Complementary data

1

2 3

4 5 6

7

8

Table S1 Best-tuned parameter values for each model of XGBoost.

mtry	trees	min_n	tree_depth	learn_rate	sample_size
7	1407	4	6	0.0095	0.7304

Table S2 Contingency table for categorised predictions (independently > 666 mg*g/L and < 939 mg*h/L AUCs from test and external dataset using XGBoost daptomycin AUC prediction model), applicable in the case of a Staphylococcus aureus infection for a MIC at breakpoint of 1.

TEST SET (n = 1188)					
	Reference AUC Reference AUC		Performances for AUC > 666 mg*h/L:		
	< 666 mg*h/L	> 666 mg*h/L	Se 97.70 %		
Predicted AUC < 666 mg*h/L	390	18	Sp 95.82 %		
Predicted AUC > 666 mg*h/L	17	763	PPV 97.82 % PNV 95.59 %		

	Reference AUC Reference AUC Performances for A		Performances for AUC < 939 mg*h/L:	
	< 939 mg*h/L	> 939 mg*h/L	Se 95.27 %	
Predicted AUC < 939 mg*h/L	625	32	Sp 93.98 %	
Predicted AUC > 939 mg*h/L	31	500	PPV 95.13 % PNV 94.16 %	

EXTERNAL DATASET (n = 194) Reference AUC **Reference AUC** Performances for AUC > 666 mg*h/L: < 666 mg*h/L > 666 mg*h/L Se 100 % Sp 16.67 % Predicted AUC < 666 mg*h/L 1 0 PPV 97.41 % Predicted AUC > 666 mg*h/L 5 188 PNV 100 %

	Reference AUC Reference AUC		Performances for AUC < 939 mg*h/L:		
	< 939 mg*h/L	> 939 mg*h/L	Se 78.00 %		
Predicted AUC < 939 mg*h/L	39	0	Sp 100 %		
Predicted AUC > 939 mg*h/L	11	144	PPV 100 % PNV 92.90 %		

Table S3 Performances of the machine learning models in the training and testing datasets to estimate daptomycin AUCs obtained from two samples.

	Daptomycin q.d. 2 samples					
	Training_glm	Testing_glm	Training_rf	Testing_rf	Training_svm	Testing_svm
RMSE, mg*h/L ^a	88.2	86.2	80.3	79.9	88.2	85.9
R2 ^a	0.981	0.981	0.984	0.984	0.981	0.982
Relative MPE, %	1.25	0.95	0.93	0.65	1.22	0.93
Relative RMSE, %	10.8	9.67	8.27	7.98	10.9	9.73
Number of MPE of	178 (5.01 %)	47 (3.96 %)	101 (2.84 %)	20 (2 26 %)	180 (5.08 %)	EO (4 21 9/)
the ± 20% interval n	1/8 (5.01%)	47 (5.96 %)	101 (2.84 %)	28 (2.36 %)	180 (5.08 %)	50 (4.21 %)

¹² AUC, area under the curve; MPE, mean prediction error; RMSE, root mean square error; glm, generalized

10

11

¹³ linear model; rf, random forest; svm, support vector machine.

^aValues obtained after 10-fold cross-validation. ^b SDs obtained after 10-fold cross-validation