An Integrative Machine Learning approach to discover determinants of AMR for Highly Important Pathogens.

Raghav Awasthi[1], Samprati Agarwal [2], Vaidehi Rakholia [2], Lovedeep Singh Dhingra [2], Tavpritesh Sethi [1,2*],

1 Indraprastha Institute of Information Technology Delhi, New Delhi, India
2 All India Institute of Medical Sciences, Department of Pediatrics, New Delhi, India

Abstract

Background:

Antimicrobial resistance (AMR) is a complex multifactorial outcome of health, socio-economic and geopolitical factors. Therefore, tailored solutions for mitigation strategies could be more effective in dealing with this challenge. Knowledge-synthesis and actionable models learned upon large datasets is critical in order to diffuse the risk of entering into a post-antimicrobial era.

Objective:

This work is focused on learning determinants of AMR for critically important pathogens from large datasets.

Methods:

We integrated large data sets in order to find the unbiased determinants of AMR . We chose a Bayesian Decision Network (BDN) approach within the causal modeling framework because these allow intuitive visualization, exploration and incorporation of expert-insights and sanity checks into the modeling steps. A BDN is a Bayesian Network (BN) embellished with utility and decision nodes that are set by the decision-maker (clinician/policymaker) to assign the value.Finally Integrating Bayesian networks with classical machine learning approaches lead to effective modeling of the level of AMR.

Results:

From MAR (multiple Antibiotic Resistance) scores We found developing countries at high risk of AMR compare to developed countries for all the critically important pathogens. Also, Principal components analysis (PCA) revealed that Governance, finance, and disease burden variables have a strong association with AMR. We further quantified the impact of determinants in a probabilistic way and observed that Heath system access and government effectiveness is a strong factor to reduce AMR. which is confirmed by what-if analysis. And finally, our supervised machine learning models have shown decent performance with the highest on **Staphylococcus aureus**. for the Staphylococcus aureus, our model predicted Ceftrolirne and Oxacillin with highest AUROC 0.94 and 0.89 respectively.

Introduction

Antimicrobial resistance is the reduction in the efficacy of antimicrobials in treating infections due to pathogen evolution under strong selection pressures given the repeated and heavy consumption of antibiotics. Even though we have witnessed a significant improvement in global health, there are still millions out there who do not have adequate access to health services [1]. Factors such as rising incomes, incessant infectious diseases, poor or marginalized populations having limited access to primary health care and consumption of antibiotics without any prescription have been compounding the problem of antimicrobial resistance in the low-income and middle-income countries [2]. Ensuring proper sanitation, better governance, increased focus on public health care and regulation of the private health care sector are a must to combat the spread of antimicrobial resistance [3]. While its disastrous effect on health outcomes can be understood by an estimated 10 million deaths, the cost of treating resistant infections is estimated to reach the US $100 trillion by the year 2050 [4,5,6]. To prevent these catastrophic consequences, it is essential to develop evidence-based policies for AMR mitigation. The role of government (policymakers) is multifold in improving antibiotic stewardship policies, increasing antibiotic use surveillance and financing these along with funding development of new drugs [7]. There arises a need to formulate governance frameworks to help the policymakers in designing, monitoring national action plans for tackling antimicrobial resistance at all levels: local, regional, national and global [8]. It is thus essential to understand/ quantify the roles of these factors in the development of resistance. Tackling antibiotic resistance is a high priority for WHO. A global action plan on antimicrobial resistance, including antibiotic resistance, was endorsed at the World Health Assembly in May 2015. The global action plan aims to ensure the prevention and treatment of infectious diseases with safe and effective medicines. WHO has been leading multiple initiatives to address antimicrobial resistance like World Antimicrobial Awareness Week, The Global Antimicrobial Resistance Surveillance System (GLASS), Global Antibiotic Research and Development Partnership (GARDP), Interagency Coordination Group on Antimicrobial Resistance (IACG) with the motivation to improve awareness and understanding of antimicrobial resistance, to strengthen surveillance and research, to reduce the incidence of infection, to optimize the use of antimicrobial medicines and to ensure sustainable investment in countering antimicrobial resistance. A political declaration endorsed by Heads of State at the United Nations General Assembly in New York in September 2016 signalled the world's commitment to taking a broad, coordinated approach to address the root causes of antimicrobial resistance across multiple sectors, especially human health, animal health and agriculture. WHO is supporting the Member States to develop national action plans on antimicrobial resistance, based on the global action plan. Primary health care can play a vital role in tackling antimicrobial resistance. Community engagement and empowerment is essential to prevent common health problems without the unnecessary use of antimicrobials. Multisectoral action on antimicrobial resistance to limit the usage of antibiotics in the agricultural sector and ensuring equitable and good quality primary health care to all can act as an effective response to the antimicrobial resistance [9]. In this work, we quantified the relationships between Governance, Global AMR and microbiomes. We created a compendium of Bayesian AI models to predict drug-combination patterns that are most likely to mitigate AMR in given machine learning models for resource, drug-combinations, Priority pathogen list and Critically Important Antibiotics. And finally we used supervised machine learning models and predicted the susceptibility of antibiotics.

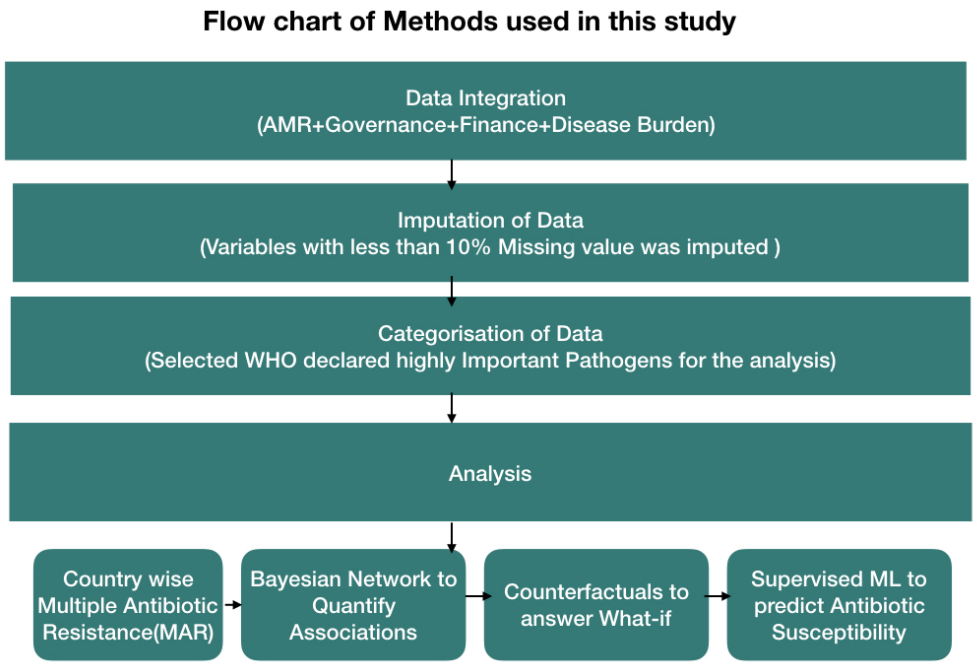


Figure 1:Flowchart showing pipeline used in our analysis.

Methods

In this study, we developed a novel pipeline **Fig1** to understand the robust connection among Antibiotic Resistance and Disease burden, Governance, Finance and other Socioeconomic indicators.and using these indicators prediction of susceptibility of antibiotics.

Data Integration

The AMR base data was extracted from the Dream Challenges which have 633820 no. of isolates. The characteristics for the same have been displayed in **Table 1**. We have integrated the AMR base data with the WGI (World Governance Indicators) data [10], the GBD (Global Burden of Disease Study) data [11] and the Finance data [12].

Age.Group				
0 to 2 Years	452 (1.26%) (5.98%)	15329 (42.83%) (5.51%)	20012 (55.91%) (5.75%)	35793 (100.00%) (5.65%)
13 to 18 Years	99 (0.73%) (1.31%)	6348 (47.06%) (2.28%)	7043 (52.21%) (2.02%)	13490 (100.00%) (2.13%)
19 to 64 Years	2174 (0.74%) (28.77%)	130162 (44.04%) (46.80%)	163201 (55.22%) (46.88%)	295537 (100.00%) (46.63%)
3 to 12 Years	242 (0.90%) (3.20%)	12613 (46.88%) (4.53%)	14052 (52.22%) (4.04%)	26907 (100.00%) (4.25%)
65 to 84 Years	1560 (0.75%) (20.65%)	86561 (41.63%) (31.12%)	119801 (57.62%) (34.41%)	207922 (100.00%) (32.80%)
85 and Over	344 (0.80%) (4.55%)	23406 (54.29%) (8.42%)	19364 (44.91%) (5.56%)	43114 (100.00%) (6.80%)
Unknown	2685 (24.28%) (35.53%)	3709 (33.54%) (1.33%)	4663 (42.17%) (1.34%)	11057 (100.00%) (1.74%)
Phenotype				
Not Given	4246 (1.37%) (56.19%)	133118 (42.91%) (47.86%)	172852 (55.72%) (49.65%)	310216 (100.00%) (48.94%)
(BL Neg)	233 (1.11%) (3.08%)	8696 (41.35%) (3.13%)	12100 (57.54%) (3.48%)	21029 (100.00%) (3.32%)
(BL Pos)	57 (0.95%) (0.75%)	2500 (41.47%) (0.90%)	3471 (57.58%) (1.00%)	6028 (100.00%) (0.95%)
ESBL	406 (1.00%) (5.37%)	17829 (43.93%) (6.41%)	22352 (55.07%) (6.42%)	40587 (100.00%) (6.40%)
MRSA	516 (0.92%) (6.83%)	22555 (40.29%) (8.11%)	32910 (58.79%) (9.45%)	55981 (100.00%) (8.83%)
MSSA	646 (1.12%) (8.55%)	24202 (41.94%) (8.70%)	32864 (56.94%) (9.44%)	57712 (100.00%) (9.11%)
non ESBL	1452 (1.02%) (19.22%)	69228 (48.66%) (24.89%)	71587 (50.32%) (20.56%)	142267 (100.00%) (22.45%)

Table 1. Characteristic of AMR Data

AMR WGI :

WGI data covers six dimensions of governance for over 200 countries over the period 1996-2018. The six dimensions of governance are as follows:

1) Control of Corruption
2) Voice and Accountability
3) Political Stability and Absence of Violence/Terrorism
4) Government Effectiveness
5) Regulatory Quality
6) Rule of Law

We picked only variables Rank and Estimate from each sheet corresponding to the above-mentioned dimensions of governance. Here, Rank refers to the "Percentile rank among all countries (ranges from 0 (lowest) to 100 (highest) rank)" while Estimate refers to the "Estimate of governance (ranges from approximately -2.5 (weak) to 2.5 (strong) governance performance)". These were then left merged with the AMR data set by country and year, thus giving us the resultant AMR WGI data set.

AMR WGI GBD :

The GBD data set estimates the burden of diseases for over 195 countries. It consists of a total of 334 covariates data files for the period 1980-2017. In the Gender variables, some covariates have information about gender specificity and some have information common for both. So in the latter case, we regenerated the same information for both genders and then combined the data in rows. Similarly, all covariates have three types of Age-Group:

1) All Ages
2) In the form of intervals

Priority category of Pathogens	Pathogens	Samples in High Income Country	Samples in Middle Income Country
Critical	Pseudomonas aeruginosa	11373	13266
Critical	Klebsiella pneumoniae	11901	15437
Critical	Escherichia coli	14864	18125
High	Staphylococcus aureus	21456	27314
Critical	Acinetobacter baumannii	4468	5816
High	Enterococcus faecium	2385	1980
Critical	Enterobacter cloacae	7462	6860

Table 2. Sample Wise distribution of Who Declared Critically Important Pathogens in High and Middle Income Country

3) Age-standardized in covariates where age groups are in the form of intervals

We then grouped by age, gender, country and summarised by taking the mean. After having made these changes in the covariates files, these were then left merged with the AMR data by country, age and gender, taking only the value of covariates, thus giving us the resultant AMR WGI GBD data set.

AMR WGI GBD FINANCE :

We reshaped the finance data such as columns country, year and then merged with the above mentioned AMR_WGI_GBD data set by country and year. We discarded all those variables which had more than 10% missing data, thus giving us the resultant AMR WGI GBD Finance data set.

Data Imputation and Discretisation

We replaced the missingness in the AMR phenotypic data by 'Not_Given' and antibiotic_I by "Not_Tested" and the missingness in the variables of merged(amr_wgi_gbd_finance) datasets using a state-of-the-art Random Forest approach [13] that took into account the associations between variables while inputting the data. After the imputation of data, we divided every numerical variable into three disjoint intervals i.e. Low, Medium and High using a KNN based algorithm for the purpose of discretisation.

WHO Declared Critical Important Species Selection:

Further, we filtered data as per the WHO declared Critically Important and highly Important Pathogens [14] in **Table 2**. The pathogens we took into consideration included Pseudomonas aeruginosa, Klebsiella pneumoniae, Escherichia coli, Acinetobacter baumannii, Enterobacter cloacae, Staphylococcus aureu and Enterococcus faecium. We considered their samples pertaining to high-income and middle-income countries.

Analysis

MAR Score :

The MAR (Multiple Antibiotic Resistance) [15] index of a single isolate is defined as a/b, where, a is the number of antibiotics which the isolate is resistant to and b is the number of antibiotics tested.

Bayesian Network Analysis:

Bayesian Network that learn the latent structure in complex data, represent it as compact graphical representations and allow inferential and intervention modeling .

Interpretability and explainability are key challenges in AI-based decision models. Recent years have seen a revival of causal networks that can be learned directly and reliably from data as a quintessential approach towards achieving explainability and trust for complex problems faced by society. In this study, Bayesian network was learned directly from complex multivariate data for explainable intervention modelling. Firstly, we took all the variables of WGI WDI FINANCE AMR and learn one-time Bayesian network [17]. Then the Markov blankets of all antibiotics were calculated from the one-time network and then the final network was learned considering only the Markov blanket variables. The inferences were then calculated from this network.

(Counterfactual)What-if Analysis:

We imagined the hypothetical situation of the variables (derived from the Markov Blanket of Bayesian Network) and performed counterfactual analysis using the R package Counterfactual [18] to answer the hypothetical 'what if' questions. For example, what would be the Resistance of Antibiotics if countries with poor health system access prevailed the characteristics of countries which have advanced health system access.

Prediction of antibiotic susceptibility using supervised Machine Learning

Identification of isolates susceptible to certain antibiotics is essential in fighting against antibiotic-resistant pathogens. so we extarcted $GBD, WGI, Finance$ indicators (which was in the markov blanket of antibiotics in the Pathogen wise Bayesian network), demographic and clinical information of isolates as a predictor of antibiotic susceptibility . Data were partitioned into training (80%) and testing (20%) sets and the class imbalance was corrected using the Synthetic Minority Oversampling Technique(SMOTE) []. Different supervised machine learning models - Random-Forest(RF), Support vector machine (SVM), logistic, naive-Bayes, were learned for predicting the response to mental health indicators using the Scikit-learn library []in Python.

Results

To identify the Actionable Global determinants of AMR, we first recognised the pattern of AMR spread across the globe. We took into consideration the WHO declared critically important pathogens and performed network analysis to identify pathogen wise AMR Mitigator. For each identified mitigator, we determined the impact in Middle and High-income countries separately. This was followed by a counterfactual analysis to measure the hypothetical inferences. and fianlly using identified GBD,WGI,and finance determeninats we predicted the susceptibilty of antibiotics in the selected pathogens.

Global Prevalence of Multiple Antibiotic Resistance of Critically Important Pathogens:

To understand the spread pattern of Antibiotic resistance of different pathogens, we calculated the Multiple Antimicrobial Resistance (MAR) country-wise **Fig 2**. This revealed that every species had a different pattern of resistance spread. For Enterococcus, MAR was found in the range [0.24,0.57] with the highest value in Vietnam (0.57) and the lowest value in Venezuela (0.24). For Acinetobacter baumannii, MAR value was found in the range (0,0.86] with the highest value in Vietnam (0.86). For E coli, MAR value was found in the range [0.032, 0.45] with highest values in

Indonesia (0.45), India (0.40) and the lowest value in Norway (0.032). For the klebsiella pneumoniae, MAR was found in the range [0.09,0.64] with the highest value in Serbia (0.64) and the lowest value in Japan (0.09). For Pseudomonas aeruginosa, MAR value was observed in the range [0.016,0.75] with the highest value in El Salvador (0.75) and the lowest value in Norway (0.016).

Component Level Bayesian Network Revealed Strong Connection of Antibiotics with Independent WGI , Finance , GBD data sets .

In order to check the directional relationship among the independent datasets i.e. WGI, GBD, Finance and AMR, we performed Principal component analysis (PCA) in the Numerical data set and Multiple correspondence analysis(MCA) in Categorical data sets. We picked 2 components which captured higher variance of datasets and with the help of these 2 components, on a one time network **Figure 3** we found a strong association of governance, finance and GBD principal components with Antibiotics components.

Probabilistic Impact Quantification of Governance, Finance sociodemographic, Socioeconomic and Burden of Disease on Antibiotic Resistance

For WHO declared critical Species we learned Robust Bootstrapped Network. From the Network Inference governance of the country and Health System Access, we found a strong Influencer of Antibiotic resistance in all species. We found the probability of Cefepime resistance in E coli species to be 21% higher in a country where Voice and Accountability was low as compared to a country where Voice and Accountability was High. The probability of piperacillin resistance in Acinetobacter baumannii species is 50% lower in a country where there is High level of government effectiveness. The probability of cefepime resistance is 12.98% lower in the species Pseudomonas aeruginosa and 13% lower in the species staphylococcus aureus in a country where there is High level of government effectiveness as compared to a country having a low level of government effectiveness. Similarly, the probability of Amikacin resistance is 12.17% lower in the species Pseudomonas aeruginosa and 11.6% lower in the species staphylococcus aureus in a country where there is High level of government effectiveness as compared to a country having a low level of government effectiveness. We also found that good Health system Access of a country also significantly decreases Ceftraixone and meropenem resistance in Species. The probability of ceftriaxone resistance in countries with good health system access was observed to be 14.74% lower in E Coli, 36% lower in Acinetobacter baumannii, 12.47% lower in Enterobacter Cloacae, 15% lower in Klebsiella pneumoniae while the probability of Meropenem resistance in Klebsiella pneumoniae was observed to be 3.43% higher in countries with good health system access. Improvements in water sanitation and hygiene (WASH) are critical elements of preventing infections and reducing the spread of antimicrobial resistance (AMR) as identified in the Global Action Plan to combat AMR. In our analysis, we found out that the probability of Cefepime resistance in Klebsiella pneumonia is 19.87% higher in countries where the level of Unsafe_Wash_Sanitation is high. Our analysis quantified the effect of fruit consumption on Cefepime resistance in Klebsiella pneumoniae. High Food consumption decreased 13.04% chance of Cefepime resistance.

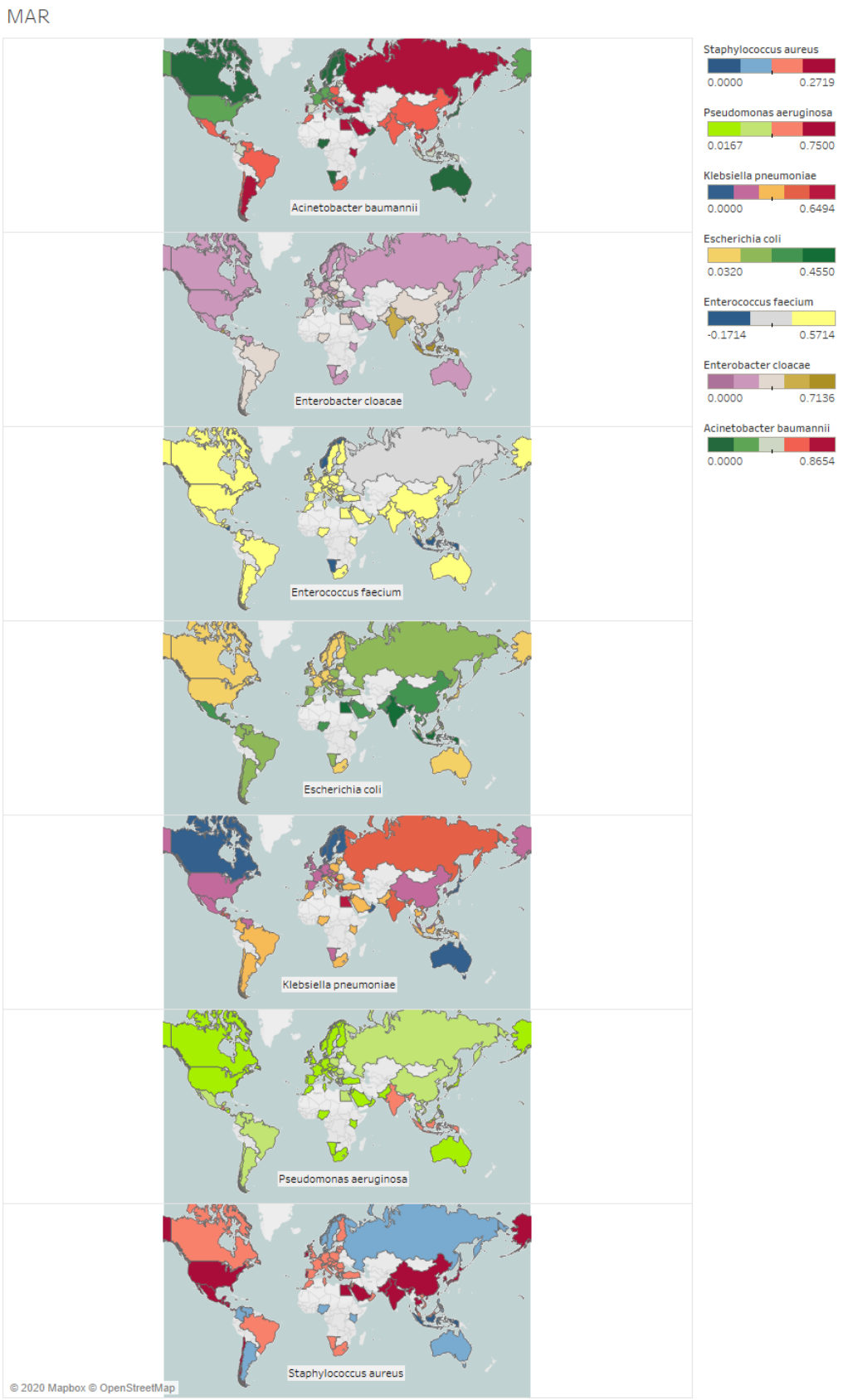


Figure 2: Global Prevalence of Antibiotic Resistance in different Pathogens .

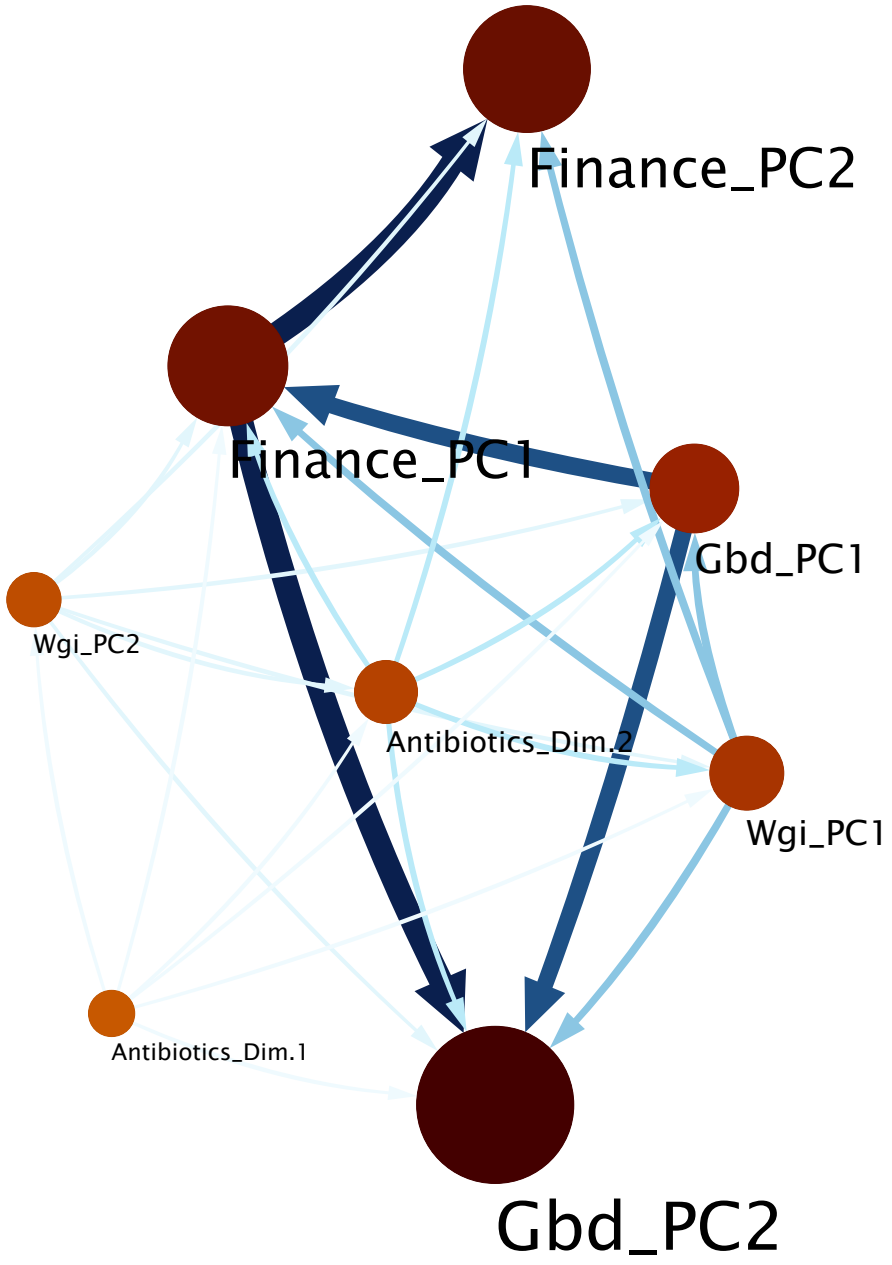


Figure 3: Bayesian Network of Principal Components reveals strong connection of AMR with WGI,Finance,GBD

Species	Variables (Parent)	Antibiotics (Child)	Pr(Antibiotics=Resistant | $Variable = Low$)	Pr(Antibiotics=Resistant | $Variable = High$)
E Coli	Mean_Temp_Long_Term	Cefepime	0.15	0.32
	Voice and Accountability	Cefepime	0.33	0.12
	Fruits_G_Adj	Cefepime	0.2	0.2
	Health_System_Access_Capped	Cefriaxone	0.3302	0.1828
	Health_System_Access_Capped	Ceftazidime	0.1357	0.189
	Dengue_Outbreak	Ceftazidime	0.0978	0.1652
	Health_System_Access_Capped	Meropenem	0.0139	0.006
Acinetobacter baumannii	Mean_Temp_Long_Term	Piperacillin Tazobactam	0.5	0.7489
	Government_Effectiveness	Piperacillin Tazobactam	0.79	0.29
	Hospital_Bed_per1000	Meropenem	0.66	0.47
	Health_System_Access_Capped	Meropenem	0.38	0.67
	Health_System_Access_Capped	Ceftriaxone	0.7	0.34
EnteroBacter Cloacae	Rota Coverage Prop	Meropenem	0.0238	0.0319
	Health System Access Capped	Meropenem	0.05	0.0262
	Sev Scaler Diarrhea	Cefepime	0.09	0.2218
	Health System Access Capped	Cefrtiaxone	0.4107	0.286
	Health System Access Capped	Minocycline	0.1255	0.0508
Klebsiella pneumonia	Health System Access Capped	Cefriaxone	0.39	0.24
	GFDD_EI_04	Cefriaxone	0.2856	0.1
	Fruits_G_ADJ	Cefepime	0.4111	0.2807
	Mean_Temp_Long_term	Cefepime	0.2865	0.3339
	Unsafe_Wash_Sanitation	Cefepime	0.2134	0.4121
	Health System Access Capped	Ceftazidime	0.1984	0.3516
	Dengue_Outbreak	Ceftazidime	0.2186	0.25
	pop_Dens_under_150_Psqkm_Pct	Ceftazidime	0.226	0.3005
	Fruits_G_ADJ	Piperacillin Tazobactam	0.229	0.2195
	Health System Access Capped	Piperacillin Tazobactam	0.1996	0.2046
	Health System Access Capped	Amoxicillin Clavulanate	0.2889	0.2697
	Rule of Law	Amoxicillin Clavulanate	0.3396	0.119
	pop_Dens_under_150_Psqkm_Pct	Amoxicillin Clavulanate	0.2205	0.3291
	Health System Access Capped	Meropenem	0.0587	0.093
Pseudomonas aeruginosa	Health System Access Capped	Meropenem	0.1826	0.2413
	Hospital Beds per 1000	Meropenem	0.2865	0.2381
	Health System Access Capped	Cefepime	0.2151	0.1353
	Government Effectiveness	Cefepime	0.2	0.0702
	Government Effectiveness	Amikacin	0.1417	0.02
Staphylococcus aureus	Government Effectiveness	Cefepime	0.2	0.07
	Government Effectiveness	Amikacin	0.1417	0.0257
	Health System Access Capped	Cefepime	0.2151	0.1353
	Health System Access Capped	Meropenem	0.1826	0.2413
	Hospital Beds Per 1000	Meropenem	0.2865	0.2381
Enterococcus faecium	Pop_Dens_150_300_Psqkm_Pct	Vancomycin	0.135	0.352
	Prop_Urban	Vancomycin	0.13	0.29
	Health System Access Capped	Penicillin	0.7107	0.4042
	GFDD_OI_01	Penicillin	0.3882	0.6233
	Health System Access Capped	Minocycline	0.1868	0.1979

Table 3. Inferences from Pathogen wise Bayesian Network

Counterfactual Analysis quantified the reduction in Ceftriaxone Resistance if countries with poor Health System Access follow advanced health system Access.

we performed a counterfactual analysis on the ceftriaxone resistance for different species in the Middle and High-income countries. For the species Enterobacter Cloacae, counterfactual effect (what would be the Ceftriaxone Resistance if poor Health system follows the characteristics of advanced health system Access) of Health system Access on Ceftriaxone resistance was observed to be consistent in the Middle and High-income countries with median quantile effect in Middle-income countries to be 0.007 with 95% CI [-0.05,0.0.7] and in High-income countries to be 0.02 with 95% CI [0.003,0.04]. For Acinetobacter Baumannii and E Coli, we found the counterfactual effect of Health system Access on Ceftriaxone in the Middle-income countries to be much higher than that in the High-income countries. The median quantile effect in the Middle-income countries was observed to be -0.9 with 95% CI [-0.17,-0.02] in the case of Acinetobacter Baumannii and -0.03 with 95%CI [-0.16,0.10] in the case of E Coli. The median quantile effect in the High-income countries was observed to be 0.05 with 95% CI [-0.03,013] in the case of Acinetobacter Baumannii and -0.004 with 95% CI [-0.03,0.02] in the case of E Coli.

Prediction of Antibiotic Susceptibility:

Overall we have found that random forest model are best performing models. Among all the pathogens Interestingly we have observed that our prediction models have best preformed for the **Staphylococcus aureus** [Fig5] which is known as the most dangerous of all of the many common staphylococcal bacteria and often cause skin infections. for the Staphylococcus aureus our model predicted Ceftrolirne and Oxacillin with heighest AUROC 0.94 and 0.89 respectively(from RF model).Also we have found

our model performed well for the cefepime and ceftazidime susceptibility prediction for *Klebsiella pneumoniae* with AUCROC 0.88 ,0.92 respectively(From RF model). From our prediction models we have also found that Meropenem Sucepetibilty is highly predictble for the *Escherichia coli* (AUROC=0.93) compared to other pathogens *Pseudomonas aeruginosa*(AUROC=0.75) , *klebsiella pneumoniae* (AUROC=0.81),*Acinetobacter baumannii* (AUROC=0.75). futher we analysed the feature importance of prediction model(RF) for the Meropenem susceptibilty for we found that *scherichia coli* and intrestingly we found *Health system access* of a country a most importnat feature [Supp Fig 2]. Altough our model has not showed high performance for pathogens like Pseudomonas aeruginosa.but overall our models have showed decent and actionable performance performance.

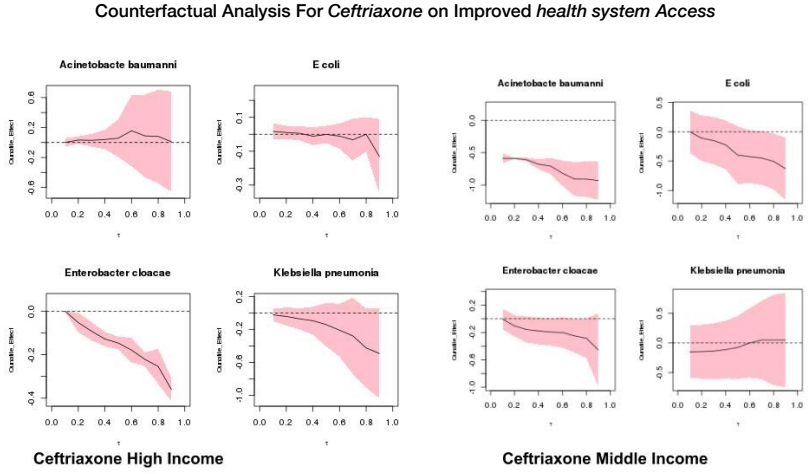


Figure 4: Counterfactual Analysis answered the question what could be resistance of Ceftriaxone205and if countries with poor health system access having High Health system access. Analysis was performed sepratly for the middle and high income countries.

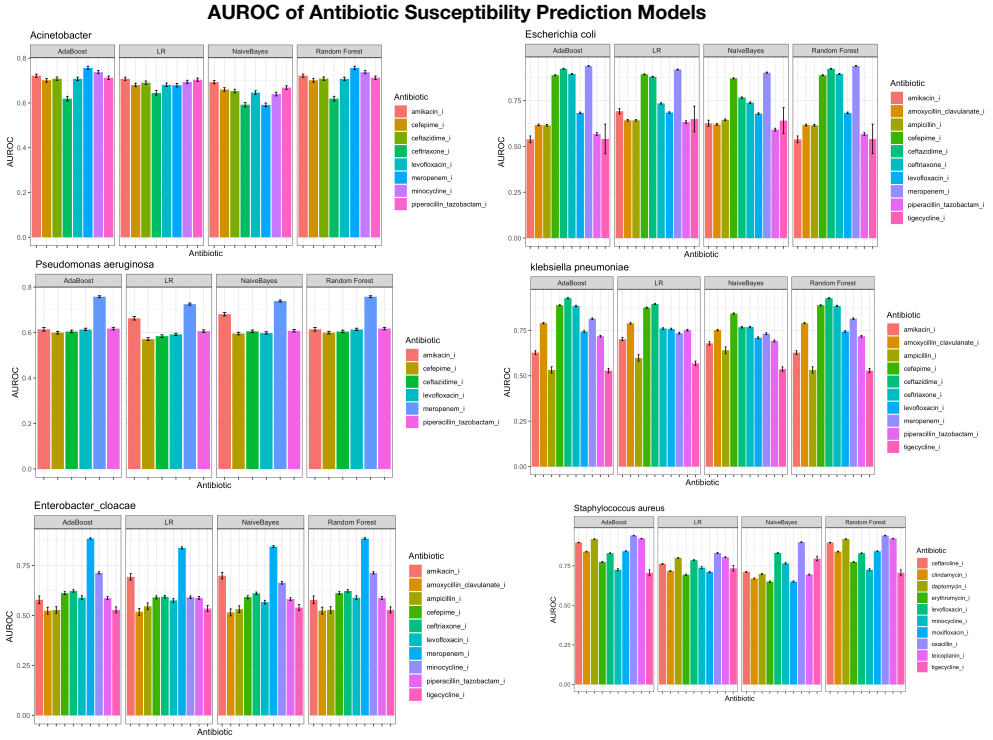


Figure 5: *AUROC* of Prediction models.Error bars in the figure denote the 95% CI AUROC

Discussion

Antibiotic resistance is rising to alarmingly high levels in all parts of the world. New resistance mechanisms are emerging and spreading globally, threatening our ability to treat common infectious diseases. A growing list of infections – such as pneumonia, tuberculosis, blood poisoning, gonorrhoea and foodborne diseases – are becoming harder and sometimes impossible to treat as antibiotics become less effective. The emergence and spread of the resistance are worse in those regions where the antimicrobial medicines can be bought for human or animal use without a prescription. Similarly, in countries without standard treatment guidelines, antibiotics are often over-prescribed by health workers and veterinarians and overused and misused by the public. Certain economic and governance factors such as Political Stability and Absence of Violence/Terrorism, Control of Corruption, Voice and Accountability, Government Effectiveness, Regulatory Quality and Rule of Law play a key role in designing such frameworks which can help curtail the spread of Anti-microbial resistance. By consolidating four different datasets (the AMR, GBD, WGI and Finance data sets) into one, this study aims to find the actionable global determinants. Taking into consideration the prevalence of MAR score across the globe, our findings report that MAR score is non-uniform across the globe for critically important pathogens. This finding calls upon devising policies that are effective at each country level. Using AI techniques like Bayesian Network Analysis, we found a strong connection of AMR with the WGI, Finance and GBD datasets which motivated us to proceed with Counterfactual analysis of Ceftriaxone. As per the results, the counterfactual effect of Health system Access on Ceftriaxone in the Middle-income countries turned out to be much higher than that in the High-income countries. With the on-going advancements in medicine, we have breakthrough treatment techniques from complex organ transplants to robotic surgeries. All of these have been possible by keeping bacterial infections under control. But the rising instances of Antibiotic resistance, which are accelerated by the misuse and overuse of antibiotics, may render simple bacterial diseases untreatable. The time calls upon strategies at every stratum of the society to reduce the impact and limit the spread of the resistance. Global governance systems and finance regulators, industrial stakeholders, medical experts and scientists need to stand united to tackle this problem first hand. The Global Reference List includes priority indicators pertaining to four domains namely health status, service coverage, health systems and risk factors which countries can use to monitor their health priorities at national and sub-national levels [19]. Providing timely access to the healthcare system, judicious and lawful uses of medicinal resources, awareness regarding public hygiene with strict action against unlawful practices, such as drug distribution without prescription, should be taken at all costs [20]. There are a few limitations of our work. Some confounders are still missing such as government effectiveness and health-system influenced AMR. Thus, a complete pathway has not been explained due to the dearth of data. We had data only for 72 countries which included only Middle-income and High-income countries. But Governance finance-related Intervention might also be required in case of Low-income countries. Our work is focused on high-level policy-making, so it does not include biological aspects such as Microbiome role for AMR.

Acknowledgments

This work was partially supported by the Wellcome Trust/DBT India Alliance Fellowship IA/CPHE/14/1/501504 awarded to Tavpritesh Sethi and the Center for Artificial Intelligence at IIIT-Delhi. . Authors also acknowledge Prof. Rakesh Lodha from All India Institute of Medical Sciences, New Delhi, for his valuable inputs. we also

acknowledge Aditya Nagori , Mehrab Singh Gill, Arshita Bhatt. for their support in figure preparation , data collection and helping in manuscript writing.

References

1. Tracking universal health coverage: 2017 global monitoring report. Geneva: World Health Organization and International Bank for Reconstruction and Development/The World Bank; 2017. (http://pubdocs.worldbank.org/en/193371513169798347/2017-global-monitoring-report.pdf, accessed 17 September 2018)

2. Singh PK. A universal good: How increased health coverage can help beat back antimicrobial resistance. World Health Organization, Regional Office for South-East Asia, 2017 (http://www.searo.who.int/mediacentre/features/2017/how-increased-health-coverage-can-help-beat-back-amr-resistance/en/, accessed 17 September 2018)

3. Collignon P, Beggs JJ, Walsh TR, Gandra S, Laxminarayan R. Anthropological and socioeconomic factors contributing to global antimicrobial resistance: a univariate and multivariable analysis. Lancet Planet Health. 2018 Sep;2(9):e398-e405. doi: 10.1016/S2542-5196(18)30186-4. PMID: 30177008.

4. Krumperman, PAUL H. Multiple antibiotic resistance indexing of Escherichia coli to identify high-risk sources of fecal contamination of foods. Appl. Environ. Microbiol. 46.1 (1983): 165-170.

5. Naylor, Nichola R., et al. Estimating the burden of antimicrobial resistance: a systematic literature review. Antimicrobial Resistance Infection Control 7.1 (2018): 58

6. Smith, R. D., and Joanna Coast. The economic burden of antimicrobial resistance: why it is more serious than current studies suggest(2012).

7. Maragakis, Lisa L., Eli N. Perencevich, and Sara E. Cosgrove. Clinical and economic burden of antimicrobial resistance. Expert review of anti-infective therapy 6.5 (2008): 751-763.

8. Anderson, M., Schulze, K., Cassini, A., Plachouras, D., Mossialos, E. (2019) A governance framework for development and assessment of national action plans on antimicrobial resistance. The Lancet Infectious Diseases, 19(11), e371-e384.

9. World Health Organization. (2018). Antimicrobial resistance and primary health care. World Health Organization. https://apps.who.int/iris/handle/10665/326454. License: CC BY-NC-SA 3.0 IGO

10. Worldwide Governance Indicators (www.govindicators.org), The World Bank

11. Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2017 (GBD 2017) Covariates 1980-2017. Seattle, United States: Institute for Health Metrics and Evaluation (IHME), 2018.

12. Global Financial Development Database (GFDD), The World Bank.

13. D.J. Stekhoven and P. Buhlmann. MissForest—non-parametric missing value imputation for mixed-type data. Bioinformatics 2012; 28(1):112–118

14. World Health Organization. WHO list of critically important antimicrobials for human medicine (WHO CIA list). 5th revision 2016. World Health Organization.

15. Randall LP, Woodward MJ. The multiple antibiotic resistance (mar) locus and its significance. Res Vet Sci 2002; 72: 87–93

16. J. A. G ´amez, J. L. Mateo, and J. M. Puert. Learning Bayesian networks by hill climbing: efficient methods based on progressive restriction of the neighborhood Data Mining and Knowledge Discovery, vol. 22, no. 1,pp. 106–148, Jan. 2011. [Online]. Available: https://doi.org/10.1007/s10618-010-0178-6

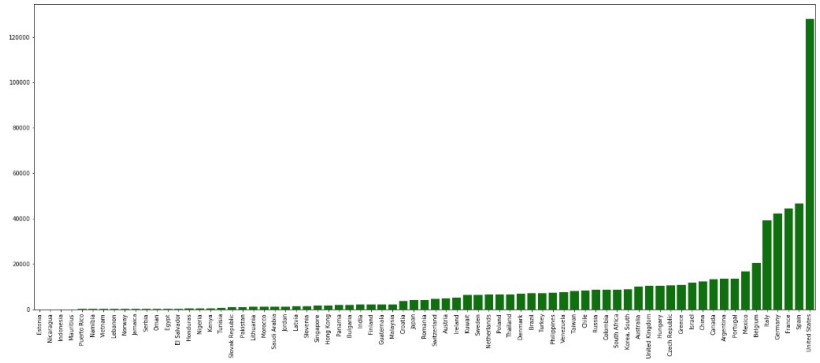
17. Friedman N, Koller D. Being Bayesian about Network Structure. arXiv:13013856 [cs, stat]. 2013 [cited 19 May 2020]. Available: http://arxiv.org/abs/1301.3856

18. T. Sethi and S. Maheshwari. wiseR: A Shiny Application for End-to-End Bayesian Decision Network Analysis and Web-Deployment Nov. 2018. [Online]. Available: https://CRAN.R-project.org/package=wiseR10

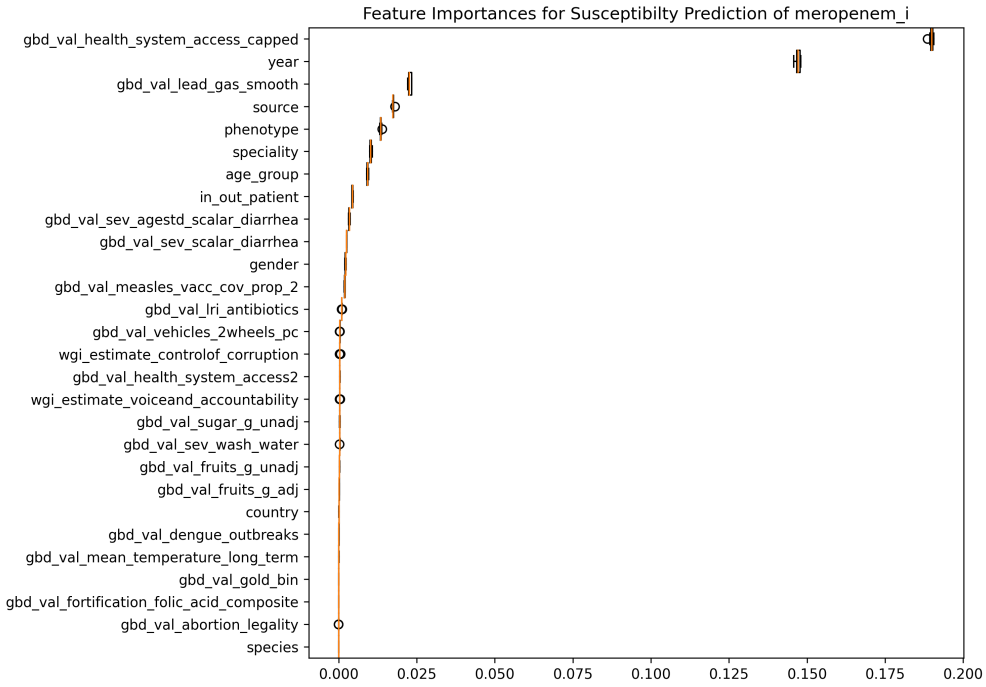
19. Mingli Chen, Victor Chernozhukov, Ivan Fernandez-Val and Blaise Melly, 2017. Counterfactual analysis in R: a vignette. CeMMAP working papers CWP64/17, Centre for Microdata Methods and Practice, Institute for Fiscal Studies

20. World Health Organization. Global Reference List of 100 Core Health Indicators (plus health-related SDGs). 2018.

21. Laxminarayan, R., Duse, A., Wattal, C., Zaidi, A. K., Wertheim, H. F., Sumpradit, N., ... Greko, C. (2013). Antibiotic resistance—the need for global solutions. The Lancet infectious diseases, 13(12), 1057-1098.



Supp Fig 1:Country wise samples in AMR data



Supp Fig 2: Feature Importance for the Prediction of Meropenem Susceptibilty agains E.Coli obtained from best performing the Random Forest