MIC Predictions for Four β-lactam Agents for *Escherichia coli* and *Klebsiella pneumoniae* from a Large Surveillance Program Using Genomic Data and a Machine Learning Model

Cory Kromer-Edwards[1,2], Suely Oliveira[2], Mariana Castanheira[1]

[1] JMI Laboratories, North Liberty, Iowa, USA
[2] Computer Science Department, University of Iowa, Iowa City, Iowa, USA

Objective

To predict the MIC values for four β-lactam agents for *E. coli* and *K. pneumoniae* isolates using MIC and genomic data from the SENTRY Antimicrobial Surveillance Program using the Random Forrest machine learning algorithm.

Methods

A total of 3054 *E. coli* and 2940 KPN isolates from 2016 and 2017 were susceptibility tested using the CLSI reference broth microdilution method.

Isolates displaying β-lactam and/or aminoglycoside resistance were submitted to whole genome sequencing for the identification of genes encoding β-lactamases.

MIC and genetic results from 2016 and 2017 were used to train the Random Forrest machine learning algorithm.

- MIC results from 2016 to 2018 were predicted for ceftriaxone, ceftazidime, cefepime, and meropenem.
- Error rates were calculated according to CLSI M23 criteria upon comparison to BMD MICs.

Results

Table 1. Error rates for predicted versus tested MIC values for four β-lactam agents against *E. coli* and *K. pneumoniae* isolates

Organism (no. of isolates) Breakpoints[a]	Error type	Ceftriaxone		Ceftazidime		Cefepime		Meropenem	
		No. of isolates	%	No. of isolates	%	No. of isolates	%	No. of isolates	%
E. coli (3054)									
CLSI	Minor	13	1.4	90	9.4	136	14.2	3	0.3
	Major	141	14.8	162	17.0	36	3.8	459	48.0
	Very Major	0	0	41	4.3	42	4.4	0	0
EUCAST	Minor	0	0	77	8.1	35	3.7	8	0.8
	Major	145	15.2	111	11.6	20	2.1	458	47.9
	Very Major	1	0.1	42	4.4	60	6.3	0	0
ECV	Minor	0	0	0	0	0	0	0	0
	Major	128	13.4	28	2.9	0	0	701	73.3
	Very Major	0	0	0	0	0	0	4	0.4
K. pneumoniae (2940)									
CLSI	Minor	18	1.8	50	4.9	72	7.1	142	13.9
	Major	28	2.8	41	4.0	45	4.4	414	40.6
	Very Major	8	0.8	28	2.8	18	1.8	4	0.4
EUCAST	Minor	0	0	13	1.3	31	3.0	59	5.8
	Major	30	2.9	25	2.5	32	3.1	419	41.1
	Very Major	9	0.9	31	3.0	25	2.5	9	0.9
ECV	Minor	0	0	0	0	0	0	0	0
	Major	19	1.9	6	0.6	0	0	482	47.3
	Very Major	4	0.4	12	1.2	0	0	9	0.9

[a] Criteria published by CLSI (2022), EUCAST (2022) and ECVs determined using the entire dataset according to CLSI M57

Conclusions

- Among machine learning prediction methods, the Random Forrest algorithm is capable of learning complex data representations to make accurate predictions that generate random decision trees.
- The Random Forrest machine learning algorithm was able to predict MICs for ceftriaxone for *K. pneumoniae* with acceptable error rates.
- For other cephalosporins, ECVs generated acceptable error rates, but not clinical breakpoints.
- Meropenem MIC predictions had high error rates, potentially due to the small sample of carbapenem-resistant isolates in the studied population.
- Machine learning algorithms should be further explored to predict MIC results, but the use of breakpoints that are established from clinical outcomes and PK/PD outcome studies might not be ideal to interpret these predictions.

Acknowledgements

The authors thank the participants of the SENTRY Antimicrobial Surveillance Program and the JMI Laboratories staff for the testing of these organisms.

Contact

Mariana Castanheira, PhD
mariana-castanheira@jmilabs.com