

AMR Surveillance Open Data Re-use Challenge

Proposal Title: Novel approach to analysis of AMR: looking at the composite resistance phenotype.

Date of Submission (dd-mmm-yy): 30th July 2023

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

Team Member Name	Role in the Data Challenge	Affiliation	Email	Country
Shraddha Karve	Conceptualization of subtype analysis, formulation of EOI and final report	Trivedi School of Biology, Ashoka University	shraddha.karve@ashoka.edu.in	India
Rintu Kutum	Data analysis and final report	Department of Computer Science, Ashoka University	rintu.kutum@ashoka.edu.in	India
Devojit Sarma	Conceptualization of climatic correlations and final report	ICMR, NIRE H	devojit.sarma@icmr.gov.in	India
Vasundhara Karthikeyan	Data wrangling and analysis	Trivedi School of Biology, Ashoka University	vasundhara.karthikeyan@ashoka.edu.in	India
Ragul N	Data wrangling and analysis	Department of Computer Science, Ashoka University	ragul.n_ug23@ashoka.edu.in	India

Add more rows if needed

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

<input type="checkbox"/>	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): _____

Objectives

The escalating challenge of antimicrobial resistance (AMR) poses a significant global concern for public healthcare systems^{1,2}. Current AMR surveillance and molecular mechanism studies traditionally focus on specific drug-bug combinations, like carbapenem-resistant *Klebsiella pneumoniae*, designated as a priority pathogen by the WHO³. While wastewater and environmental surveillance aim to detect *Klebsiella* species and genes conferring carbapenem resistance⁴, it is known that resistance genes for one antibiotic often coexist with genes for resistance to others⁵. To address these complexities, we propose a novel analysis approach using *Klebsiella pneumoniae* as a model. We consider the resistance profile of an isolate for a set of common antibiotics across two datasets, ATLAS and GEARS. We term this composite phenotype, encompassing resistance/sensitivity to a group of antibiotics, a 'subtype' of the pathogen. Our primary objective is to track and study the prevalence of different subtypes across time and space, enabling a more comprehensive understanding of AMR dynamics. We then explore the impact of climatic parameters on the prevalence of different *Klebsiella pneumoniae* subtypes, aiming to uncover additional insights into antibiotic resistance patterns. Rising temperatures and climate change have been associated with recent antibiotic resistance developments, as bacterial growth and genetic material dissemination are closely tied to temperature conditions⁶. Heavy rainfall has been linked to bacterial mutagenesis and antibiotic resistance gene expression⁷. Rising local temperatures in the United States and Europe have shown correlations with increased antibiotic resistance at the population level in various pathogens^{8,9}.

Methods:

1) Generating unique subtypes and their frequencies for different datasets

We chose a priority pathogenic species, *Klebsiella pneumoniae* as an example for the kind of analysis we propose. We used Pfizer-Atlas and Venatorx datasets which included data on *K. pneumoniae* from diverse specimens like blood, urine, pus, etc. The MIC values in the ATLAS and the GEARS datasets were interpreted using the CLSI 2020 and 2021 breakpoint criteria respectively. To ensure consistency, we evaluated data from 2018 to 2021 for the subtype study, allowing for a comprehensive comparison of results between the two datasets.

We next shortlisted the antibiotics which were common for *K. pneumoniae* across the two chosen datasets. These were cefepime, ceftazidime, gentamicin, levofloxacin, piperacillin-tazobactam. Antibigram of every isolate were then represented by a string of five 'R' (for resistant) or 'S' (for sensitive) denoting the resistance status of that isolate for that antibiotic. We call every unique string of R and S as a subtype.

2) Testing the impact of climatic parameters on the subtype prevalence

We chose five different climatic parameters, two were the proxy of precipitation (daily average rainfall in mm and percent relative humidity) and other three were the proxy of temperature (maximum, minimum and mean temperature). Daily average values of rainfall (mm), maximum, minimum and mean temperature (°C), and RH (%), all measured at 2 metres from ground level, were collected from 1st January 2018 to 31st December 2021 and yearly average values were estimated. The parameters were extracted from the National Aeronautics and Space Administration (NASA) Langley Research Center Prediction of Worldwide Energy Resource Project (NASA. POWER LARC Data Access Viewer, Available at <https://power.larc.nasa.gov/data-access-viewer/> accessed on 29 date 07 Month 2023, which is based on Modern-Era Retrospective analysis for Research and Applications version 2 (MERRA-2) dataset merged with Goddard Earth Observing System Model (GEOS) version 5.12.4. Using the API (<https://power.larc.nasa.gov/docs/tutorials/service-data-request/api/>) we retrieved the data using the Python request module.

The climate parameters namely, relative humidity, specific humidity and precipitation were \log_{10} transformed and later were used for regression analysis.

3) Multivariate Regression analysis

Regression analysis was performed on the prevalence of thirty-two subtypes of *Klebsiella pneumoniae* across 63 countries from 2018 to 2021 with six climate variables (768 regression analysis in total). P-value less than 0.05 was considered statistically significant for a given regression model.

All the analysis was done using Python (python 3.1.12) and R statistical programming language (R 4.2.1). For reproducibility of the work we have used the Conda environment. Data analysis and visualization in Python was done using statsmodels (0.14.0), pandas (1.5.3), matplotlib (3.7.1), seaborn (0.12.2) and geopandas (0.9.0) modules . Data visualization in R was done using ggplot2 (3.4.0), ggpubr (0.4.0) and RColorBrewer (1.1.3) packages.

Results:

Antibiotic subtype shows significant association with climate variables:

The subtype study carried out using the merged ATLAS and GEARS dataset and focusing on five antibiotics (cefepime, ceftazidime, gentamicin, levofloxacin, and piperacillin-tazobactam) identified 33 distinct subtypes of *Klebsiella pneumoniae* between 2018 and 2021 (Table 1 and Figure 1). Notably, the analysis highlighted that the extreme subtypes, RRRRR (~13%) and SSSSS (~55%), accounted for approximately 70% of the isolates (Table 1). We then explored the relationship between the regional patterns (averaged) of the most prevalent subtypes in the merged dataset (RRRRR and SSSSS) and the global climatic variables including precipitation, relative humidity (RH) and temperature (mean, minimum and maximum temperature) averaged over 2018-2021. Precipitation and relative humidity showed higher correlations with many of the subtypes than the temperature variables over the years (Figure 2). Owing to the consistent strength of correlation of precipitation and relative humidity with the most prevalent subtypes (RRRRR and SSSSS) across the years, we focused our analysis on the association between these parameters and the predominant subtypes.

Table 1: Prevalence of 13 subtypes of *K. pneumonia* from 2018 to 2021.

Subtype	2018 (%)	2019 (%)	2020 (%)	2021 (%)
SSSSS	55.79	48.23	56.32	56.66
SSSSR	2.39	2.88	3.24	2.65
SRSSS	1.85	1.25	1.83	2.99
SRSRR	0.39	0.59	0.42	0.94
SRRSS	0.69	0.72	0.92	1.11
RSSRR	0	0	0.17	0.26
RSRRR	0.08	0.07	0.08	0.09
RRSSS	3.32	2.69	2	1.88

RRSSR	0.77	1.11	0.92	1.37
RRSRS	1.77	2.36	2	0.85
RRSRR	5.56	7.6	5.32	7.08
RRRSS	2.39	2.23	1.91	1.19
RRRRR	11.96	14.94	11.56	13.99
Total isolates (N)	2592	3052	2404	2344

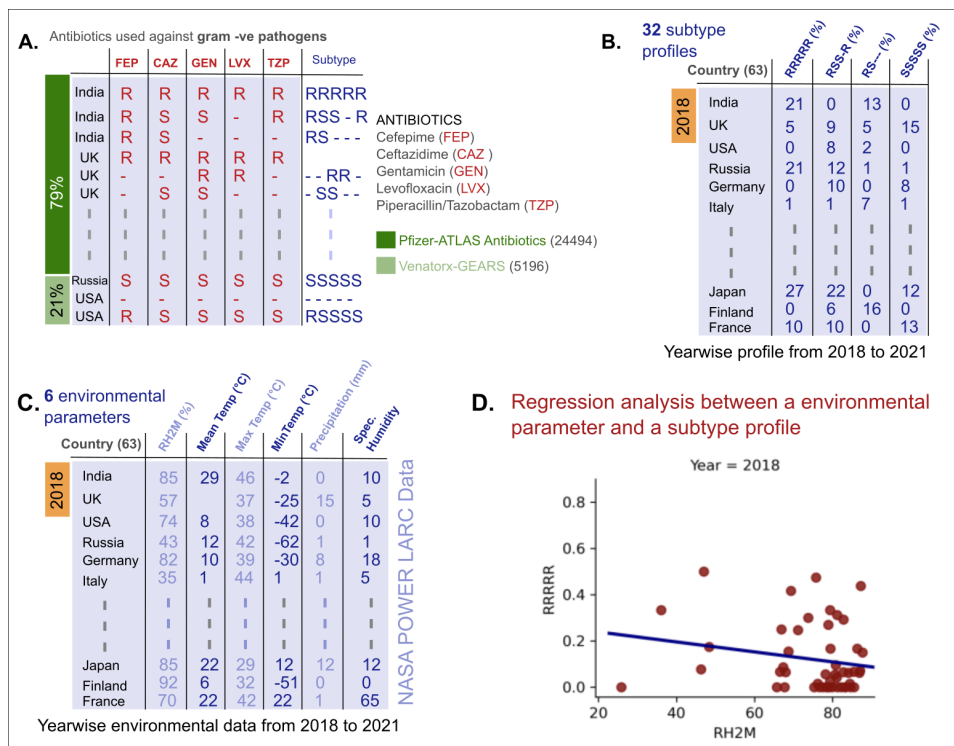


Figure 01: Complete workflow of integration of AMR data with environmental parameters. (A) Illustration showing conversion for gram negative pathogens antibiotics data (FEP, CAZ, GEN, LVX, and TZP) to subtype for clinical isolates for different countries, (B) Illustration showing year wise (2018) prevalence of 32 subtypes for the given antibiotics, (C) Illustration showing 6 environmental parameters across countries for a given year (2018), (D) Regression analysis showing relationship between RRRRR subtype prevalence with relative humidity (RH2M).

Figure 02: Bubble plot representing significant association post regression analysis across different subtypes of *K. pneumoniae* and global climatic parameters from 2018 to 2021 using ATLAS and GEARS data. The color intensity represents the R-squared value and size of the point represents the $-\log_{10}(\text{p-value})$, where bigger size implies statistical significance.

Association of relative humidity with all-susceptible-subtype (SSSSS) and all-resistance-subtype (RRRRR)

Post the multivariable regression analysis, we found that there is a strong association between the all-susceptible subtype and the relative humidity (Figure 3) and this was consistent across all the regions. We see that there is a significant inverse relationship between RH and SSSSS subtype with time (2021). We saw a similar trend between RH and RRRRR subtype across all the regions. With 2020, showing a significant inverse relationship between the RRRRR subtype and RH (Figure 4).

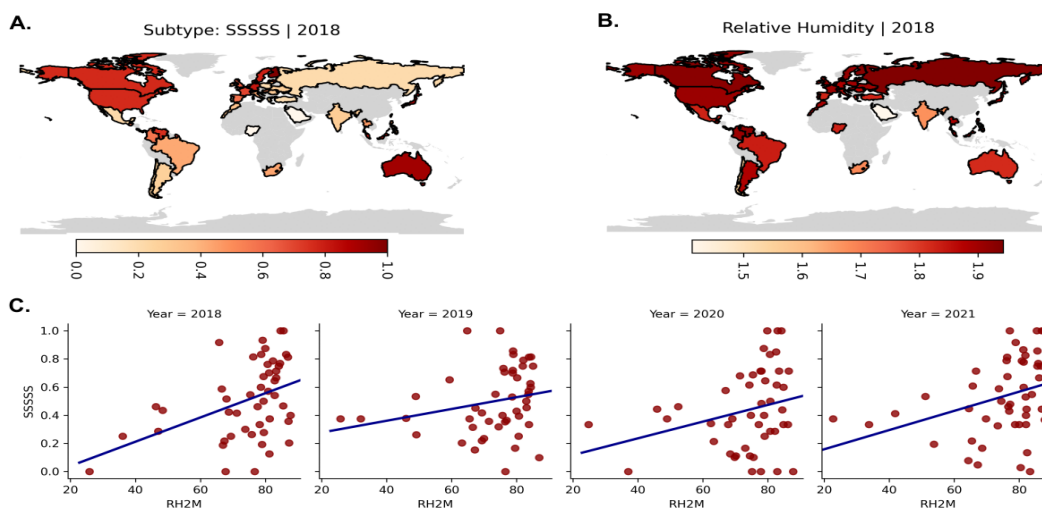


Figure 03: Regression analysis between all-susceptible-subtype prevalence (0-1) with relative humidity (\log_{10} transformed). Heatmap showing (A) prevalence of all-susceptible-

subtype (SSSSS) and (B) $\log_{10}(\text{relative humidity})$ for 2018 across 63 countries. (C) Scatterplot between SSSSS subtype and relative humidity (RH2M) with the regression line from 2018 to 2021 (for 2018 and 2021 corresponding p-values were less than 0.05).

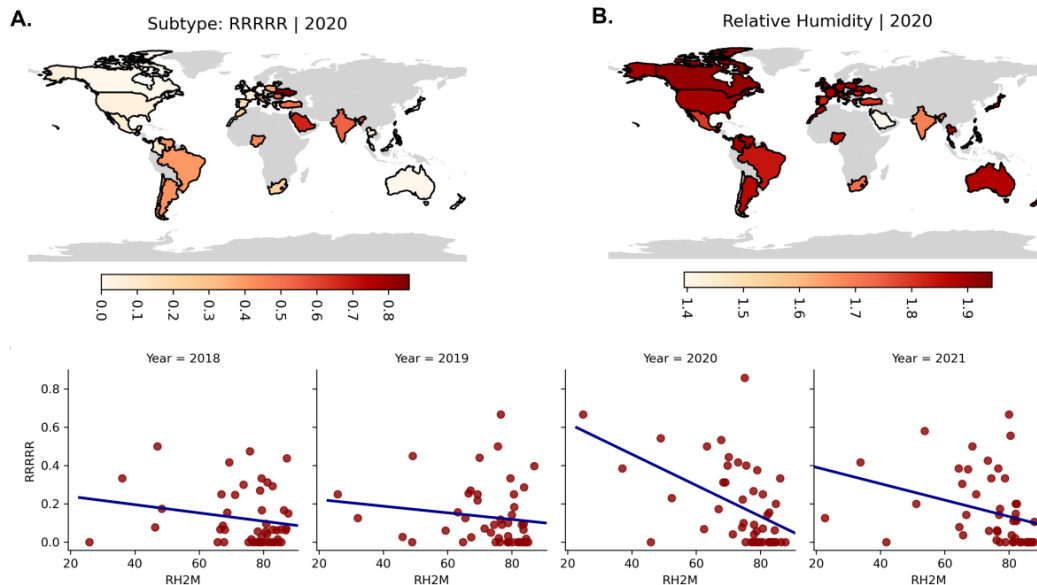


Figure 04: Regression analysis between all-resistance-subtype prevalence (0-1) with relative humidity (\log_{10} transformed). Heatmap showing (A) prevalence of all-resistance-subtype (RRRRR) and (B) $\log_{10}(\text{relative humidity})$ for 2018 across 63 countries. (C) Scatterplot between RRRRR subtype and relative humidity (RH2M) with the regression line from 2018 to 2021 (for 2020 corresponding p-value was less than 0.05).

Impact of the work:

Majority of the AMR studies consider a drug-bug combination as the standard unit of analysis. We propose extending this approach to include the resistance status against all the relevant antibiotics. This proposed 'subtype analysis' is more relevant for molecular mechanisms of resistance as well as resistance evolution as it considers the entire resistance phenotype of the isolate. One logistic limitation of this approach is that only those antibiotics may be considered which are common across the different datasets or specimens that one aims to study.

The analysis of spatiotemporal interaction between prevalent subtypes with climate variables is an example of how subtype analysis can be integrated into a one-health paradigm. It can potentially provide important insights into the intricate relationships that lead to the transmission and persistence of antibiotic resistance in the environment. This understanding can guide targeted therapies and strategies to address the multifaceted challenge of antibiotic resistance while also preserving antibiotic efficacy for future

generations. Future work in this area can focus on defining the most relevant proxies for climatic parameters.

References

- Burnham, J. P. (2021). Climate change and antibiotic resistance: a deadly combination. *Therapeutic Advances in Infectious Disease*, 8. <https://doi.org/10.1177/2049936121991374>
- Frost, I., Kapoor, G., Craig, J., Liu, D., & Laxminarayan, R. (2021). Hot Topic Status, challenges and gaps in antimicrobial resistance surveillance around the world. *Journal of Global Antimicrobial Resistance*, 25, 222–226. <https://doi.org/10.1016/j.jgar.2021.03.016>
- Icmr. (2021). *Annual Report 2021: Antimicrobial Resistance Research and Surveillance Network*.
- MacFadden, D. R., McGough, S. F., Fisman, D., Santillana, M., Brownstein, J. S., & Clim Chang Author manuscript, N. (2018). *Antibiotic Resistance Increases with Local Temperature*. 8(6), 510–514. <https://doi.org/10.1038/s41558-018-0161-6>
- McGough, S. F., MacFadden, D. R., Hattab, M. W., Mølbak, K., Santillana, M., Sarah, M. F., Derek, M. R., Mohammad, H. W., Kåre, M., & Mauricio Rates, S. (n.d.). *Rates of increase of antibiotic resistance and ambient temperature in Europe: a cross-national analysis of 28 countries between 2000 and 2016*. 1. <https://doi.org/10.2807/1560>
- Murray, C. J., Ikuta, K. S., Sharara, F., Swetschinski, L., Robles Aguilar, G., Gray, A., Han, C., Bisignano, C., Rao, P., Wool, E., Johnson, S. C., Browne, A. J., Chipeta, M. G., Fell, F., Hackett, S., Haines-Woodhouse, G., Kashef Hamadani, B. H., Kumaran, E. A. P., McManigal, B., ... Naghavi, M. (2022). Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet*, 399(10325), 629–655. [https://doi.org/10.1016/S0140-6736\(21\)02724-0](https://doi.org/10.1016/S0140-6736(21)02724-0)
- Pal, C., Bengtsson-Palme, J., Kristiansson, E., & Larsson, D. G. J. (2015). Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. *BMC Genomics*, 16(1), 1–14. <https://doi.org/10.1186/S12864-015-2153-5/FIGURES/9>

Ratkowsky, D. A., Olley, J., McMeekin, T. A., & Ball, A. (1982). Relationship between temperature and growth rate of bacterial cultures. *Journal of Bacteriology*, 149(1), 1–5. <https://doi.org/10.1128/JB.149.1.1-5.1982>

WHO publishes list of bacteria for which new antibiotics are urgently needed. (n.d.). Retrieved July 31, 2023, from <https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed>