (MICs) of bacterial isolates from clinical samples, are widely available through routine surveillance. MIC distributions inform antibiotic dosing in clinical care by determining cutoffs to define isolates as susceptible or resistant. However, differences in MIC distributions between patient populations could indicate strain variation and hence differences in transmission, infection or selection.

We exploit the Vivli AMR register's wealth of MIC and metadata in an innovative way to explore MIC distribution variation by key population groups such as age, sex and infection type across time. We developed a generalisable methodology with open-source code which characterises MIC distributions across user-defined groups to gain insights into different infecting strains and selection pressures.

We found clear differences between MIC distributions across various patient groups for a subset of bacteria-antibiotic pairings. For example, within *Staphylococcus aureus,* MIC distributions by age group and infection site displayed clear trends, especially for levofloxacin with higher resistance levels in older age groups, which appeared more often in men. This trend could reflect greater use of fluoroquinolones in adults than children but also reveals an increasing MIC level with age, suggesting either transmission differences or accumulation of resistance effects. We also observed high variations by income groups, regions, and over time, with the latter likely linked to changes in surveillance.

Such comparisons could reveal hidden transmission sources and effects of antibiotic use in different population groups and highlight opportunities to improve stewardship programmes and interventions, resulting in implementable impacts. Our methodology could be simple to apply in low-resource settings where raw MIC data is regularly generated but could be more efficiently exploited.

Introduction: The global burden of antimicrobial resistance (AMR) is growing, yet there remain huge unknowns concerning the drivers and dynamics of its spread. Due to a lack of understanding of whether antimicrobial use or transmission drives more AMR infections, we cannot optimally target interventions, with many stewardship interventions failing to impact AMR rates [1]–[3]. Evidence suggests that locally targeted interventions using local data to implement cost-effective solutions are urgently required, particularly in low-resource settings [4].

Minimum inhibitory concentration (MIC), defined as the lowest concentration of a chemical, usually a drug, which prevents visible *in vitro* growth of bacteria [5], [6], is one of the most routinely collected types of AMR surveillance data globally. Its primary use is to determine susceptibility phenotype (susceptible/resistant), used for clinical purposes and in AMR surveillance studies. MIC data, however, contains more information that could be useful in monitoring differences in resistance evolution and transmission between patient populations. Also, MIC is an absolute measurement (though there can be differences by measurement assay) and therefore allows for greater comparison over time than susceptibility phenotype, which is dependent on ever-developing break-point definitions [7]–[9].

To our knowledge, historical exploration of MIC data has been limited to examination of MIC "creep" (changes over time e.g. [10], [11]) and country-level comparisons [12]. One database, Pfizer's Antimicrobial Testing Leadership and Surveillance (ATLAS) programme [13], contains 17 years of cumulative global MIC and rich metadata. Catalán et al. (2022) used the ATLAS database to demonstrate that global changes in MICs for particular bug-drug combinations over time can be predicted using the dataset, identifying some gradual global increases (MIC creep) [14]. Similarly, Kenyon et al. (2019) used historical *Neisseria gonorrhoea* MIC distribution data from five countries to track the replacement of susceptible strains by increasingly resistant ones, raising questions about whether resistance is reversible [12]. These studies highlight the value of MIC analyses, in particular with ATLAS, and how spatiotemporal analyses of MIC distributions can be used to fill knowledge gaps regarding AMR evolutionary dynamics.

However, MIC distribution comparisons could be extended to explore more subtle and local differences in resistance evolution as MIC data is widely, rapidly and cheaply available. Differences in MIC distributions from bacteria isolated from different patient groups could suggest different transmission sources or more rapid resistance evolution in different groups. This could be driven by known differences in antibiotic effects or exposures between groups and could be used to measure the impact of stewardship measures that target groups differently. Moreover, it could reveal hotspots for interventions and allow for enhanced analysis of pathways to infection. For example, comparing MIC distributions between sites of infection may reveal that isolates from wound infections are distinct from those from bloodstream infections, suggesting a need to find other sources for infection and hence the correct point to intervene.

Preliminary work by our research group has found distinct resistance prevalence patterns by age and sex for different antibiotic–bacteria combinations at the national level, matching simple European-level comparisons [15]. This could be linked to a wide range of factors such as antibiotic exposure, contact patterns and infection incidence which vary by antibiotic and bacteria.

Here, we explore MIC distribution differences within a novel stratification framework to explore multinational demographic patterns across multiple longitudinal databases within the Vivli AMR register, including ATLAS. This framework could then be applied locally in future studies.

Aim: To develop and apply a method for comparing MIC distributions by population groups to suggest commonalities and variation in transmission, infection or selection.

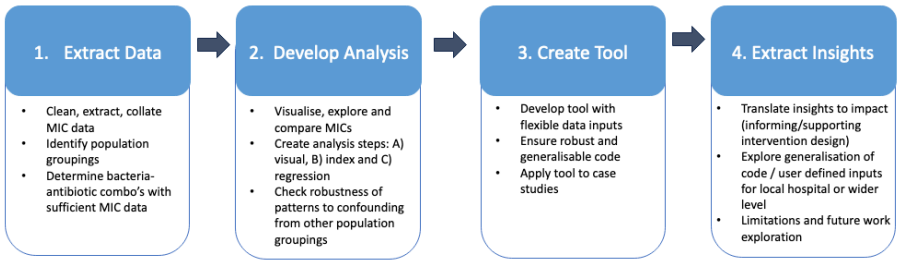

Figure 1: Objectives for MICAG data challenge exploration of MIC distribution differences.

Methods: Our analyses followed four main objectives (Figure 1). Code for data cleaning and analyses were written in *R* [16] and are available on our GitHub repository [17].

Databases and cleaning: We explored the six available databases in the Vivli AMR Register and incorporated their data in our analyses if they included (a) MIC values and (b) had metadata other than just country. This excluded the DREAM database. We combined the other five databases: ATLAS, SOAR 201818, KEYSTONE, SIDERO-WT and GEARS. The MIC data was cleaned by excluding all alpha-numeric values (often appearing to refer to genetic determinants, e.g. "OXA-"). Values with a qualifier (e.g. < or ≤) were set at the numeric value in the qualifier (e.g. <8 was set to 8). The data on "body site" or "source" was grouped from 177 unique locations into five key "infection site" types ("blood" / "urine" / "respiratory" / "gastrointestinal" / "wound") (assuming isolates come from infection), which covered 85% of the original data [17]. As the most extensive dataset (ATLAS) had age groups instead of age, we used these throughout (<=2, 3-12, 13-18, 19-64, 65-84, 85+).

Of the 398 unique species, over 277 had fewer than 1000 antibiotic susceptibility results. Due to limited time constraints, we focused on the four bacteria with more than 1.4 million results (*Staphylococcus aureus* (2.5M)*, Escherichia coli* (1.7M)*, Pseudomonas aeruginosa* (1.6M) and *Klebsiella pneumoniae* (1.4M)) as the next most represented bacteria had less than 750,000. 20 antibiotics had susceptibility results for all four bacteria. Our final dataset had 7,419,954 susceptibility results.

Our analyses consisted of 3 steps: visual analysis of differences between groups, an index calculation to identify bacteria-antibiotic-groups with substantial differences, and a regression analysis to explore confounding. The groups we investigated were: age group, sex, infection site, year, database,

income group, WHO region and country. Results of the first four groups are detailed in this report, while the others can be found on our GitHub repository [17].

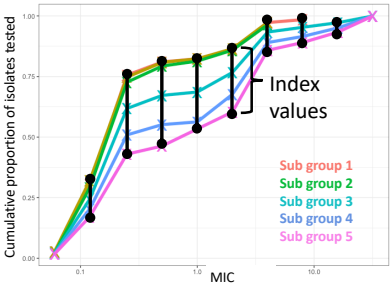

Figure 2: Mock example Index calculation for MIC distribution differences across sub groups

Visual analysis: We chose to use cumulative plots to allow for easy comparison with such variability of step changes in MIC (Figure 2). *Index analysis:* We developed an index (Figure 2) to allow for rapid comparison between groups which calculates the maximum difference in cumulative population proportions across groups at each MIC value. We defined a bacteria-antibiotic pair to have notable differences in MIC distributions by groups if they had 4 or more index values greater than 0.1 (i.e. at 4 or more MIC values there was a > 10% difference in cumulative population values across the groups). *Regression Analysis:* To determine if differences between groups were confounded by any other groups (e.g. sex by age etc.) we ran a proportional odds regression model.

Results: ***Case study: age and sex:*** MIC distributions within isolates of *S. aureus, E. coli, K. pneumoniae* and *P. aeruginosa* varied substantially by age (Figure 3A&C). Of those bacteria-antibiotic combinations with differences in MIC distributions by age and sex, most were in *S. aureus* (50% across all data), followed by *P. aeruginosa*. We highlight three antibiotic combinations across two bacteria to demonstrate the MIC differences by group. Ampicillin does not have notable trends by age and sex (Figure 3) in either *S. aureus* or *E. coli*, whereas levofloxacin-*S. aureus* shows the strongest differences.

Case study: infection site: Clear MIC distribution differences could be seen between isolates taken from different infection sites (Figure 3B&D). Again, we saw bigger differences in the MIC distributions by infection site for *S. aureus*, followed by *K. pneumoniae* and *E. coli,* with few in *P. aeruginosa* (Figure 3B). A clear difference in MIC distributions can be seen for ampicillin resistance in *S. aureus* with higher numbers of isolates at 1mg/L for "blood" and "gastro" isolates vs "respiratory", "urine" and "wound".

Case study*: *high level spatial and time variation: Substantial differences were seen in the MIC distributions across income groups & WHO regions (data not shown, [17]). Tracking the index over time for all bacteria-antibiotic combinations (Figure 4A) showed no clear trend in index with time. However clear differences in MIC distribution by age and sex could be seen for some bacteria-antibiotic pairs at different time points. For levofloxacin resistance in *S. aureus,* there was substantial long-term data (Figure 4A) with differences in MIC distribution by age over time. Still, frequent data changes make interpretation difficult: the number of isolates varied substantially over time (Figure 4B).

Case study: analysis of confounding: In the above analyses, some individual-level variables (age, sex, infection site, space and time) could interact to confound true relationships - particularly in the analyses by age group and time. To address this, we demonstrate the application of regression analysis for levofloxacin resistance in *S. aureus* (Table 1). This showed a strong impact of age with increased MIC value, ranging from odds of 0.559 (ages 3-12) to 2.067 (ages 85+) compared to adults (age 19-64), a small but significant impact of sex (male odds 1.053) and a varied impact of infection site (highest odds in urine) (Table 1).

Variable	Value	Std. Error	p value	Odds
Age: 0 to 2 Years	-0.469	0.027	0	0.626
Age: 3 to 12 Years	-0.512	0.029	0	0.599
Age: 13 to 18 Years	-0.412	0.035	0	0.663
Age: 65 to 84 Years	0.393	0.015	0	1.481
Age: 85 and Over	0.726	0.026	0	2.067
Sex: M	0.051	0.013	0	1.053
Infection site: blood	-0.032	0.024	0.176	0.968
Infection site: gastro	-0.136	0.037	0	0.873
Infection site: respiratory	0.251	0.021	0	1.285
Infection site: urine	0.468	0.041	0	1.598
Infection site: wound	0.027	0.02	0.17	1.028

Table 1: *Results of regression for levofloxacin resistance in S. aureus. Base categories were Adults (age 19-64), Female and unknown sources Note abbreviations: "gastro = "gastrointestinal", "M" = "male".*

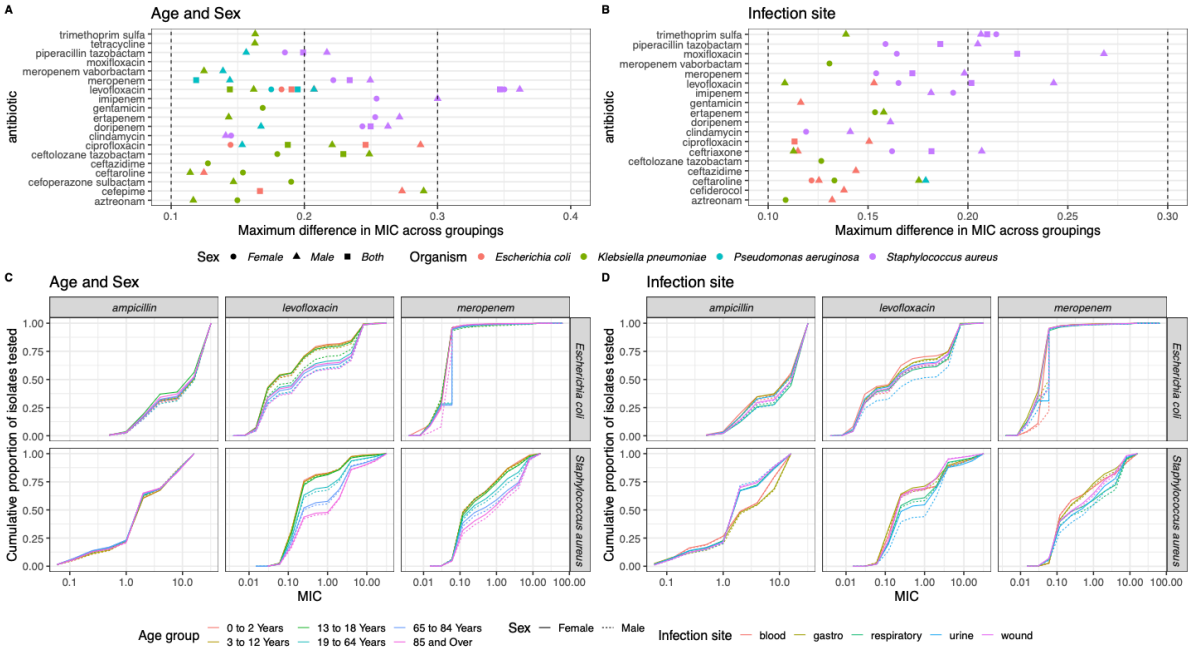

Figure 3: *Examples of variation in MIC distribution by age (left) and infection type (right). A&B: The bacteria-antibiotic pairs with notable differences in MIC distributions (defined as a difference between cumulative population values of more than 0.1 for more than three MIC values). C&D: Plots of cumulative sum of isolates tested by MIC for example, bacteria-antibiotic combinations highlight the variation in the MIC distribution by age (C), infection type (D) and sex (linetype). Only isolates from the five labelled infection sites were included in the index calculation. Note abbreviations: "gastro = "gastrointestinal".*

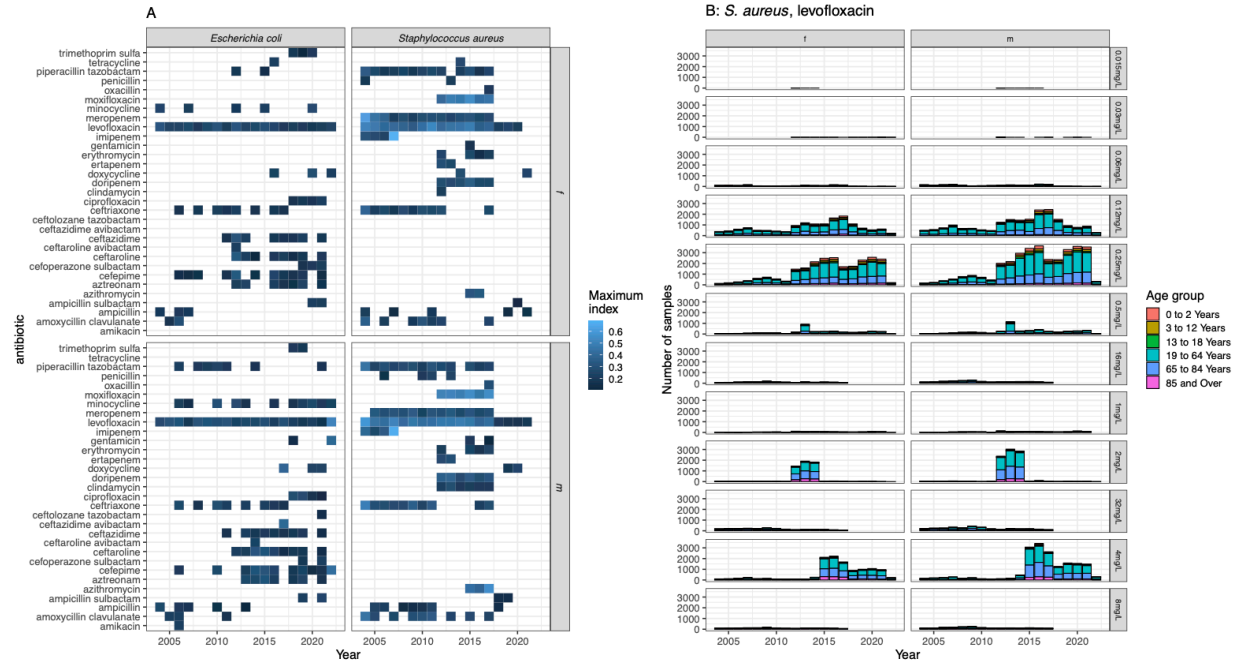

Figure 4: High variability in time is likely linked to changes in data surveillance. A: Tracking the index over time for all bacteria-antibiotic combinations by age group and sex (row) showed a high variation in availability of data and no clear trend in index with time. B: For levofloxacin resistance in S. aureus clear differences in several isolates tested at each MIC (panels) within each age group (colour) and sex (column) over time.

Conclusions: MIC distributions have been used to explore differences in resistance levels between countries and over time. Here we show that they can be used to further explore differences between patient groups (e.g. age groups). For many bacteria-antibiotic combinations, we found substantial differences in MIC distributions between patient groups, identified and confirmed through index value calculation and data visualisation. All four analysed bacteria showed substantial variation in MIC distributions by age and infection site for multiple antibiotics.

Most notably, large variation exists between the MIC distributions of *S. aureus*, which we highlight for the fluoroquinolone levofloxacin across age and infection site. This could be linked to the limited use of quinolones in children due to safety concerns [18], but then we might expect a step change between children and adults, not the observed continual differential increase with age (Table 1). Levofloxacin is not usually used to treat *S. aureus* infections, so whilst these patterns may not be a priority for clinical care, they still reveal patterns that point to either transmission or selection variation by age. This could be used to design age-structured stewardship patterns or provide better understanding of infection pathways.

For our analysis we combined multiple large multinational datasets in the longitudinal Vivli AMR register. These varied both in terms of which isolates were tested (we assumed for simplicity these were all infecting rather than colonising isolates) and the way testing was conducted (e.g. to calculate MIC) [6], [11]. This is a major limitation of routine surveillance and limits the applicability of our results and usage of the Vivli AMR register. We also made the simplifying assumption that we could remove character symbols (e.g. "<" and ">=") from MIC measurements. There will likely also be strong sampling biases in the data that we are unable to account for, due to the different aims for which the data was originally collected. Using multiple Vivli AMR register datasets, instead of focusing on one, helped us to counter this bias to an extent. Overall, however, our analyses point to a new dimension for which existing routine AMR data can be harnessed to understand differences between patient groups and to test hypotheses.

A key step in overcoming these limitations would be to apply our methodology to locally collected data and verify if the patterns seen in these global data hold at the local level. Exploring MIC distribution differences using the Vivli AMR register datasets has enabled us to demonstrate population group differences and to develop a framework that should now be used to complement local analyses. By providing a baseline, these open data are a key resource for AMR researchers, which do not require sometimes difficult negotiations with clinical settings and added data curation burden. At a local level, comparisons between MIC distributions and those in the Vivli AMR register data could point to transmission source differences (e.g. between wards) or to where AMR evolution is accelerated.

In the future, discussion with healthcare epidemiologists, pharmacists and clinicians will help us to better understand the most relevant patterns for intervention and stewardship as our team consisted of mathematical modellers and microbiologists. Our open-source code requires further adaptation and could be developed into a more user-friendly tool but is already flexible for other patient group stratifications. Applying our framework to local data would also produce more robust, actionable conclusions that can account for sampling frameworks and enhance routinely collected MIC data.

Impact: Due to the averaging of AMR data down to discrete, clinical decision-making thresholds, there has been a lack of analysis at the ecological level of AMR diversity. In this project, we exploit the data available through the Vivli AMR register to take the **innovative** step of exploring MIC distributions by different patient groups. We demonstrated that for certain bacteria, different strains are likely circulating in different age groups. This work is **generalisable** as it could be applied to any user-defined grouping, e.g. hospital setting. It could be used to explore MIC distribution variation to gain insight into phenotypic variation and hence variation in transmission and selection across wards, hospitals or at the national level. The **impact** would be to exploit the differences or similarities in MIC distributions to say where and when interventions could impact resistance evolution, antibiotic selection or transmission. This could be differential targeting of interventions by subgroups or differential antibiotic stewardship guidelines.

References:

[1] M. Akpan, R. Ahmad, N. Shebl, and D. Ashiru-Oredope, 'A Review of Quality Measures for Assessing the Impact of Antimicrobial Stewardship Programs in Hospitals', *Antibiotics*, vol. 5, no. 1, p. 5, Jan. 2016, doi: 10.3390/antibiotics5010005.

[2] E. S. F. Orubu *et al.*, 'Assessing Antimicrobial Resistance, Utilization, and Stewardship in Yemen: An Exploratory Mixed-Methods Study', *Am. J. Trop. Med. Hyg.*, vol. 105, no. 5, pp. 1404–1412, Nov. 2021, doi: 10.4269/ajtmh.21-0101.

[3] E. J. Septimus, 'Antimicrobial Resistance', *Med. Clin. North Am.*, vol. 102, no. 5, pp. 819–829, Sep. 2018, doi: 10.1016/j.mcna.2018.04.005.

[4] J.-B. Ronat *et al.*, 'AMR in low-resource settings: Médecins Sans Frontières bridges surveillance gaps by developing a turnkey solution, the Mini-Lab', *Clin. Microbiol. Infect.*, vol. 27, no. 10, pp. 1414–1421, Oct. 2021, doi: 10.1016/j.cmi.2021.04.015.

[5] B. Kowalska-Krochmal and R. Dudek-Wicher, 'The Minimum Inhibitory Concentration of Antibiotics: Methods, Interpretation, Clinical Relevance', *Pathogens*, vol. 10, no. 2, p. 165, Feb. 2021, doi: 10.3390/pathogens10020165.

[6] J. M. Andrews, 'Determination of minimum inhibitory concentrations', *J. Antimicrob. Chemother.*, vol. 48 Suppl 1, pp. 5–16, Jul. 2001, doi: 10.1093/jac/48.suppl_1.5.

[7] 'eucast: Clinical breakpoints and dosing of antibiotics'. https://www.eucast.org/clinical_breakpoints (accessed Jul. 26, 2023).

[8] 'Updating Breakpoints in Antimicrobial Susceptibility Testing', *ASM.org*. https://asm.org:443/Articles/2022/February/Updating-Breakpoints-in-Antimicrobial-Susceptibili (accessed Jul. 26, 2023).

[9] W. M. Johnson, J. A. Clark, K. Olney, D. R. Burgess, and D. S. Burgess, 'Changing times: The impact of gram-negative breakpoint changes over the previous decade', *Antimicrob. Steward. Healthc. Epidemiol. ASHE*, vol. 2, no. 1, p. e165, Oct. 2022, doi: 10.1017/ash.2022.301.

[10] B. Jiang *et al.*, 'A 5-year Survey Reveals Increased Susceptibility to Glycopeptides for Methicillin-Resistant Staphylococcus aureus Isolates from Staphylococcus aureus Bacteremia Patients in a Chinese Burn Center', *Front. Microbiol.*, vol. 8, p. 2531, 2017, doi: 10.3389/fmicb.2017.02531.

[11] B. Edwards, K. Milne, T. Lawes, I. Cook, A. Robb, and I. M. Gould, 'Is vancomycin MIC "creep" method dependent? Analysis of methicillin-resistant Staphylococcus aureus susceptibility trends in blood isolates from North East Scotland from 2006 to 2010', *J. Clin. Microbiol.*, vol. 50, no. 2, pp. 318–325, Feb. 2012, doi: 10.1128/JCM.05520-11.

[12] C. Kenyon, J. Laumen, D. Van Den Bossche, and C. Van Dijck, 'Where have all the susceptible gonococci gone? A historical review of changes in MIC distribution over the past 75 years', *BMC Infect. Dis.*, vol. 19, no. 1, p. 1085, Dec. 2019, doi: 10.1186/s12879-019-4712-x.

[13] Pfizer, 'Antimicrobial Testing Leadership and Surveillance (Atlas)'. https://atlas-surveillance.com/ (accessed Jul. 26, 2023).

[14] P. Catalán, E. Wood, J. M. A. Blair, I. Gudelj, J. R. Iredell, and R. E. Beardmore, 'Seeking patterns of antibiotic resistance in ATLAS, an open, raw MIC database with patient metadata', *Nat. Commun.*, vol. 13, p. 2917, May 2022, doi: 10.1038/s41467-022-30635-7.

[15] OECD, 'Stemming the Superbug Tide: Just A Few Dollars More', OECD Health Policy Studies, OECD Publishing, Paris, 2018. [Online]. Available: https://doi.org/10.1787/9789264307599-en.

[16] R Core Team, ‘R: A language and environment for statistical computing’. R Foundation for Statistical Computing, Vienna, Austria, 2021. [Online]. Available: https://www.R-project.org/.
[17] N. Waterlow, A. Clements, and G. Knight, ‘MICAG source code’. Jul. 2023. [Online]. Available: https://github.com/aj-clements/Vivli-AMR-KG
[18] K. Patel and J. L. Goldman, ‘Safety Concerns Surrounding Quinolone Use in Children’, *J. Clin. Pharmacol.*, vol. 56, no. 9, pp. 1060–1075, Sep. 2016, doi: 10.1002/jcph.715.