

# AMR Data Collection, Curation, Analysis and Antibigram Development

Julhas Sujan  
Data Science Engineer

August 05, 2023

# By end of this presentation, you will be able to learn:



## How is measured AMR rates?

Estimate AMR prevalence of pathogens



## AMR data analysis using different tools

Tools, Trends, patterns, antibiograms



## Importance and sources of AMR data

Identification of data source and variables



## Antibiograms, Dashboard and Research

Develop antibiogram and publish manuscript



## AMR information cycle, Data collection, curation

Basic cleaning



## AI, Machine learning and Data science

Drug discovery, Prediction and Mobile application

# Why is AMR data important

*“On the basis of our predictive statistical models, there were an estimated 4.95 million (3.62–6.57) deaths associated with bacterial AMR in 2019, including 1.27 million deaths attributable to bacterial AMR.”*

**THE LANCET** Submit Article Log in Register Subscribe Claim 🔍 ☰

ARTICLES | VOLUME 399, ISSUE 10325, P629-655, FEBRUARY 12, 2022

## Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis

Antimicrobial Resistance Collaborators <sup>†</sup> • [Show footnotes](#)

**Open Access** • Published: January 19, 2022 • DOI: [https://doi.org/10.1016/S0140-6736\(21\)02724-0](https://doi.org/10.1016/S0140-6736(21)02724-0) • [Check for updates](#)

PDF [3 MB] Figures Save Reprints Request

Study type	Number of study-location-years	Sample size	Sample size events	1: region	2: infectious syndrome	3: case-fatality ratio	4: pathogen distribution	5: antibiotic use	6: prevalence of resistance	7: drug profile
Multiple cause of death	2980	120,871,372	Deaths							
Hospital discharge	391	192,533,435	Discharges							
Microbial or laboratory data with outcome	1102	3,060,802	Isolates							
Microbial or laboratory data without outcome	2302	145,067,113	Isolates							
Literature studies	607	701,356	Cases, isolates, or pathogen-drug compatibility tests							
Single drug resistance profiles	158	86								
Pharmacoepidemiology	86									

PlumX Metrics

- Summary
- Introduction
- Methods
- Results

**Summary**

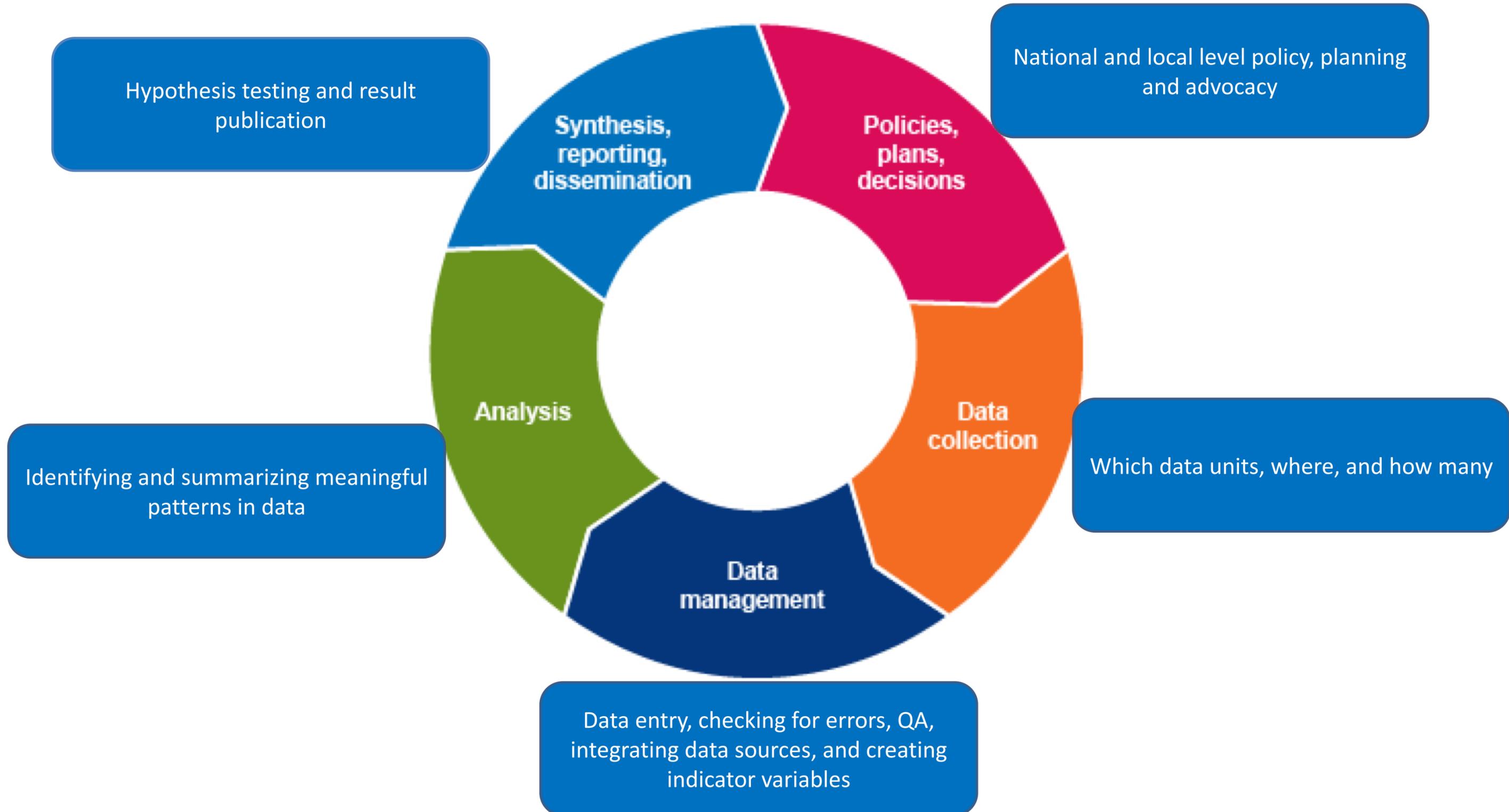
### Background

Antimicrobial resistance (AMR) poses a major threat to human health around the world. Previous publications have estimated the effect of AMR on incidence, deaths, hospital length of stay, and health-care

Request Your Institutional Access to the *Lancet* journal

ADVERTISEMENT

# AMR information cycle



# Sources of data and necessary information (variables)

## Laboratory based surveillance data variables

- ▶ Patient ID
- ▶ Sex
- ▶ Age
- ▶ Unit
- ▶ Ward
- ▶ Department
- ▶ Sample type
- ▶ Sample collection date
- ▶ Date of admission
- ▶ Organism
- ▶ AST results as RIS or Zol

Minimum data variables

Manual  
Registers

Laboratory  
Information  
System

VITEK 2 or  
MYLA  
Software

Excel Sheet/  
Others

**Primary:** surveys, interviews, routine record keeping, laboratory tests or experiments.  
**Secondary:** Existing data are accessed and analysed by someone other than the person(s) who collected the data.

# Example of a Sample data

File Home Insert Page Layout Formulas Data Review View Help Tell me what you want to do														
C27		fx 70												
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Identification number	Sex	Age	Specimen_date	Year	Sample	Organism	AMC	AMK	CIP	CRO	CTX	GEN	SXT
2	_9242362412_	m	70	1/2/2017	2017	Pus	Staphylococcus aureus		S	S			S	R
3	_9139642321_	f	34	1/2/2017	2017	Pus	Staphylococcus aureus		S	R			S	S
4	_8021891052_	m	52	1/2/2017	2017	Pus	Staphylococcus aureus		S	R			S	R
5	_1617418339_	m	27	1/2/2017	2017	Pus	Staphylococcus aureus		S	S			S	S
6	_1501885197_	m	48	1/2/2017	2017	Wound Sv	Staphylococcus aureus	R	S	R			S	R
7	_1301121252_	m	60	1/2/2017	2017	Pus	Staphylococcus aureus		S	R			S	R
8	_6584123381_	f	70	1/2/2017	2017	Urine	Klebsiella sp.		S	S	S		S	R
9	_5245247211_	m	60	1/2/2017	2017	Urine	Klebsiella sp.		S	R	S		S	R
10	_5140148802_	f	12	1/2/2017	2017	Wound Sv	Klebsiella sp.		R	R	R	R	R	R
11	_5120871702_	f	63	1/2/2017	2017	Urine	Klebsiella sp.		R	R	R		R	R
12	_4620205931_	m	63	1/2/2017	2017	Pus	Klebsiella sp.	R	S	R	R	R	S	R
13	_2442203663_	f	47	1/2/2017	2017	Pus	Klebsiella sp.	R	S	S	S	S	S	S
14	_1240941431_	m	65	1/2/2017	2017	Urine	Klebsiella sp.		S	R	S		S	R
15	_6328111119_	f	70	1/2/2017	2017	Blood	Escherichia coli		S	S	S		S	S
16	_4718848462_	f	57	1/2/2017	2017	Urine	Escherichia coli		S	R	S		S	R
17	_3955138154_	f	35	1/2/2017	2017	Urine	Escherichia coli		S	R	S		S	S
18	_3673267750_	m	65	1/2/2017	2017	Urine	Escherichia coli		S	R	R		S	S
19	_2526268228_	m	58	1/2/2017	2017	Pus	Escherichia coli		R	S	S	S	R	R
20	_1641141162_	f	27	1/2/2017	2017	Urine	Escherichia coli		S	R	S		S	S
21	_1531292172_	f	40	1/2/2017	2017	Urine	Escherichia coli		R	R	R		R	S
22	_1516984671_	f	70	1/2/2017	2017	Urine	Escherichia coli		S	S	S		S	S
23	_0186641342_	f	43	1/2/2017	2017	Urine	Escherichia coli		S	R	R		R	S
24	_6783982251_	f	40	1/3/2017	2017	Pus	Staphylococcus aureus	S	S	S			S	R
25	_6393247351_	f	55	1/3/2017	2017	Wound Sv	Staphylococcus aureus	R	S	R			R	R
26	_3108178241_	m	70	1/3/2017	2017	Pus	Staphylococcus aureus		S	R			S	R

# Sources of data and necessary variables ...

## General Information

Case ID, Case date, Hospital registration number of the patient, Treatment received from, Department, IPD patient ward number

## Previous admission detail

Status, Hospital information, Date of admission, Date of discharge, Number of days previous hospitalisation

## Patient profile

Name, Age, Gender, Father/ Husband's name, Mother's name, Occupation, Educational status, Mailing address, Mobile number of the respondent, Name and mobile number of family member

## Comorbidity

Diabetes, Chronic Kidney Disease, Chronic Liver Disease, CKD, Cancer, Other

## Clinical information

Urinary Tract Infection, Diarrhoea, Lower Respiratory Tract Infection, Wound Infection, Septicemia

## Specimen

Urine, Stool, Wound swab/ pus, Sputum, Blood, Endotracheal aspirate and other

## Antibiotic treatment history

Treated with antibiotic status, Type of antibiotics prescribed, Antibiotic name, Number of days taken antibiotic, Number of doses of antibiotic taken per day, No of missed doses

## Organism and AST results

- Identified organisms
- Zone of inhibition value or
  - RIS result
  - Both

# How is measured AMR data?

**Study objective:** To estimate the country/hospital level AMR prevalence and trends for WHO priority pathogens and other clinically important and frequently isolated pathogens as well as to enable the comparison of countries on spatiotemporal maps<sup>1</sup>.

$$\text{AMR rate} = \frac{\text{No. of non-susceptible isolates}}{\text{No. of tested isolates}} \times 100 \text{ ( CI 95\% )}$$

The AMR rate is the extent to which a pathogen is resistant to a particular antimicrobial agent or class as is determined by the proportion of isolates that are non-susceptible (i.e., either intermediate or resistant) over a one-year period

\*\* CIs for AST data can be constructed using the Wilson score method

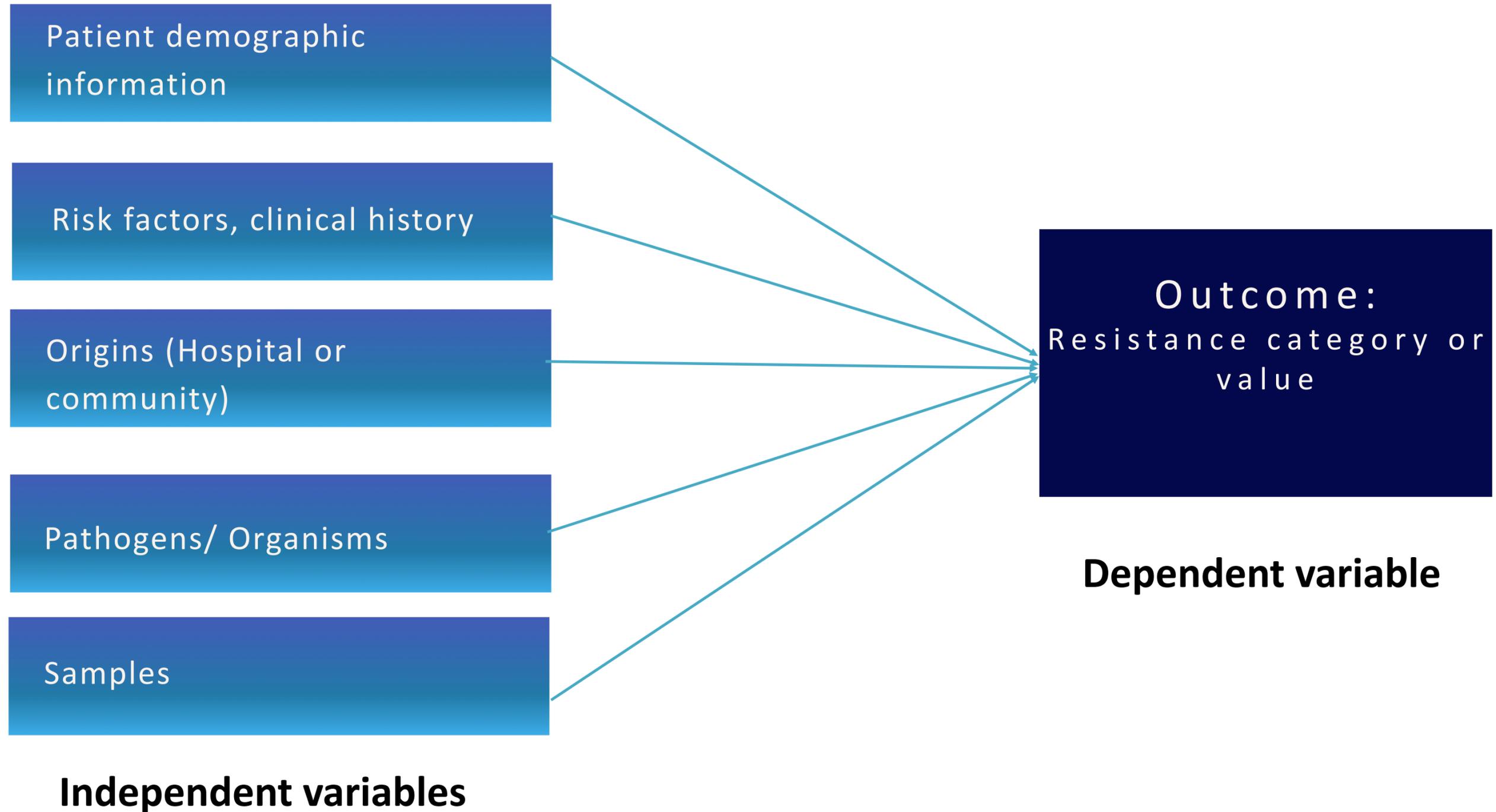
<sup>1</sup>[https://aslm.org/wp-content/uploads/2023/07/AMR\\_REPORT\\_BURKINA\\_FASO.pdf?x89467](https://aslm.org/wp-content/uploads/2023/07/AMR_REPORT_BURKINA_FASO.pdf?x89467)

# Drivers of Antimicrobial Resistance

To determine the association between AMR and its potential drivers, the following patient and country-level factors were considered:

- **Patient-level factors:** demographics (age and gender), diagnosis, comorbidities, antimicrobial usage, presence of device (catheter, central line or ventilator) and origin of infection (hospital or community)
- **Country-level factors:** Global Health Security index scores on AMR prevention, primary education, GDP per capita, physician and nurse density, disease prevalence and antibiotic consumption in defined daily dose (DDD) per 1 000 inhabitants (the country-level associations are presented separately at a regional or continental level)

# AMR data analysis (Conceptual framework)



# AMR data analysis (Descriptive stats)

- **Descriptive analysis for categorical variables**
  - Count: AST results showed that 38 of the 120 MRSA isolates were resistant to levofloxacin
  - Proportion: AST results showed that 31.7% of the 120 MRSA isolates were resistant to levofloxacin
  - Frequency: Found 120 positive isolates
- **Descriptive analysis for numeric variables**
  - Arithmetic mean
  - Geometric mean
  - Median
  - Mode
- **Measures of central tendency**
  - Percentiles
  - Interquartile range
  - Variance
  - Standard variance

*For manuscript: A large number of MRSA isolates showed resistance to levofloxacin (83.9%), ciprofloxacin (82%), erythromycin (77.7%) and clindamycin (72.3).*

# AMR data analysis (Inferential stats)

- **Allow us to make predictions about a population, based on data collected from a sample**
- **Discuss in detail in the next session**

# AMR data curation and analysis tools

## R-Programming

- R Package for AMR:  
<https://msberends.github.io/AMR>
- Example:  
<https://julhas.com/blog/amr-data-analysis-using-r>



Research | Conferences | Presentations | DHIS2 Profile | DHIS2 Trainingland | DHIS2 ePortal | JS Edutech | MPH Materials | QAAPT | Hiking | Travelling | Running | Blog

◀ Back to blog

### AMR Data Analysis Using R

Posted: 23 December 2021  
Author: Julhas Sujan

#### About AMR and R package

R package to simplify the analysis and prediction of Antimicrobial Resistance (AMR) and to work with microbial and antimicrobial data and properties, by using evidence-based methods. Copyright by: <https://msberends.github.io/AMR/index.html#copyright>

#### Outline

• R packages installation

## Stata/ SPSS

- Available commands
- Example:  
<https://julhas.com/blog/amr-data-analysis-using-stata>



Research | Conferences | Presentations | DHIS2 Profile | DHIS2 Trainingland | DHIS2 ePortal | JS Edutech | MPH Materials | QAAPT | Hiking | Travelling | Running | Blog

◀ Back to blog

### AMR Data Analysis Using Stata

Posted: 22 December 2021  
Author: Julhas Sujan

#### About AMR and Stata

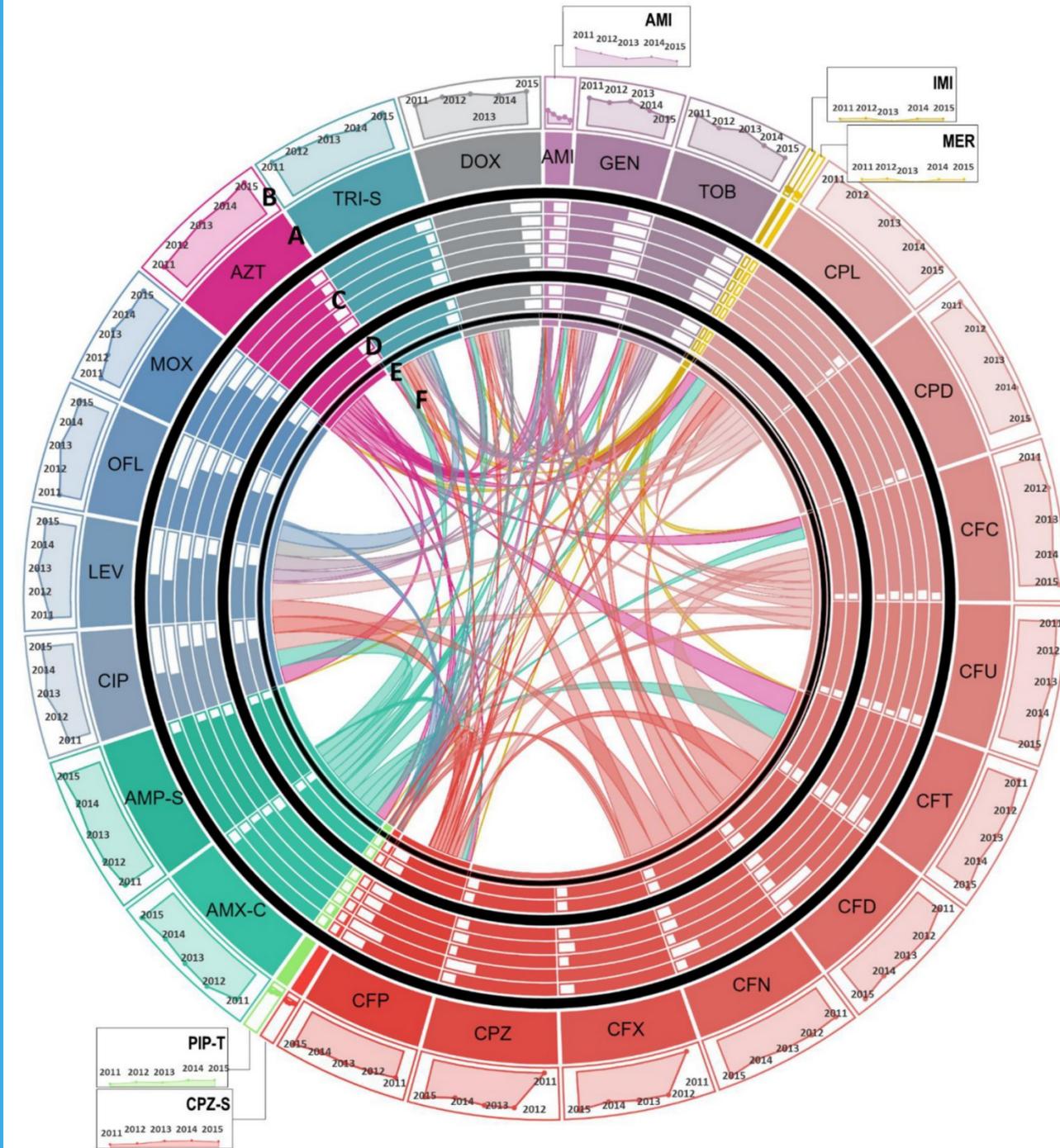
The IBM Stata software will help us to analyze the antimicrobial resistance data. We can use cross tabulation, pearson's chi-squared test, bar, pie, line, box, histogram, and regression analysis to determine frequency distributions.

#### Outline

## Python

- Python commands

- A:** Each section of the diagram represents the resistance observed in *E. coli* against the antibiotic. Size of each section is proportional to the proportion of *E. coli* resistant to the antibiotic over the study period. Antibiotics of the same class are shown in similar colors.
- B:** Line graphs show temporal trends of proportion of resistant *E. coli* in a clockwise direction from 2011 to 2015.
- C:** Bar charts show the comparison of susceptibility to resistant strains in patients of different age groups. Moving from out to inward, bars represent proportion of resistant *E. coli* reported in children <5 years of age, young adults between 6 to 18 years, middle aged 19 to 45 years old, 45 to 65 years old patients, and elderly over 65 years of age, respectively.
- D:** Gender-wise comparison to susceptibility to resistant *E. coli* is shown in form bars. Outer circle and inner circle shows proportion of resistant *E. coli* isolated from women vs. men, respectively.
- E:** For co-resistance analysis, antibiotics belonging to the same class with same susceptibility profile for all isolates of *E. coli* were merged into a single variable.
- F:** Proportion of *E. coli* isolates resistant to one antimicrobial resistant to another antimicrobial are shown in the connections. The area covered by the connection on E is proportional to the level of co-resistance observed. Co-resistance proportions were scaled down to 1/15<sup>th</sup> of the actual overlap for visualization.



**Fig: Antimicrobial resistance in *Escherichia coli* (*E. coli*).**

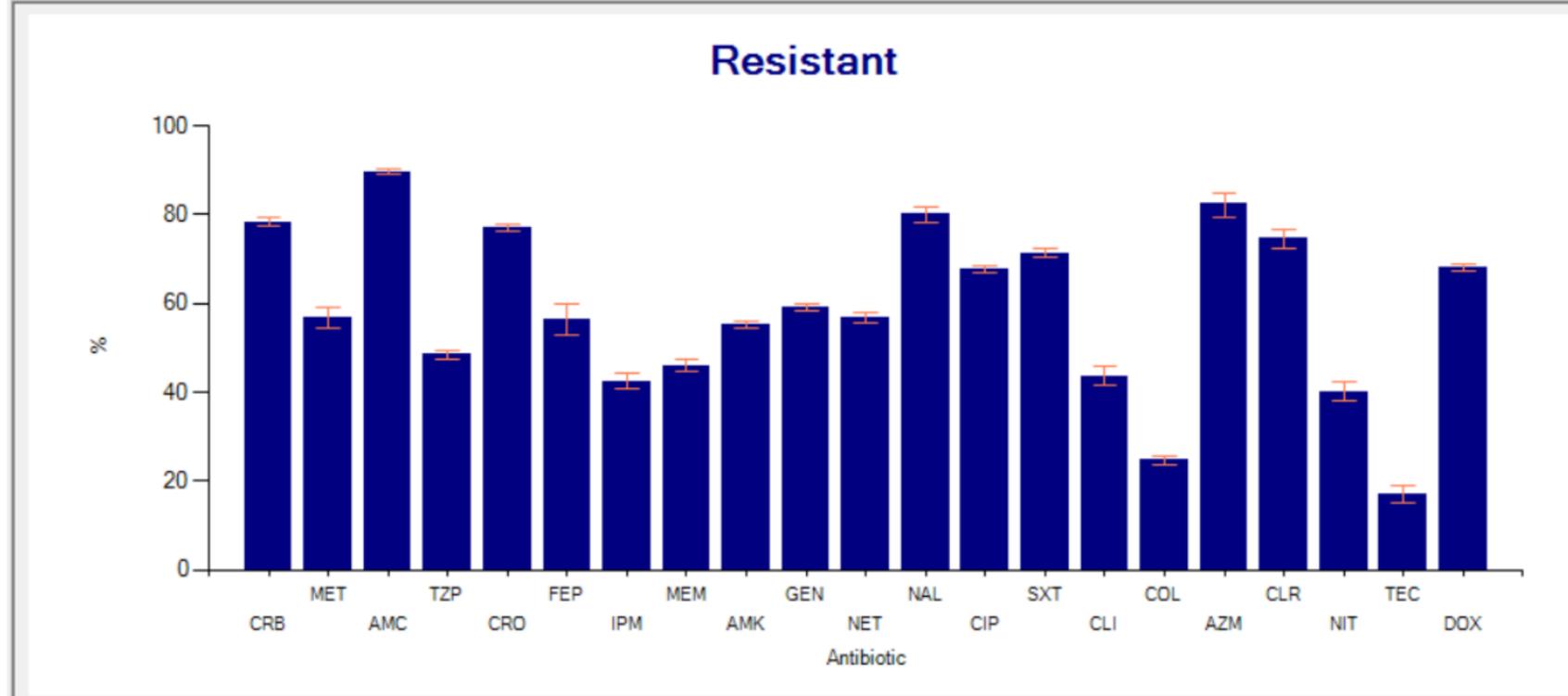
# WHONET

Analysis results - Organism = All organisms (n=42016 Isolates)

File Edit

Copy table Copy graph Save table Save graph Continue  Show hidden columns

Organisms	Code	Antibiotic name	Site of infection	Breakpoints	Number	%R	%I	%S	%R 95%C.I.	%S 95%C.I.	Number	6	7	8	9	10	11	12	13	14	15	16	17
All organisms	CRB_ND100	Carbenicillin		(None)	8146	78.3	0.7	21	77.4-79.2	20.1-21.9	0												
All organisms	MET_ND5	Methicillin		10 - 13	1700	56.7	0.4	42.9	54.3-59.1	40.5-45.3	0												
All organisms	AMC_ND20	Amoxicillin/Clavulanic acid		14 - 17	10240	89.6	0.2	10.2	89.0-90.2	9.7-10.8	0												
All organisms	TZP_ND100	Piperacillin/Tazobactam		18 - 20	9393	48.5	2.1	49.4	47.5-49.5	48.4-50.4	1												
All organisms	CRO_ND30	Ceftriaxone		20 - 22	12916	77	0.5	22.5	76.2-77.7	21.8-23.2	0												
All organisms	FEP_ND30	Cefepime		19 - 24	749	56.5	0.5	43	52.8-60.0	39.4-46.6	0												
All organisms	IPM_ND10	Imipenem		20 - 22	3117	42.4	2.2	55.3	40.7-44.2	53.6-57.1	1												
All organisms	MEM_ND10	Meropenem		20 - 22	4997	46	1.1	53	44.6-47.4	51.6-54.3	0												
All organisms	AMK_ND30	Amikacin		15 - 16	13947	55.3	0.6	44.1	54.5-56.1	43.2-44.9	0												
All organisms	GEN_ND10	Gentamicin		13 - 14	13531	59	0.3	40.6	58.2-59.9	39.8-41.5	0												



RIS

- Resistant
- Intermediate
- Susceptible
- Unknown
- Number tested

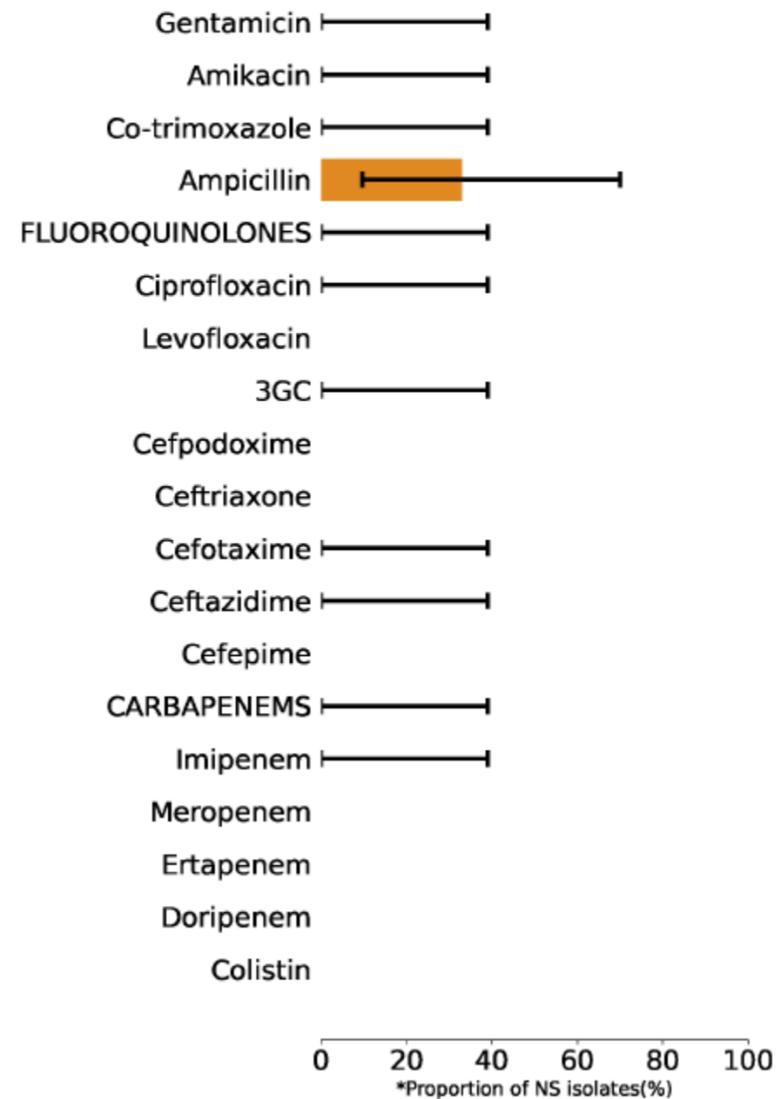
Test measurements

- Piperacillin/Tazobactam
- Imipenem

## Section [2]: Isolate-based surveillance report

Blood: *Escherichia coli*

( No. of patients = 6 )



Antibiotic agent	Proportion of NS isolates (n)	95% CI
Gentamicin	0% (0/6)	0% - 39%
Amikacin	0% (0/6)	0% - 39%
Co-trimoxazole	0% (0/6)	0% - 39%
Ampicillin	33% (2/6)	10% - 70%
FLUOROQUINOLONES	0% (0/6)	0% - 39%
Ciprofloxacin	0% (0/6)	0% - 39%
Levofloxacin	NA	-
3GC	0% (0/6)	0% - 39%
Cefpodoxime	NA	-
Ceftriaxone	NA	-
Cefotaxime	0% (0/6)	0% - 39%
Ceftazidime	0% (0/6)	0% - 39%
Cefepime	NA	-
CARBAPENEMS	0% (0/6)	0% - 39%
Imipenem	0% (0/6)	0% - 39%
Meropenem	NA	-
Ertapenem	NA	-
Doripenem	NA	-
Colistin	NA	-



Dashboard

Custom Reports

Dataset Import

Settings

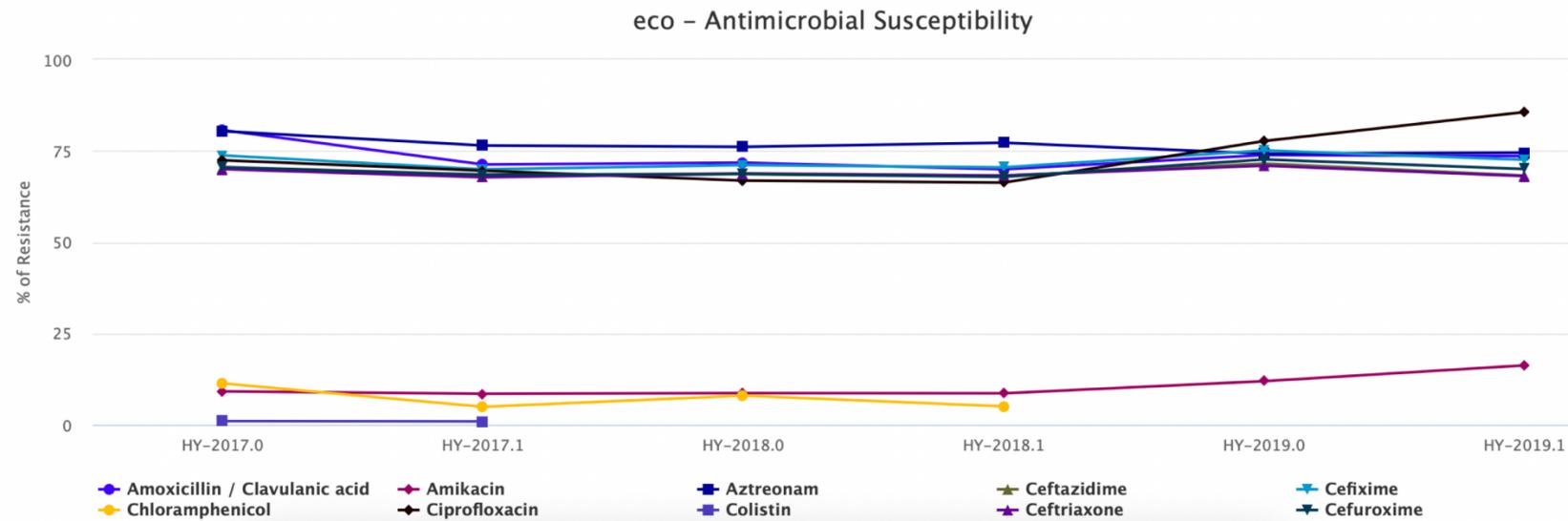
Admin Reports

Category

Need any support? Click here

### Antibiotic susceptibility trends analysis

QAAPT\_Demo\_Dataset Escherichia coli Amoxicillin / Clavulanic ac 5



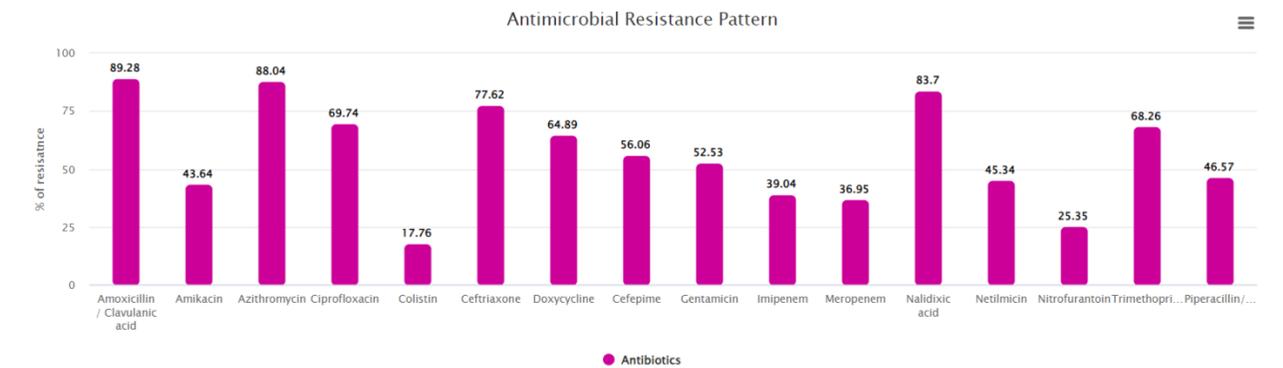
### Antibiotic resistant pattern analysis

Escherichia coli Amoxicillin / Clavulanic ac 5

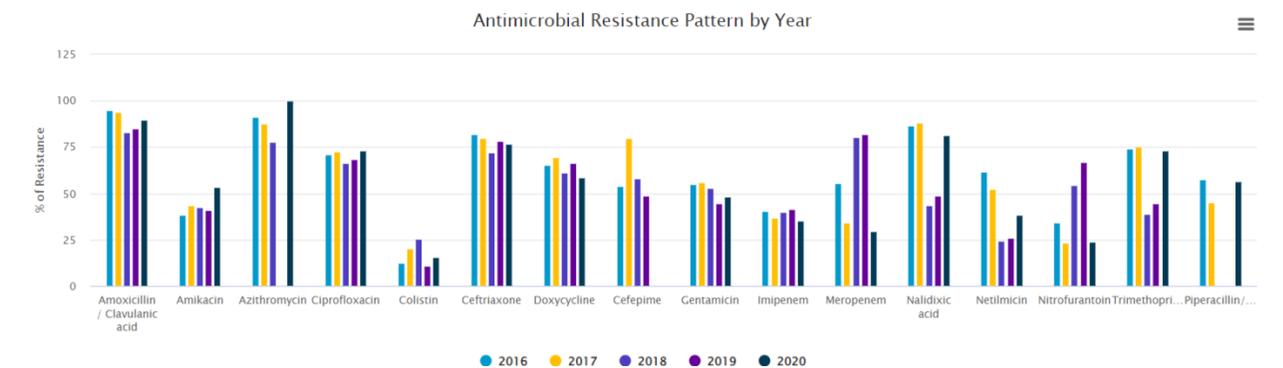
Antibiotic name	Percentage (%)	Number of Resistant Isolates	Total test	Combined Value
Amikacin	43.64	1424	3263	43.64 (1424/3263)
Amoxicillin / Clavulanic acid	89.28	1933	2165	89.28 (1933/2165)
Azithromycin	88.04	162	184	88.04 (162/184)
Cefepime	56.06	37	66	56.06 (37/66)
Ceftriaxone	77.62	2487	3204	77.62 (2487/3204)
Ciprofloxacin	69.74	1987	2849	69.74 (1987/2849)
Colistin	17.76	322	1813	17.76 (322/1813)
Doxycycline	64.89	2114	3258	64.89 (2114/3258)
Gentamicin	52.53	1653	3147	52.53 (1653/3147)

Showing 1 to 16 of 16 entries

### Antimicrobial resistance pattern - cumulative



### Antimicrobial resistance patterns by year



# Other platforms

- ATLAS: Antimicrobial Testing Leadership And Surveillance
- <http://www.bccdc.ca/health-professionals/data-reports/antimicrobial-resistance-utilization/antimicrobial-resistance-dashboard>
- <https://amrmap.net/>
- <https://dashboard.globalamrhub.org/>
- <https://amrhub.ru/>
- <https://amrcloud.net/en/>
- <https://amr.tghn.org>
- <https://www3.paho.org/data/index.php/en/mnu-topics/antimicrobial-resistance/572-amr-acerca-en.html>
- NARMS: <https://www.fda.gov/animal-veterinary/antimicrobial-resistance/national-antimicrobial-resistance-monitoring-system>

# Antibiogram components

- Table 1: Distribution of bacterial growth in different specimen
- Table 2: Pattern of organisms isolated
- Table 3: Distribution of bacteria isolated from OPD, IPD
- Table 4: Distribution of sex and specimens
- Table 5: Antibiotic susceptibility pattern of major Gram-positive organisms
- Table 6: Antibiotic susceptibility pattern of major Gram-negative organisms
- Table 7: Distribution of organisms and specimens

# Example of cumulative antibiograms

Organism	No. of isolates	Less restricted antibiotics																		Restricted Antibiotics									
		Ampicillin		Amoxicillin / clavulanate		Cefazolin		Tri methoprim		Sulfamethoxazole & tri methoprim		Nitrofurantoin		Gentamicin		Ticarcillin / clavulanate		Piperacillin / tazobactam		Amikacin	Ceftriaxone	Meropenem	Norfloxacin	Vancomycin					
		%S	n	%S	n	%S	n	%S	n	%S	n	%S	n	%S	n	%S	n												
<i>Escherichia coli</i>	821	64	821	84	821	87	816	77	821	79	801	96	821	97	821	76	802	92	53	100	803	98	804	100	804	94	806	R	
<i>Klebsiella pneumoniae</i>	133	R		89	133	85	131	84	133	88	130			95	133	88	130	98	118	99	130	88	130	100	130	85	130	R	
<i>Pseudomonas aeruginosa</i>	105	R		R		R		R		R				92	105	60	104	95	97	95	100	R		91	105	90	100	R	
<i>Enterococcus faecalis</i>	80	100	80			R						100	80								R						100	79	
<i>Proteus mirabilis</i>	58	93	58	100	58	88	58	90	58	91	58	R		98	58	100	58	100	54	100	58	100	58	100	58	100	58	R	
ESCAPP group	68	R						79	68	81	67			93	68	75	68	79	62	100	68	81	68	99	68	90	68	R	

■ <70% of isolates sensitive   
 ■ 70-89% of isolates sensitive   
 ■ >90% of isolates sensitive   
  not tested or not clinically effective   
  antibiotic not recommended to be used in children without specialist advice  
R intrinsic resistance

*E. coli* ESBL = 2.8% of isolates

*K. pneumoniae* ESBL = 10.5% of isolates

CRE- 1 isolate of *Enterobacter cloacae*

*Enterococcus* sp - 85 isolates

*Enterococcus faecalis* - 80 isolates (VRE = 0% of isolates)

*Enterococcus faecium* - 4 isolates (VRE = 50% of isolates)

*Staphylococcus aureus* - 27 isolates (MRSA = 26% of isolates)

- NOTE**
1. Data processed by OrgTRx (antibiogram) software to exclude multiple isolates so only the first isolate of a given species per patient per year per subtype (eg. urine) is included
  2. Only organisms with greater than 30 isolates are included (CLSI Guideline M39-A2 recommends that results should include at least 30 isolates to be considered significant)
  3. ESCAPP group includes *Enterobacter*, *Serratia*, *Citrobacter* (excluding *C. koseri*), *Aeromonas*, *Hafnia*, *Providencia*, *Pantoea* and *Morganella* spp.
  4. Antimicrobial susceptibility testing method: EUCAST microbroth dilution and disc diffusion

# Example of cumulative antibiograms ...

Hospital Antibiogram – 2019 Inpatient – All Units		TOTAL ISOLATES	Penicillin	Ampicillin	Ampicillin/sulbactam	Piperacillin/tazobactam	Ceftriaxone	Cefepime	Meropenem	Levofloxacin	Oxacillin	SXT	Nitrofurantoin	Gentamicin	Vancomycin
Period: 01/01/19 – 12/31/19															
Gram (-)	<i>Escherichia coli</i>	2215		41	54	94	85	88	100	77		76	97	91	
	<i>Klebsiella pneumoniae</i>	532			72	95	89	94	99	91		83	35	93	
	<i>Pseudomonas aeruginosa</i>	446				92		91	87	81				90	
Gram (+)	<i>Staphylococcus aureus</i>	821									62	95			100
	MRSA	312									0	94			100
	MSSA	509									100	96			100
	<i>Streptococcus pneumoniae</i>	47	97 (83) <sup>a</sup>				97 (86) <sup>a</sup>			100					100
	<i>Enterococcus spp.</i>	295		87									99		89

# Antibiogram by R

```
1 # AMR works great with dplyr, but it's not required or necessary
2 #Julhas Aug 02, 2023
3 library(AMR)
4 library(dplyr)
5
6 example_isolates %>%
7   mutate(bacteria = mo_fullname()) %>%
8   # filtering functions for microorganisms:
9   filter(mo_is_gram_negative(),
10         mo_is_intrinsic_resistant(ab = "cefotax")) %>%
11   # antibiotic selectors:
12   select(bacteria,
13         aminoglycosides(),
14         carbapenems())
15 antibiogram(example_isolates,
```

11:26 (Top Level) ↓

Console Terminal × Background Jobs ×

R 4.2.2. ~/

```
i For aminoglycosides() using columns 'GEN' (gentamicin), 'TOB' (tobramycin), 'AMK' (amikacin), and 'KAN' (kanamycin)
i For carbapenems() using columns 'IPM' (imipenem) and 'MEM' (meropenem)
# A tibble: 35 × 7
  bacteria          GEN TOB  AMK  KAN  IPM  MEM
  <chr>            <chr> <chr> <chr> <chr> <chr> <chr>
1 Pseudomonas aeruginosa I     S    NA    R    S    NA
2 Pseudomonas aeruginosa I     S    NA    R    S    NA
3 Pseudomonas aeruginosa I     S    NA    R    S    NA
4 Pseudomonas aeruginosa S     S    S    R    NA    S
5 Pseudomonas aeruginosa S     S    S    R    S    S
6 Pseudomonas aeruginosa S     S    S    R    S    S
7 Stenotrophomonas maltophilia R     R    R    R    R    R
8 Pseudomonas aeruginosa S     S    S    R    NA    S
9 Pseudomonas aeruginosa S     S    S    R    NA    S
10 Pseudomonas aeruginosa S     S    S    R    S    S
# i 25 more rows
# i Use `print(n = ...)` to see more rows
> antibiogram(example_isolates,
+             antibiotics = c(aminoglycosides(), carbapenems()))
i using column 'mo' as input for col_mo.
i For aminoglycosides() using columns 'GEN' (gentamicin), 'TOB' (tobramycin), 'AMK' (amikacin), and 'KAN' (kanamycin)
i For carbapenems() using columns 'IPM' (imipenem) and 'MEM' (meropenem)
i 502 combinations had less than minimum = 30 results and were ignored
# A tibble: 10 × 7
  `Pathogen (N min-max)` AMK  GEN  IPM  KAN  MEM  TOB
  * <chr>                <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 CoNS (43-309)          0    86   52    0    52   22
2 E. coli (0-462)       100   98  100   NA   100   97
3 E. faecalis (0-39)    0     0  100    0    NA    0
4 K. pneumoniae (0-58) NA    90  100   NA   100   90
5 P. aeruginosa (17-30) NA   100  NA    0    NA   100
6 P. mirabilis (0-34)  NA    94  94    NA   NA    94
7 S. aureus (2-233)    NA   99  NA    NA   NA    98
8 S. epidermidis (8-163) 0    79  NA    0    NA    51
9 S. hominis (3-80)    NA   92  NA    NA   NA    85
10 S. pneumoniae (11-117) 0     0  NA    0    NA    0
> |
```

# Example of AMR Dashboard

Investments in AMR R&D

Funding Distributors

> All sectors

> Animal

> Environment

> Human

> Plant

Data Trends

Our Data

Antibacterials in Clinical Development

Incentives for Antibacterial R&D

Dynamic Dashboard

## INVESTMENTS IN AMR R&D

Last updated: 18.07.2023

Parameter

Number of Projects

**Total Investments**

Currency

USD EUR

Status

Active Closed

Reset filter

\*by year = incomplete years

Funders **234**

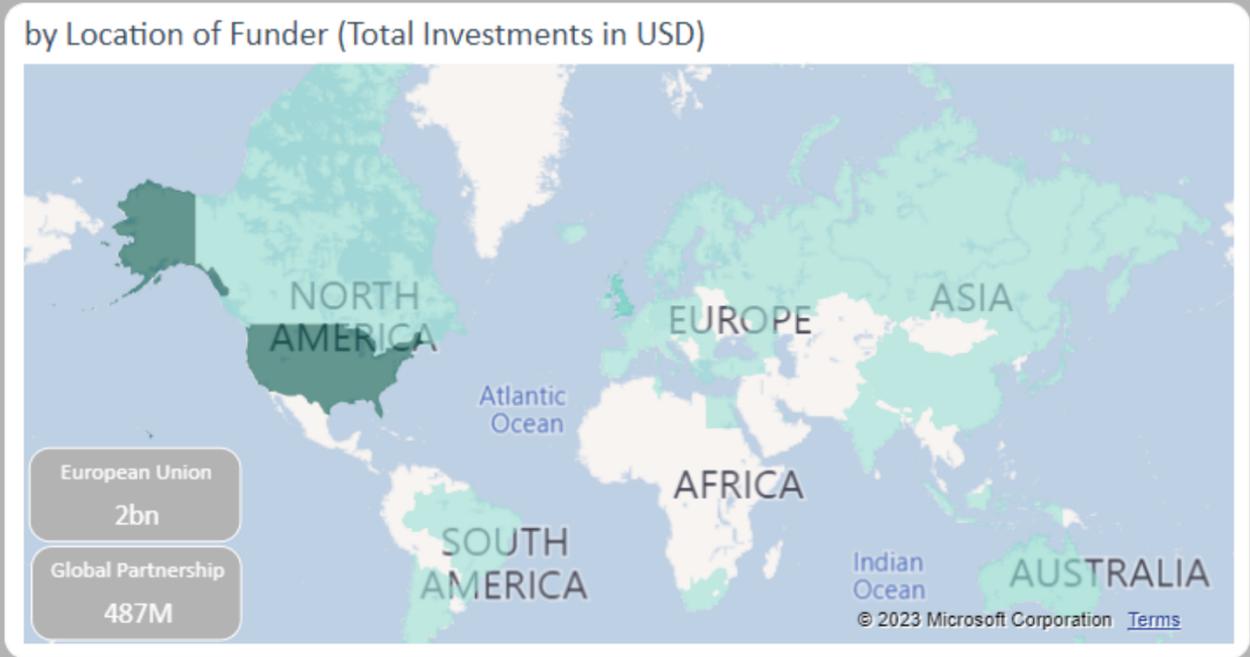
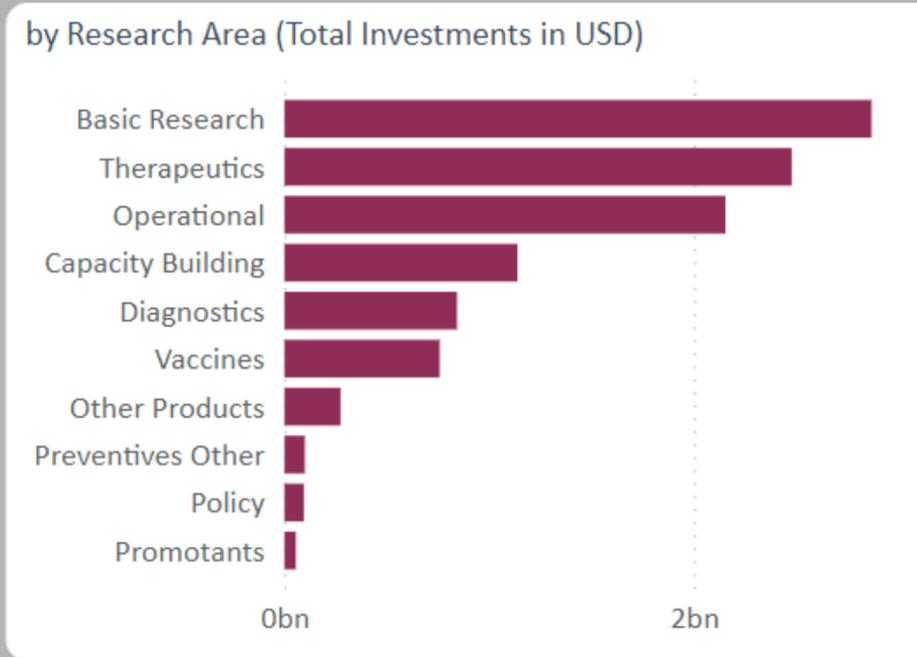
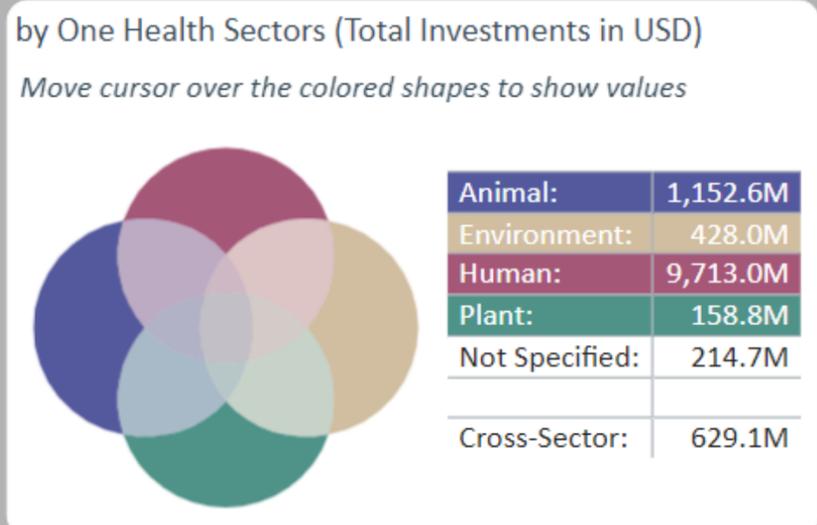
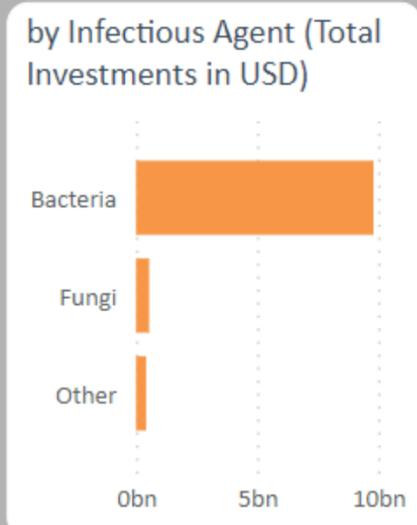
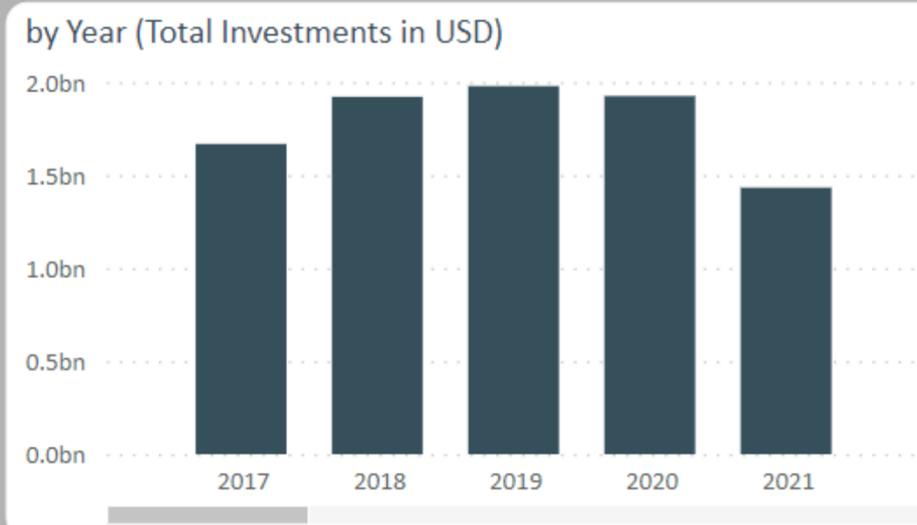
Number of Projects **13159**

Total Investment **USD 10.76bn**

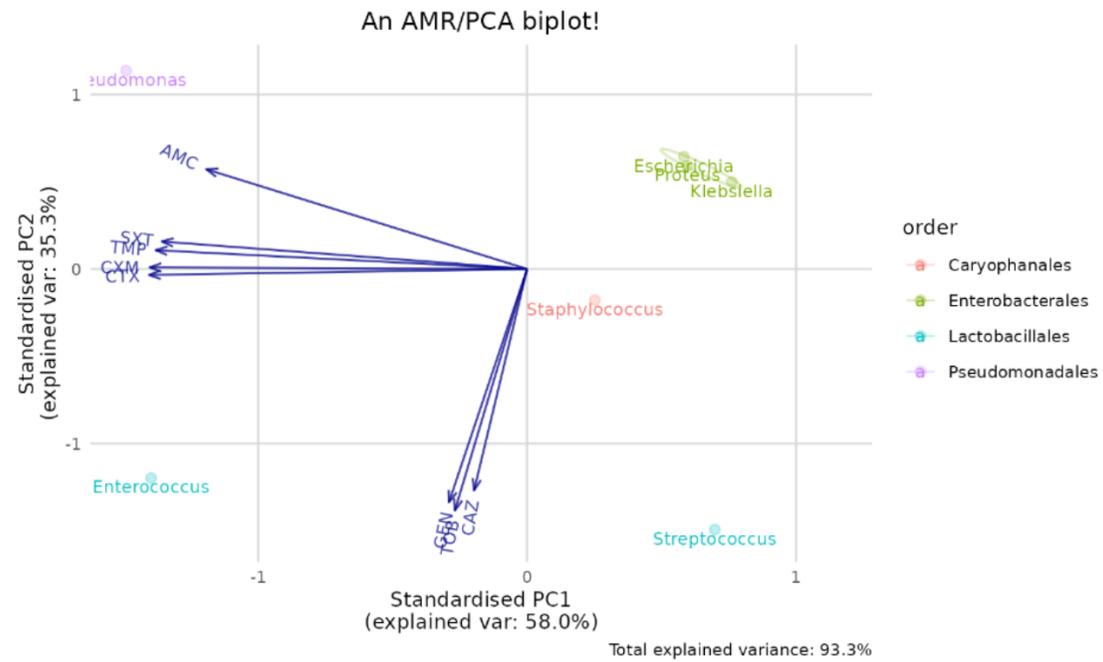
Research Organisations (RO) **2834**

Countries of RO **90**

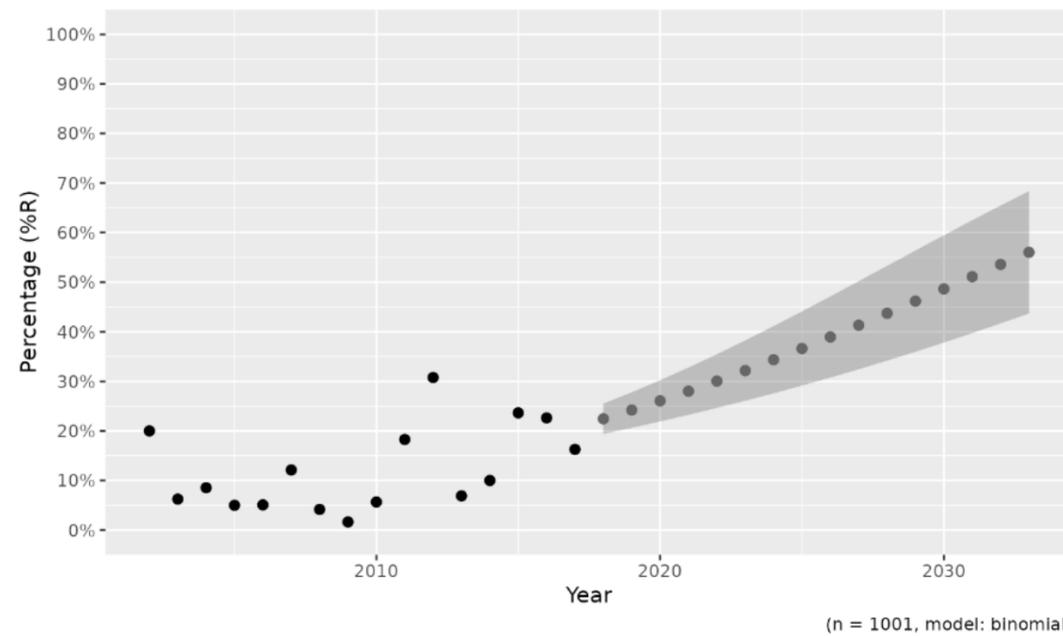
One Health Sectors **4**



# Artificial Intelligence, Machine Learning and Data Science



a) Principal Component Analysis: Example of a ML method



b) Resistance prediction of Piperacillin/tazobactam (TZM)

# Next: Descriptive and inferential statistics

Parametric or non-parametric?	Outcome variable	Number of groups <sup>1</sup>	Statistical test	Key assumptions
Parametric	Categorical: nominal with two levels (dichotomous)	Two or more	Chi-squared test	Expected frequency in any cell of a contingency table is not <5 or no more than 80% of cells have a value of <5
Non-parametric	Categorical: ordinal, or numeric when assumptions for a t-test are not met	Two groups	Mann-Whitney U test (Wilcoxon rank-sum test)	<ul style="list-style-type: none"> <li>• Row and column totals are fixed</li> <li>• Outcome can be ranked</li> </ul>
Non-parametric	Categorical: ordinal, or numeric when ANOVA test assumptions are not met	Three or more groups	Kruskal-Wallis test	Outcome can be ranked
Parametric	Numeric	Two groups	Student's t-test	<ul style="list-style-type: none"> <li>• Normal distribution of outcome variable</li> <li>• Residuals have normal distribution</li> <li>• Variance is the same in both groups (otherwise use modified t-test)</li> </ul>
Parametric	Numeric	Two or more groups	One-way ANOVA	<ul style="list-style-type: none"> <li>• Normal distribution of outcome variable</li> <li>• Variance is the same in all groups</li> </ul>
Parametric	Numeric	Two or more groups	Simple linear regression with one exposure variable	<ul style="list-style-type: none"> <li>• Normal distribution of outcome variable for a given exposure value</li> <li>• Linear relationship (roughly) between exposure and outcome (check with scatterplot)</li> <li>• Homoscedasticity: the variance of residuals is the same for any value of the exposure variable</li> </ul>
Parametric	Categorical: nominal with two levels (dichotomous)	Two groups	Binomial logistic regression	Linear relationship between the exposure and log odds

# THANK YOU

If you have any questions and queries, I will be happy to answer them during the QA session.

Feel free drop an email with your query to: [info@jaetech.co](mailto:info@jaetech.co)