

# Predicting Antimicrobial Resistance from Bacterial Genomes

(and doing something useful with your model)

---

Finlay Maguire

*finlaymaguire@gmail.com*

May 14, 2019

Donald Hill Family Postdoctoral Fellow,  
Faculty of Computer Science, Dalhousie University

# Table of contents

1. Why do we care about AMR?

2. What can we do about it?

3. Genomics

4. Predicting resistance

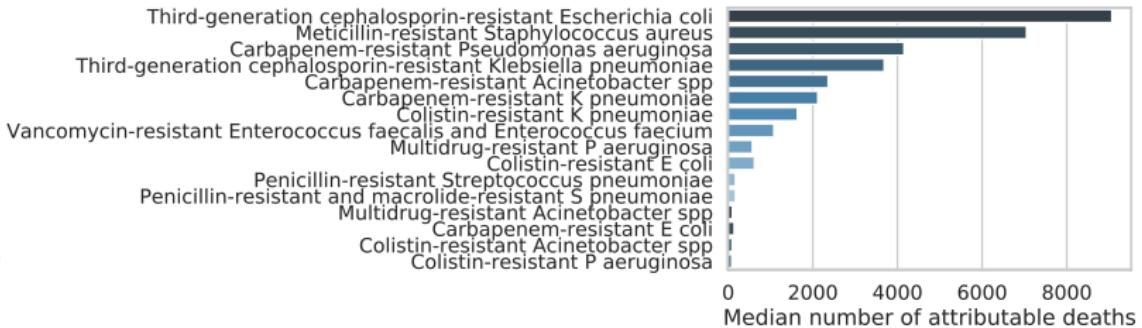
5. Going beyond the model

## Why do we care about AMR?

---

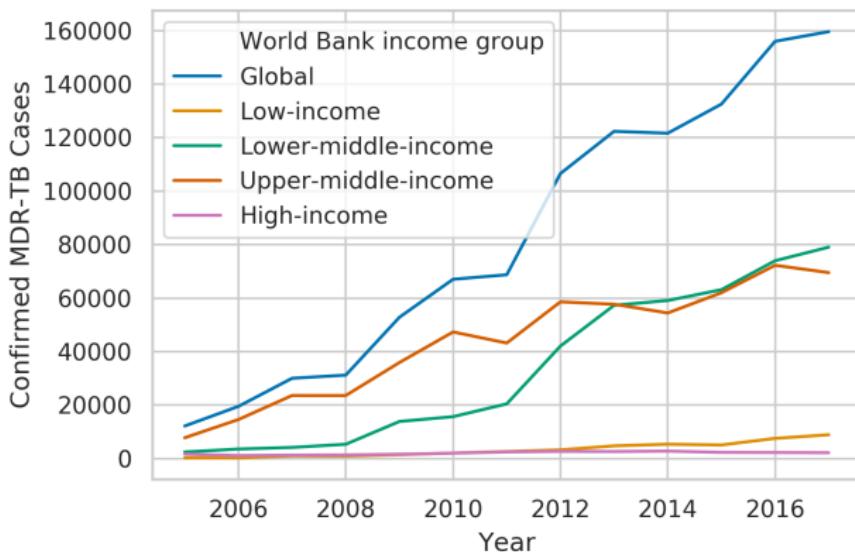
# AMR is currently a problem

EU/EAA Resistant Infections



Overall: 33110 (28480-38430), Data from Cassini et al. (2019).

# AMR is growing



WHO Global Health Observatory Data Repository.

**What can we do about it?**

---

## Improve Surveillance

- Locally: information would help improve patient health.

## Improve Surveillance

- Locally: information would help improve patient health.
- Nationally: health policies and responses to emergencies.

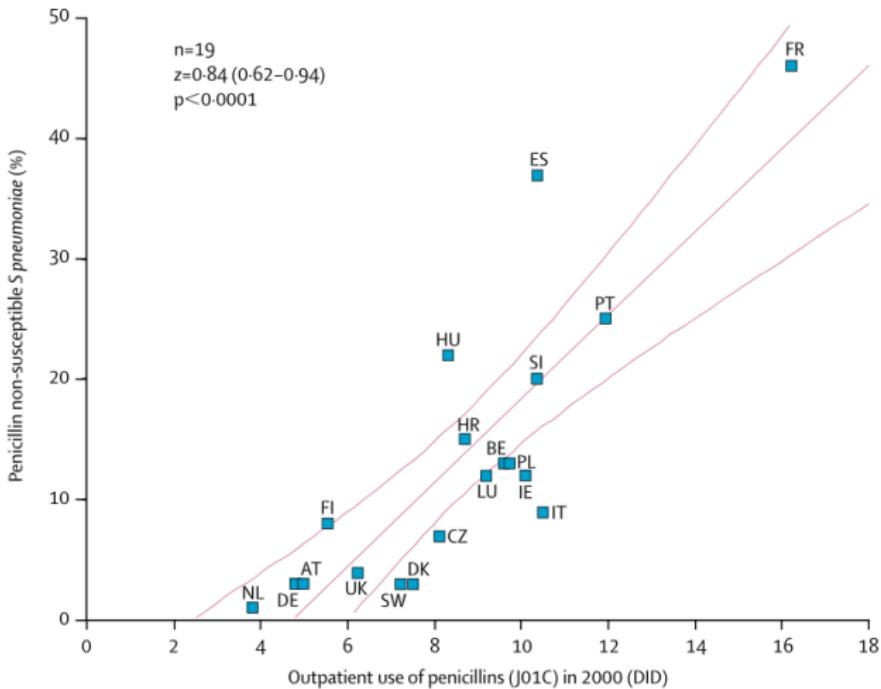
## Improve Surveillance

- Locally: information would help improve patient health.
- Nationally: health policies and responses to emergencies.
- Globally: emerging threats and longterm trends.

## Improve Surveillance

- Locally: information would help improve patient health.
- Nationally: health policies and responses to emergencies.
- Globally: emerging threats and longterm trends.
- Scientifically: better understanding of underlying biology.

# Improve Diagnostics

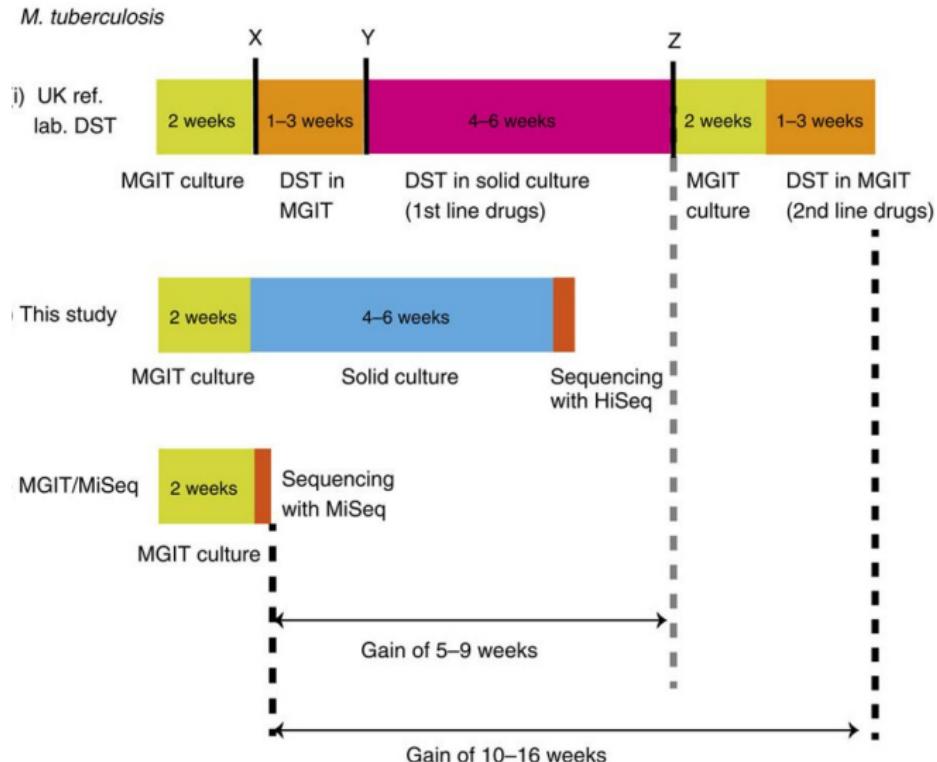


Goossens et al. (2005)

# Genomics

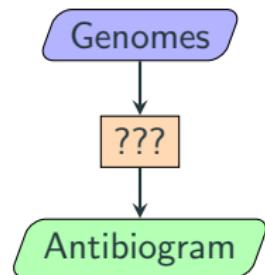
---

# Antibiotic Susceptibility Testing



Bradley et al. (2015)

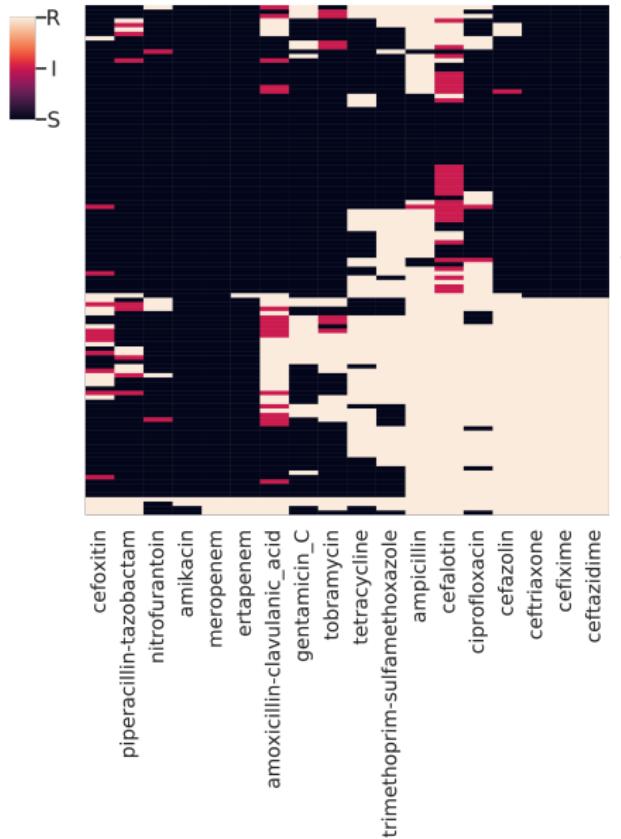
# Linking Genome to Phenotype



## Predicting resistance

---

# Datasets



- 116 *Escherichia coli*
- 127 *Pseudomonas aeruginosa*
- 97 Non-Typhoidal *Salmonella enterica*

# Linking AMR determinants to Phenotype

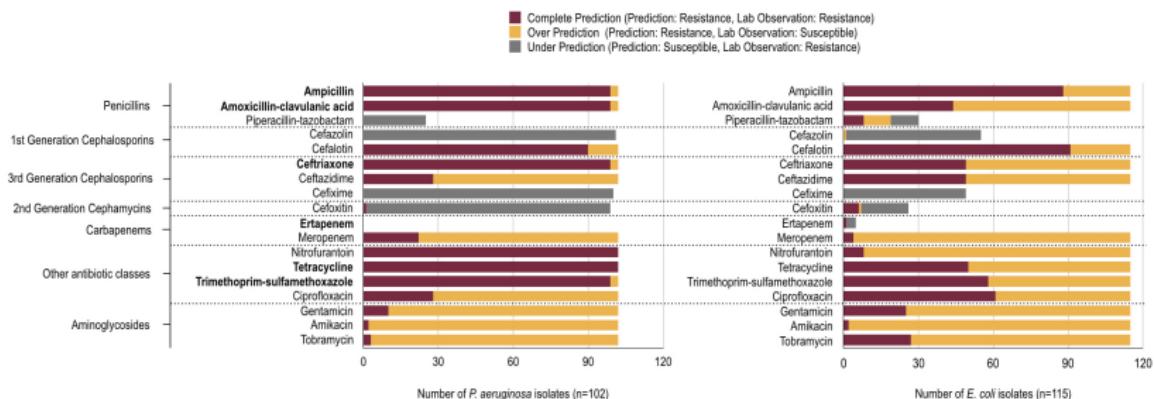


## cmlA1

[Download Sequences](#)

Accession	ARO:3002693
Definition	cmlA1 is a plasmid or transposon-encoded chloramphenicol exporter that is found in <i>Pseudomonas aeruginosa</i> and <i>Klebsiella pneumoniae</i>
AMR Gene Family	<a href="#">major facilitator superfamily (MFS) antibiotic efflux pump</a>
Drug Class	<a href="#">phenicol antibiotic</a>
Resistance Mechanism	<a href="#">antibiotic efflux</a>
Efflux Component	<a href="#">efflux pump complex or subunit conferring antibiotic resistance</a>
Classification	<a href="#">7 ontology terms</a>   <a href="#">Hide</a> + <a href="#">process or component of antibiotic biology or chemistry</a> + <a href="#">mechanism of antibiotic resistance</a> + <a href="#">determinant of antibiotic resistance</a> + <a href="#">antibiotic molecule</a> + <a href="#">antibiotic efflux</a> [Resistance Mechanism] + <a href="#">phenicol antibiotic</a> [Drug Class] + <a href="#">efflux pump complex or subunit conferring antibiotic resistance</a> [Efflux Component]
Parent Term(s)	<a href="#">2 ontology terms</a>   <a href="#">Hide</a> + <a href="#">major facilitator superfamily (MFS) antibiotic efflux pump</a> [AMR Gene Family] + <a href="#">confers_resistance_to_drug chloramphenicol</a> [Antibiotic]
Publications	Bissonnette L, et al. 1991. J Bacteriol 173(14): 4493-4502. Characterization of

# ARO Over-prediction



# Logistic Regression

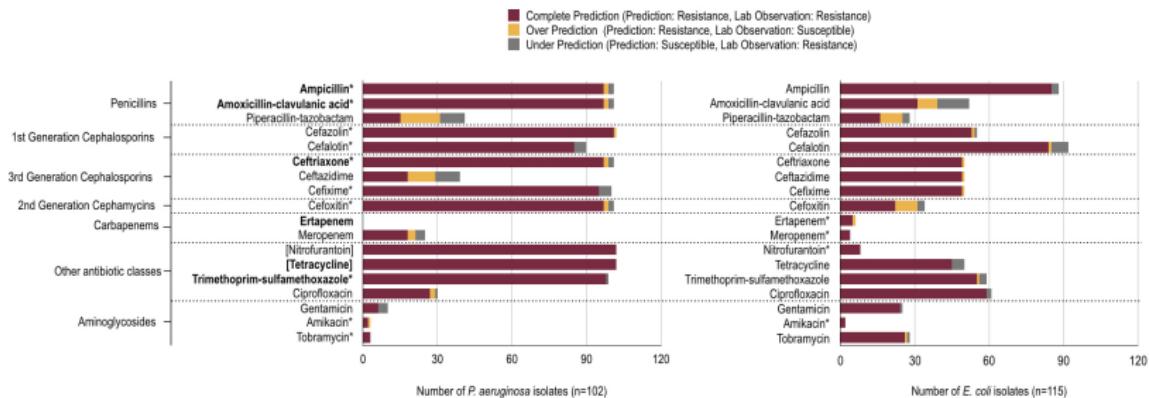


$$RGI = \begin{matrix} & amr_1 & amr_2 & \dots & amr_J \\ genome_1 & 1 & 0 & \dots & 1 \\ genome_2 & 0 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ genome_I & 0 & 0 & \dots & 1 \end{matrix}$$

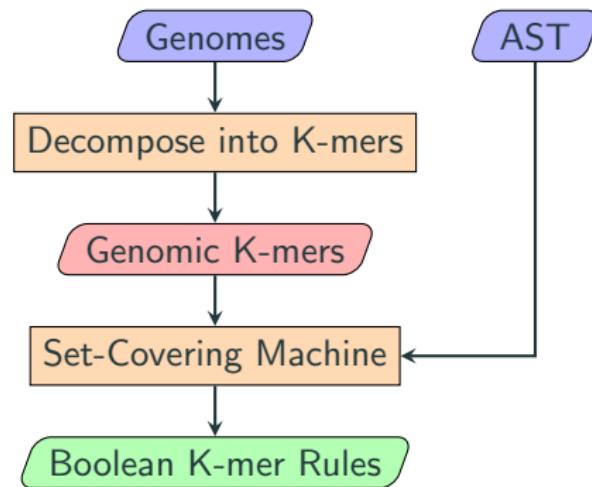
$$AST = \begin{matrix} & abx_1 & abx_2 & \dots & abx_K \\ genome_1 & S & S & \dots & R \\ genome_2 & R & R & \dots & S \\ \dots & \dots & \dots & \dots & \dots \\ genome_I & S & S & \dots & S \end{matrix}$$

$$\beta RGI = AST$$

# Logistic Regression Performance



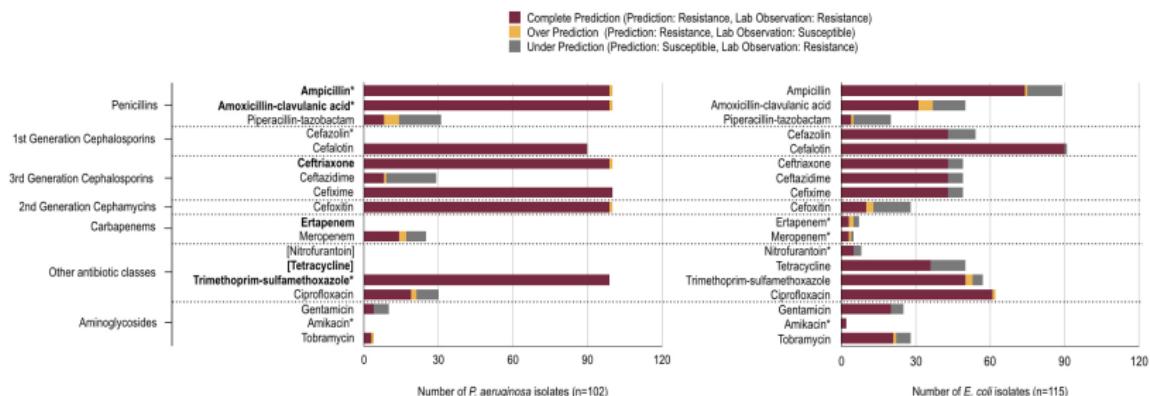
# Set-Covering Machines



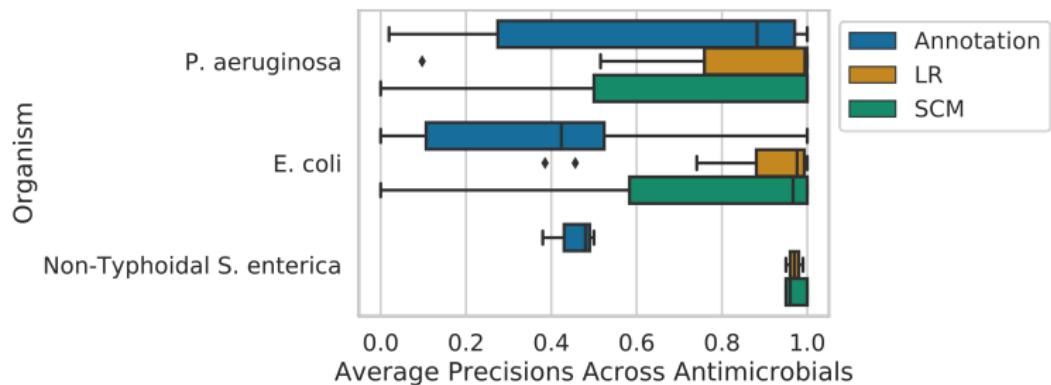
$$kmer_x \wedge kmer_y$$

$$kmer_x \vee kmer_y$$

# SCM Performance



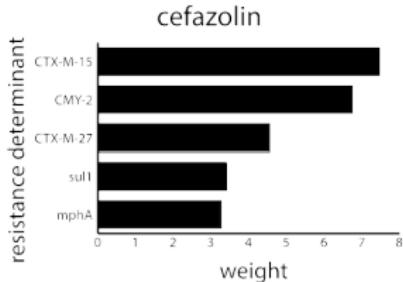
# Overall Performance



## **Going beyond the model**

---

# How is the model working?



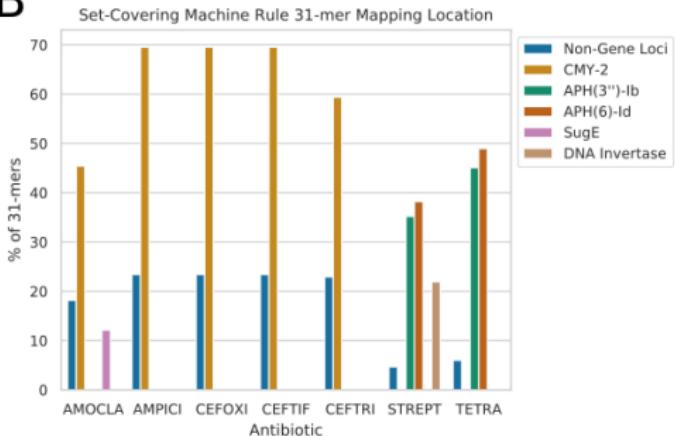
Definition	Cefazolin (INN), also known as cefazoline or cephaZolin, is a first generation cephalosporin antibiotic. It is administered parenterally, and is active against a broad spectrum of bacteria.
Drug Class	<a href="#">cephalosporin</a>
Classification	<a href="#">4 ontology terms</a>   <a href="#">Show</a>
Parent Term(s)	<a href="#">1 ontology terms</a>   <a href="#">Show</a>
Sub-Term(s)	<a href="#">4 ontology terms</a>   <a href="#">Hide</a> + <a href="#">beta-lactam_sensitive_penicillin-binding_protein_targeted_by_drug</a> + <a href="#">porin_OmpC_confers_resistance_to_drug</a> + <a href="#">SHV-1_confers_resistance_to_drug</a> + <a href="#">Nmca_confers_resistance_to_drug</a>

## Experimental Follow-up

Construct	Cefazolin MIC $\mu g/mL$
E. coli BW25113 $\Delta$ bamB $\Delta$ tolC	1-2
E. coli BW25113	4
E. coli BW25113 $\Delta$ bamB $\Delta$ tolC pGDP1:CTX-M-15	128
E. coli BW25113 pGDP1:CTX-M-15	128
E. coli BW25113 $\Delta$ bamB $\Delta$ tolC pGDP1:CMY-2	> 256
E. coli BW25113 pGDP1:CMY-2	> 256
E. coli BW25113 $\Delta$ bamB $\Delta$ tolC pGDP1:CTX-M-27	> 256
E. coli BW25113 pGDP1:CTX-M-27	> 256

# How are costly AMR genes being retained?

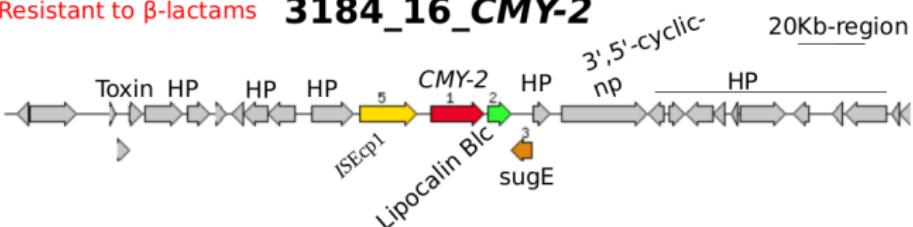
B



HP= Hypothetical proteins

Resistant to  $\beta$ -lactams

**3184\_16\_CMY-2**



# Conclusions

---

- Using annotations and ontologies works very poorly across organisms and resistance mechanisms.

# Conclusions

---

- Using annotations and ontologies works very poorly across organisms and resistance mechanisms.
- Even very simple models can greatly improve predictions.

# Conclusions

---

- Using annotations and ontologies works very poorly across organisms and resistance mechanisms.
- Even very simple models can greatly improve predictions.
- Investigation of how these models work combined with domain expertise can be hugely scientifically informative.

# Conclusions

---

- Using annotations and ontologies works very poorly across organisms and resistance mechanisms.
- Even very simple models can greatly improve predictions.
- Investigation of how these models work combined with domain expertise can be hugely scientifically informative.
- Use models but don't stop at a well-fitted model.

## Acknowledgements

---

# Acknowledgements



- McMaster University: Kara Tsang, Brian Alcock, and Andrew McArthur
- Simon Fraser University: Fiona Brinkman
- Dalhousie University: Robert Beiko
- Funding: Genome Canada, NERC Undergraduate Student Research Award, Donald Hill Family Fellowship

## References

---

- Bradley, P., Gordon, N. C., Walker, T. M., Dunn, L., Heys, S., Huang, B., Earle, S., Pankhurst, L. J., Anson, L., De Cesare, M., et al. (2015). Rapid antibiotic-resistance predictions from genome sequence data for *staphylococcus aureus* and *mycobacterium tuberculosis*. *Nature communications*, 6:10063.
- Cassini, A., Högberg, L. D., Plachouras, D., Quattrocchi, A., Hoxha, A., Simonsen, G. S., Colomb-Cotinat, M., Kretzschmar, M. E., Devleesschauwer, B., Cecchini, M., et al. (2019). Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the eu and the european economic area in 2015: a population-level modelling analysis. *The Lancet Infectious Diseases*, 19(1):56–66.

- Goossens, H., Ferech, M., Vander Stichele, R., Elseviers, M., Group, E. P., et al. (2005). Outpatient antibiotic use in europe and association with resistance: a cross-national database study. *The Lancet*, 365(9459):579–587.
- McArthur, A. G., Waglechner, N., Nizam, F., Yan, A., Azad, M. A., Baylay, A. J., Bhullar, K., Canova, M. J., De Pascale, G., Ejim, L., et al. (2013). The comprehensive antibiotic resistance database. *Antimicrobial agents and chemotherapy*, 57(7):3348–3357.
- Organization, W. H. et al. (2014). *Antimicrobial resistance: global report on surveillance*. World Health Organization.
- Silver, L. L. (2011). Challenges of antibacterial discovery. *Clinical microbiology reviews*, 24(1):71–109.

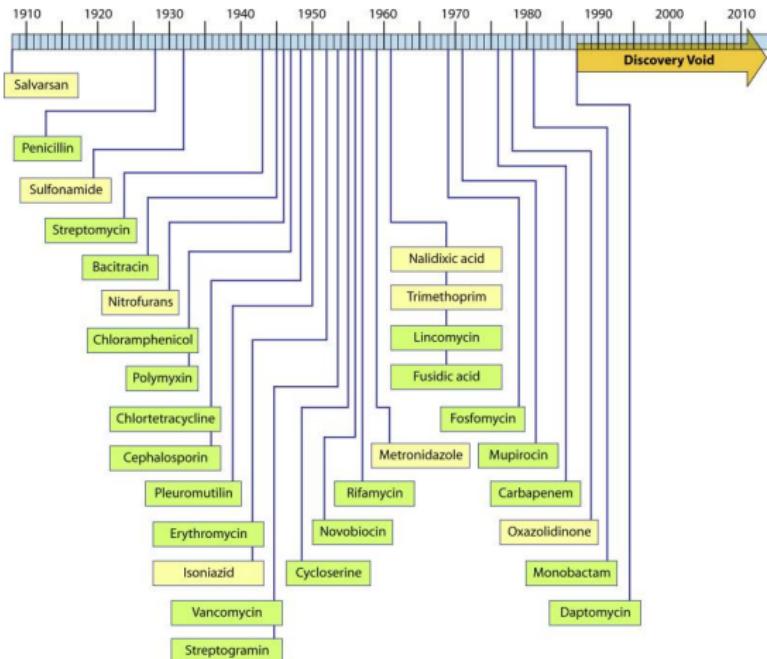
# **Backup**

## Caveat on data

	Antibacterial resistance	Studies included in SR (n)	Country income status <sup>a</sup>	
			Low income (n)	Lower-middle income (n)
<i>Escherichia coli</i>	3 <sup>rd</sup> generation cephalosporin-resistant	25	0	0
	Fluoroquinolone-resistant	12	0	0
<i>Klebsiella pneumoniae</i>	3 <sup>rd</sup> generation cephalosporin-resistant	24	0	0
	Carbapenem-resistant	13	0	0
<i>Staphylococcus aureus</i>	Methicillin-resistant	147	0	2

Studies included in systematic review on healthcare implications of AMR  
Organization et al. (2014)

# New Antibiotics



Silver (2011). Teixobactin in 2015, still in trials.