2024 AMR Surveillance Data Challenge

Proposal Title: Can artificial intelligence models be trusted? Determining the variables for building accurate prediction models for antimicrobial resistance

Date of Submission (dd-mmm-yy): 14-Sep-24

Is the team entering for the AMR Student Innovation Award? (Yes/No) : No

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

Team Member Name	**Role in the Data Challenge**	**Affiliation**	**Email**	**Country**	**Are they a student? Yes/No**
Dr. Neha Nityadarshini	Conceptualisation, writing original draft, review and editing	All India Institute of Medical Sciences, New Delhi, India	nehanitya28@gmail.com	India	No
Dr. Jaya Biswas	Writing original draft, data curation, review and editing	All India Institute of Medical Sciences, New Delhi, India	dr.jayabiswas0202@gmail.com	India	No
Dr. Tamanna Bordoloi	Data curation, review and editing	All India Institute of Medical Sciences, New Delhi, India	tamannabordoloi@gmail.com	India	No

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

	GSK – SOAR 201818
	GSK – SOAR 201910
	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

Artificial Intelligence (AI) prediction models use machine learning algorithms and statistical techniques based on historical data to predict future outcomes. Multiple recent studies have highlighted the utility and potential of AI in the field of antimicrobial resistance (AMR) [1,2]. Before implementing a prediction model in a real-world patient care setting, we need to ensure the accuracy of the models. Since AI models thrive on data, it is easy to assume that a large volume of data will build robust models but owing to the complexity of AMR patterns which keep evolving with time, a model built on recent data might be more accurate, even if the volume of data is relatively less. Our aim is to build prediction models based on both older and recent data and compare their prediction results with true results. In addition, we have designed a prototype web application which can predict susceptibility patterns. We have tried to keep a minimalistic user interface and make the prediction results more objective, accompanied by a % probability, keeping in mind their application in tertiary care centres, especially in resource-limited settings.

Methods

We used the Pfizer-ATLAS antibiotics dataset as it had the highest number of bacterial isolates data, with an extensive global coverage. We chose a set of 19 commonly used antibiotics for analysis. For each drug, we used an online AI analytics platform to design 2 prediction models [3], which can predict the susceptibility result based on species, source, speciality and patients' demographic data like age, sex, country etc. One model was based on previous 10 years data i.e. 2012 to 2021, and the other on previous 2 years data i.e. 2020 and 2021. Then we used these models to predict the susceptibility results of the isolates of year 2022. In total, we built and analysed 38 models, 2 each for the 19 drugs. Finally, we compared true susceptibility results with predicted results of both models. We categorised all the concordant results as no errors (NE). Total errors (TE) were further categorised into very major error (VME), major error (MaE) and minor error (MiE) [4]. Detailed workflow is depicted in figure 1. As an example, we have used the drug amikacin to explain the process. Same workflow was followed for all the drugs.

Results

In total, we built and analysed 38 models i.e. 2 models each for the 19 drugs (Figure 2 and Table 1). Concordance rates ranged from 99% (Linezolid-2 and Linezolid-10) to 72% (Levofloxacin-2). 2-year model performed either similar or better than 10-year model in all antibiotics except cefoxitin. Reason for this aberrant result could have been due to very less amount of data available in the dataset. In case of amoxi-clav, the concordant rate of 2-year model was 80% while it was 73% for 10-year model. Except cefoxitin and amoxy-clav, there was no significant difference between concordance rates of 10-year and 2-year model. Rate of VME were highest in clindamycin, pip-taz, co-trimoxazole, ceftazidime and levofloxacin. Overall, best concordance rates were observed for linezolid, vancomycin, teicoplanin, tigecycline and ceftazidime-avibactam. We have designed the web app as a prototype for these 5 drugs since probability of correct prediction are the best for them.

Conclusion

Looking at the results, its clear that AI models need thorough validation before implementation. A model might be accurate for one drug while having an unacceptable error rate for another. Since 2-year model performed marginally better, models based on recent data should be given priority. Building and implementation of prediction models should follow a dynamic system in which models are freshly built and validated in regular intervals rather than building a single model and relying on it for a longer duration.

Impact of the work

1. **Contribution to antimicrobial stewardship**: Typical timeline of microbiological tests starts with receiving sample on day 1, identification on day 2 and AST results on day 3. Through the web app, AI prediction models built in this study can be used to predict the susceptibility results on day 2, thus aiding in de-escalation of empirical therapy and shift to a more targeted therapy till the AST reports are available
2. **Designing antibiograms**: Above prediction models can be effectively used to design antibiograms, especially in newer health care facilities where enough data is not yet available.
3. **Aid in resource limited settings**: Resource limited settings sometimes face unavailability of drugs/media/technical expertise/facilities to perform susceptibility testing, especially for newer drugs. Prediction models can assist in decision making until resources are available or actual result is obtained from reference laboratories.
4. **Gaps identified; scope for improvement:** While performing the study, we identified several areas where a small change can greatly improve the process of building prediction models. We have enumerated a few of them along with suggestions to improve.
    a. **A universal format for datasets:** All datasets for AMR related information should have a uniform format. A single uniform format will make inter-dataset comparison extremely convenient and hence broaden the horizon for prediction models.
    b. **AI-AMR Database:** A dedicated AI-AMR database should be available which maintains a repository of predicted outcomes and details of AI models. This will enable us to monitor error rates and decide which model to apply.

Link for the web application: www.findamr.net

Tables and figures

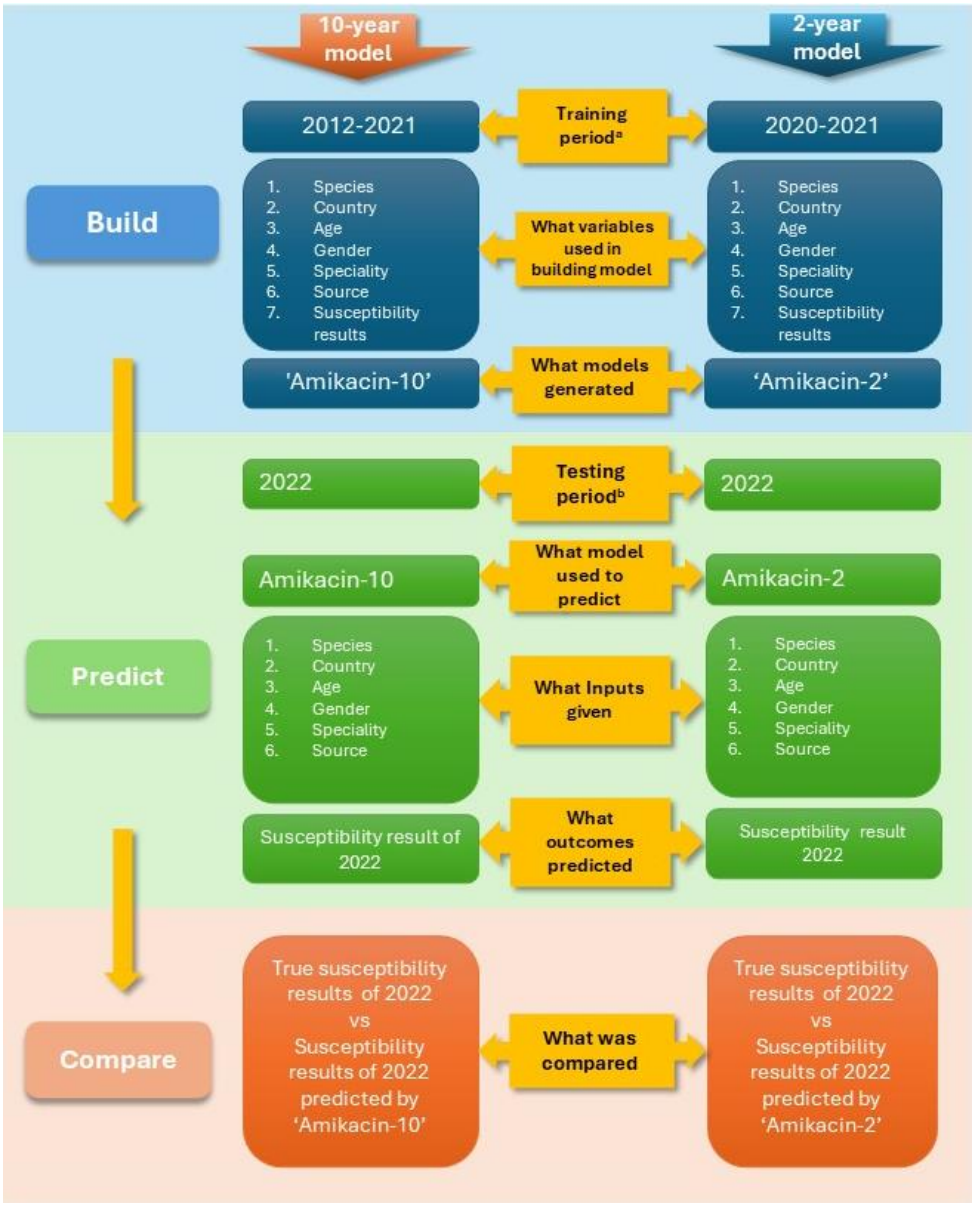

Figure 1: Outline of workflow followed to build and compare prediction models

a. The duration during which the model is trained on a set of data, where it learns patterns and relationships
b. The timeframe for evaluating a model on a separate dataset to measure its performance after training

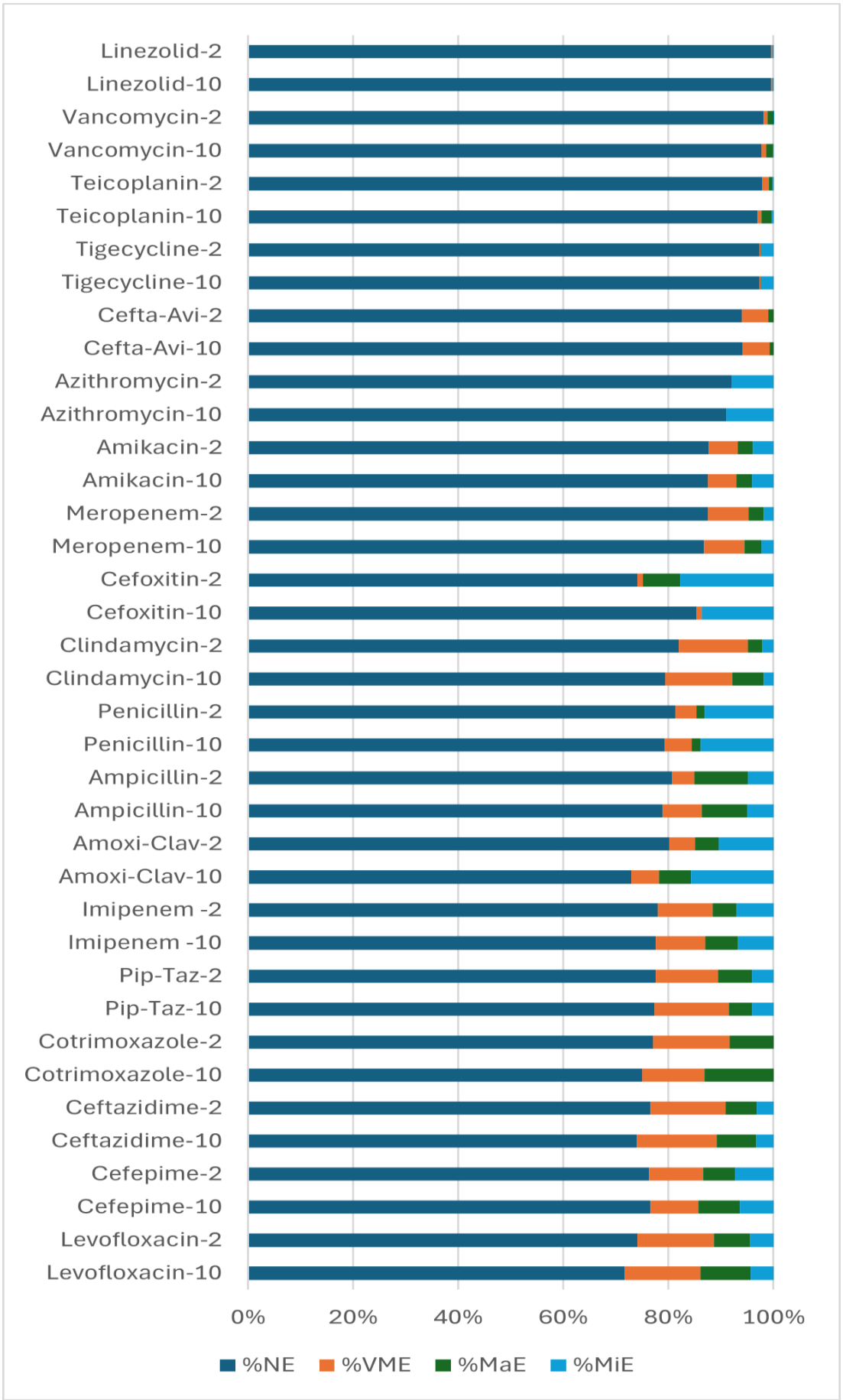

Figure 2: 10-year and 2-year models for drugs showing proportion of concordance and errors from comparison with true results of 2022 (VME: Very major error, MaE: Major error, MiE: Minor error)

Drug	Total isolates	NE	%NE	TE	%TE	VME	%VME	MaE	%MaE	MiE	%MiE
Linezolid-2	18139	18045	99%	94	1%	54	0%	0	0%	40	0%
Linezolid-10	18139	18045	99%	94	1%	54	0%	0	0%	40	0%
Vancomycin-2	18139	17781	98%	358	2%	150	1%	201	1%	7	0%
Vancomycin-10	18139	17730	98%	409	2%	149	1%	235	1%	25	0%
Teicoplanin-2	15099	14776	98%	323	2%	192	1%	86	1%	45	0%
Teicoplanin-10	15099	14629	97%	470	3%	126	1%	283	2%	61	0%
Tigecycline-2	41263	40165	97%	1098	3%	73	0%	64	0%	961	2%
Tigecycline-10	41263	40174	97%	1089	3%	80	0%	70	0%	939	2%
Cefta-Avi-2	30473	28625	94%	1848	6%	1526	5%	322	1%	0	0%
Cefta-Avi-10	30473	28657	94%	1816	6%	1573	5%	243	1%	0	0%
Azithromycin-2	524	482	92%	42	8%	0	0%	0	0%	42	8%
Azithromycin-10	524	477	91%	47	9%	0	0%	0	0%	47	9%
Amikacin-2	35724	31294	88%	4430	12%	1972	6%	1029	3%	1429	4%
Amikacin-10	35724	31240	87%	4484	13%	1946	5%	1065	3%	1473	4%
Meropenem-2	40175	35152	87%	5023	13%	3136	8%	1105	3%	782	2%
Meropenem-10	40175	34845	87%	5330	13%	3101	8%	1317	3%	912	2%
Cefoxitin-2	197	146	74%	51	26%	2	1%	14	7%	35	18%
Cefoxitin-10	197	168	85%	29	15%	2	1%	0	0%	27	14%
Clindamycin-2	15419	12634	82%	2785	18%	2039	13%	407	3%	339	2%
Clindamycin-10	15419	12248	79%	3171	21%	1962	13%	907	6%	302	2%
Penicillin-2	3721	3024	81%	697	19%	149	4%	57	2%	491	13%
Penicillin-10	3721	2950	79%	771	21%	193	5%	58	2%	520	14%
Ampicillin-2	25847	20842	81%	5005	19%	1106	4%	2650	10%	1249	5%
Ampicillin-10	25847	20371	79%	5476	21%	1963	8%	2227	9%	1286	5%
Amoxi-Clav-2	22247	17826	80%	4421	20%	1108	5%	1007	5%	2306	10%
Amoxi-Clav-10	22247	16211	73%	6036	27%	1188	5%	1363	6%	3485	16%
Imipenem -2	35723	27853	78%	7870	22%	3751	11%	1599	4%	2520	7%
Imipenem -10	35723	27713	78%	8010	22%	3352	9%	2236	6%	2422	7%
Pip-Taz-2	37076	28756	78%	8320	22%	4410	12%	2384	6%	1526	4%
Pip-Taz-10	37076	28646	77%	8430	23%	5287	14%	1643	4%	1500	4%
Cotrimoxazole-2	38486	29661	77%	8825	23%	5585	15%	3240	8%	0	0%
Cotrimoxazole-10	38486	28846	75%	9640	25%	4579	12%	5061	13%	0	0%
Ceftazidime-2	36255	27734	76%	8521	24%	5230	14%	2108	6%	1183	3%
Ceftazidime-10	36255	26793	74%	9462	26%	5545	15%	2723	8%	1194	3%
Cefepime-2	35723	27261	76%	8462	24%	3690	10%	2149	6%	2623	7%
Cefepime-10	35723	27261	76%	8462	24%	3279	9%	2786	8%	2309	6%
Levofloxacin-2	54408	40345	74%	14063	26%	7915	15%	3681	7%	2467	5%
Levofloxacin-10	54408	38999	72%	15409	28%	7815	14%	5231	10%	2363	4%

Table 1: 10-year and 2-year models for drugs showing Total errors (TE), No error (NE) from comparison with true results of 2022, along with % (VME: Very major error, MaE: Major error, MiE: Minor error)

References

1. Howard A, Aston S, Gerada A, et al. Antimicrobial learning systems: an implementation blueprint for artificial intelligence to tackle antimicrobial resistance. *Lancet Digit Health*. 2024;6(1):e79-e86. doi:10.1016/S2589-7500(23)00221-2

2. Pinto-de-Sá, R.; Sousa-Pinto, B.; Costa-de-Oliveira, S. Brave New World of Artificial Intelligence: Its Use in Antimicrobial Stewardship—A Systematic Review. Antibiotics 2024, 13,307. https://doi.org/10.3390/antibiotics13040307

3. https://app.akkio.com

4. Humphries RM, Ambler J, Mitchell SL, et al. CLSI Methods Development and Standardization Working Group Best Practices for Evaluation of Antimicrobial Susceptibility Tests [published correction appears in J Clin Microbiol. 2023 Oct 24;61(10):e0073923. doi: 10.1128/jcm.00739-23]. *J Clin Microbiol*. 2018;56(4):e01934-17. Published 2018 Mar 26. doi:10.1128/JCM.01934-17