A machine learning model that emulates experts' decision making in vancomycin initial dose planning

Tetsuo Matsuzaki, Yoshiaki Kato, Hiroyuki Mizoguchi, and Kiyofumi Yamada*

Hospital Pharmacy, Nagoya University Graduate School of Medicine, Nagoya, Aichi, 466-8560, Japan

Running title: A prediction model for vancomycin dosing

* Corresponding author: Kiyofumi Yamada

Hospital Pharmacy, Nagoya University Graduate School of Medicine, Nagoya, Aichi, 466-8560, Japan

E-mail: kyamada@med.nagoya-u.ac.jp

Phone: +81-052-744-2674

Fax: +81-052-744-2979

Abstract

Vancomycin is a glycopeptide antibiotic that has been used primarily in the treatment of methicillin-resistant *Staphylococcus aureus* infections. To enhance its clinical effectiveness and prevent nephrotoxicity, therapeutic drug monitoring (TDM) of trough concentrations is recommended.

Initial vancomycin dosing regimens are determined based on patient characteristics such as age, body weight, and renal function, and dosing strategies to achieve therapeutic concentration windows at initial TDM have been extensively studied. Although numerous dosing nomograms for specific populations have been developed, no comprehensive strategy exists for individually tailoring initial dosing regimens; therefore, decision making regarding initial dosing largely depends on each clinician's experience and expertise.

In this study, we applied a machine-learning (ML) approach to integrate clinician knowledge into a predictive model for initial vancomycin dosing. A dataset of vancomycin initial dose plans defined by pharmacists experienced in vancomycin TDM (i.e., experts) was used to build the ML model. The target trough concentration was attained at comparable rates with the model- and expert-recommended dosing regimens, suggesting

that the ML model successfully incorporated the experts' knowledge. The predictive model developed here will contribute to improved decision making for initial vancomycin dosing and early attainment of therapeutic windows.

Keywords: Machine learning, Vancomycin, Initial dosing regimen, TDM

Introduction

Vancomycin is a glycopeptide antibiotic that has been in clinical use for more than 50 years, primarily for the treatment of methicillin-resistant *Staphylococcus aureus* (MRSA) infections (1). To maximize its clinical effectiveness and avoid nephrotoxicity, therapeutic drug monitoring (TDM) is recommended. The ratio of the vancomycin area under the concentration-time curve (AUC) to the minimum inhibitory concentration (MIC), AUC/MIC, is a primary predictor of vancomycin effectiveness. However, since it is challenging to determine the AUC due to the need to obtain multiple serum vancomycin concentrations, it is recommended to monitor the trough serum concentration, which is considered to be the most accurate and practical surrogate marker for the AUC (1). The trough concentration of vancomycin should be above 10 mg/L and below 20 mg/L to avoid development of resistance and nephrotoxicity, respectively (1–3). Thus, the therapeutic range for vancomycin trough levels is 10–20 mg/L, while several guidelines have recommended vancomycin trough levels of 15–20 mg/L for serious invasive MRSA infections such as sepsis (4, 5).

Initial vancomycin dosing consists of a single loading dose followed by a series of maintenance doses; the usual loading dose and maintenance dose are 25–30 mg/kg and 15–20 mg/kg, respectively (4). Loading and maintenance doses are adjusted for patient characteristics such as age, gender, body weight (BW), and renal function. In clinical practice, initial dosing is calculated using population pharmacokinetic parameters, and numerous population pharmacokinetic studies of vancomycin have been conducted to develop nomograms designed to achieve therapeutic trough concentrations at initial TDM (6–8). However, because these nomograms were constructed and validated within specific populations, their robustness to population change is limited. To date, there is no comprehensive strategy for individually optimized initial dosing, and vancomycin dosing decisions are often made based on clinicians’ experience and expertise (hereafter, prior knowledge) (7, 9, 10). In a recent survey in the United States and Canada, many healthcare institutions indicated that institutional-level credentialing or training was important to achieve appropriate vancomycin TDM (11). Indeed, several studies demonstrated that initial dose planning by pharmacists engaged in vancomycin TDM could lead to higher target attainment rates, emphasizing the importance of prior knowledge when conducting vancomycin dose planning (12, 13).

Machine learning (ML) algorithms promote the discovery of new techniques and improve decision making on specific questions involving abundant and multi-dimensional data, and they have emerged as a promising approach for medical research and clinical care. Advances in ML have facilitated the discovery of new biomarkers and mutations related to prognosis, and also the development of automated diagnosis tools (14). Recently, ML approaches have also been used in the field of TDM. Imai and colleagues utilized an ML approach to build a nomogram for initial vancomycin dosing using a dataset of patients treated with vancomycin (6). A study by Huang et al. applied variable engineering and ML methods that enabled integration of high-dimensional data to build a predictive model for maintenance dosing (15).

Given the importance of prior knowledge in vancomycin dose planning, here we sought to integrate such knowledge into an ML model for initial vancomycin dosing. Toward this goal, we used a dataset of vancomycin initial dosing regimens that were defined by pharmacists experienced in TDM (hereafter, TDM experts). This straightforward approach yielded a predictive model that reproduces regimens defined by TDM experts. Notably, a therapeutic window was achieved at a similar rate using ML- and expert-recommended dosing regimens, whereas a previously developed ML model failed to predict dosing

96 regimens that attained a therapeutic window (6).

Materials and methods

Study subjects

This was a single-center, retrospective, observational study of hospitalized patients who received intravenous vancomycin from May 2017 to May 2021 at Nagoya University Hospital. TDM experts were defined as pharmacists who were either (i) engaged in vancomycin TDM or (ii) capable of conducting vancomycin initial dose planning with similar success as the pharmacists defined in (i). Patients who commenced vancomycin treatment with TDM expert–recommended dosing during the study period were included. The exclusion criteria were as follows: under 18 years of age; undergoing peritoneal dialysis or hemodialysis (including continuous hemodiafiltration); receiving multiple loading doses; diagnosed with amyotrophic lateral sclerosis (ALS); and missing data on gender, age, BW, body mass index (BMI), or serum creatinine.

Thirteen patients received two courses of vancomycin treatment with pharmacist-designed initial dosing regimens. In this group, each regimen was independently designed based on patient characteristics at start of vancomycin treatment, and therefore we included each of

these patients as two cases in the study dataset.

Building of the ML model

The dataset used in this study included clinical and routine laboratory data, initial dosing regimens, and serum vancomycin concentrations at initial TDM (if measured). Age, gender, BW, BMI, and creatinine clearance (CL_{CR}) calculated by the Cockcroft-Gault equation were used as parameters to predict initial dosing regimens (including loading and maintenance doses), because we routinely estimated individualized dosing regimens based on these data, as recommended by previous studies (10, 16). The dataset ($n = 106$) was divided into training and testing datasets in a ratio of 84:22 (approximately 80:20).

We developed an ML model by applying random forest (RF) classification to the dataset (see the schematic diagram in Fig. 1). An RF is an ensemble of classification (decision) trees that are generated by sampling data and features in the training dataset (17). Prediction is based on a simple majority vote by decision trees. A hyperparameter is a parameter that affects how well a model is trained, and it therefore controls the model performance. In the RF technique, hyperparameters include the number of decision trees

(n_estimator) and the maximum depth of the trees (max_depth) (18).

K-fold cross-validation was applied on RF classification to select the most appropriate value of each hyperparameter (n_estimator and max_depth) (19). In this process, the dataset of a training group was further divided into two groups, a training-validation set and a testing-validation set, in the ratio of 67:17 (approximately 80:20) using the five-fold cross-validation method (Fig. 2). The models were then built, and the accuracy scores, which were defined as the ratios of correct predictions to the total number of predictions, were measured in each training-validation/testing-validation set, and the mean accuracy score was recorded. This process was repeated with each pair of hyperparameters (n_estimator: 10, 20, 40, 80, and 160; max_depth: 2, 4, 8, 16, and 32), resulting in optimized hyperparameters that achieved the highest mean accuracy score.

Next, we developed the prediction model with optimized hyperparameters. Feature importance, which represents relative importance in terms of the accuracy of the model, was estimated during training phase (18).

Lastly, we evaluated the performance of the predictive model on the testing data. Accuracy scores for loading doses and maintenance doses were used to evaluate the model.

Estimation of trough concentrations with the ML model–recommended dosing regimen

Testing data that included serum vancomycin concentrations were used to estimate vancomycin trough concentrations with the ML model–recommended dosing regimen. In this analysis, we excluded cases in which dosing regimens were changed before initial TDM or in which TDM was conducted after just a single loading dose. We also excluded patients who developed vancomycin-associated acute kidney injury (AKI), which was defined as an increase in the serum creatine level of 0.5 mg/dL or a 50% increase from baseline in at least two consecutive measurements, as is consistent with other studies of vancomycin TDM (20, 21). The trough concentration with the ML model–recommended regimen was estimated using the following equation:

Estimated concentration

$$= \text{Actual concentration} \times \frac{\text{Daily maintenance dose with model regimen}}{\text{Daily maintenance dose with expert regimen}}$$

Comparison between the current ML model and the previous ML model reported by Imai

and colleagues was performed in patients with a BW $\geq$ 40 kg and an estimated glomerular filtration rate (eGFR) $\geq$ 50 mL/min, because the latter model was validated in this population (6).

Statical analysis

We used the Mann-Whitney test for continuous data, and Fisher's exact test or the chi-square test for categorical data. All statistical tests were two-tailed, and p values of less than 0.05 were considered statistically significant except in multiple testing (Table 5), where significance was adjusted by Bonferroni correction. Statistical analyses were performed using the Python statistics module.

Ethics

This study was approved by the ethics committee of Nagoya University Hospital (Approval No. 2021-0189).

Results

Patient characteristics

During the study period, there were 140 cases in 127 patients, with cases defined as instances in which initial dosing regimens of vancomycin were determined by TDM experts. Of these, 16 cases were in patients under 18 years of age, 14 were in patients undergoing peritoneal dialysis or hemodialysis, and two were in patients who received multiple loading doses; all of these cases were excluded from this study. One patient whose BMI value was missing and one patient with ALS were also excluded. The remaining 106 cases in 97 patients were included.

The target attainment rate at initial TDM was 71.1% in the total population, which was comparable to the rates reported in another study of expert (pharmacist)-managed vancomycin TDM (64.3%) (12), thus validating our initial dose planning (Table 1). Eligible cases were randomly assigned to a training group and a testing group in the ratio of 84:22 (approximately 80:20). Overall, patient characteristics in the training and testing groups were well balanced (Table 1).

196

197 **Building an ML model to determine the initial vancomycin dose**

198

199 Next, we sought to build an ML model for predicting individually tailored vancomycin
 200 dosing regimens to achieve the therapeutic window. To this end, we first optimized the
 201 hyperparameters for RF classification, then built a predictive model with optimized
 202 hyperparameters (Fig. 2). Individual features were ranked based on their relative
 203 importance to the accuracy of the model. For loading doses, BW was ranked as the most
 204 important feature, followed by CL_{CR} , BMI, age, and gender, while for maintenance doses,
 205 CL_{CR} was the most important feature, followed by BW, BMI, age, and gender (Fig. 3).

206

207 **ML model evaluation**

208

209 Next, we examined whether the ML model generated the same dosing regimen as TDM
 210 experts. Table 2 summarizes the performance of the model on the testing dataset. The
 211 accuracy scores for loading and maintenance doses were both 63.6% (Table 3). For loading
 212 doses, the differences in doses between the expert- and ML model–recommended regimens

213 fell within the range of 50–150%. For maintenance doses, there were two cases in which
 214 the differences in doses were larger than two-fold (cases no. 6 and no. 18). The rate of
 215 achieving a therapeutic window with ML model–recommended regimens was estimated to
 216 be 73.3%, which was comparable to that with expert-recommended regimens (80.0%, $p =$
 217 1.0; Tables 2 and 4).

218 Lastly, we compared the target attainment rates of our predictive ML model and the ML
 219 model developed by Imai and colleagues (6). Because the latter model was validated in
 220 patients with a $BW \geq 40$ kg and $eGFR \geq 50$ mL/min, we compared cases that met these
 221 criteria. In this group, the target attainment rates of the expert- and ML model–
 222 recommended regimens were 83.3% and 75.0%, respectively, again with no significant
 223 difference ($p = 1.0$; Tables 2 and 5). Importantly, the target attainment rate of regimens
 224 recommended by the previous ML model was much lower, at only 33.3%; this indicates the
 225 poor ability of that model to predict appropriate dosing regimens, at least using the current
 226 testing dataset. However, it should be noted that the differences in target attainment rates
 227 compared to our expert- and ML model–recommended regimens did not reach statistical
 228 significance ($p = 0.0361$ and 0.0995 , respectively. Note that significance was set at 0.0167
 229 after Bonferroni correction).

Discussion

Although considerable effort has been made to develop population-specific dosing nomograms for vancomycin, strategies to determine individually optimized initial dosing regimens remain controversial. In practice, decision making about initial dose planning depends on each clinician's experience regarding the clinical assessment of renal function and the determination of which dosing nomograms to use.

Here, we sought to build an ML model that emulates experts' decision making concerning initial vancomycin dosing. Toward this end, subjects were limited to those who received an initial dosing regimen defined by TDM experts. In contrast with the previous study of Imai et al. (6), we did not exclude patients with a low BW (< 40 kg) or renal dysfunction (eGFR < 50 mL/min) due to the prevalence of such patients and our desire to assess the robustness of the model to changes in patient characteristics.

Feature importance was largely consistent with predictive covariates conventionally used to determine appropriate dosing regimens: BW and CL_{CR} were the most important features for loading and maintenance doses, respectively (10, 22). Although loading doses are generally determined based on actual BW, our results indicate that CL_{CR} is also a key feature for

predicting loading dose (23). This reflects adjustment of loading doses based on renal function, which has been proposed in recent studies (7, 10).

Our ML model scored 63.6% in testing accuracy for both loading and maintenance doses.

These modest accuracy scores may have been due to the difficulty of multi-class classification tasks (24); in the current dataset ($n = 106$), there were six and nine classes (doses) of loading and maintenance doses, respectively (data not shown). In addition, the fact that this study had a relatively small dataset ($n = 106$), which generally leads to overfitting, may have further contributed to a decrease in predictive accuracy (25).

The target attainment rates were similar for our expert- and ML model-recommended regimens (80.0% and 73.3%, respectively; Table 4). This result indicates that our ML model can competently predict individually tailored vancomycin dosing regimens and thus achieve a therapeutic window. Importantly, the target attainment rate of the ML model developed by Imai et al. was 33.3% (6), which is much lower than that of our model (Table 5). This implies that our model has better predictability, but due to the small sample sizes in this study, the differences in target attainment rates between the two models did not reach statistical significance ($p = 0.0995$). The discrepant predictive accuracy of the two models is largely attributed to the differences in the training datasets. Differences in ML algorithms

264 also likely played a role; the decision tree algorithm used in Imai et al. is a classification
265 tree (a nomogram), which is easy to comprehend but is more susceptible to overfitting than
266 the RF algorithm used in the present study (6, 17).

267 The limitations of this study are as follows. First, TDM results were lacking in many study
268 subjects due to the discontinuation of vancomycin treatment before TDM (mainly due to
269 de-escalation). This hindered the evaluation of ML model–recommended dosing regimens
270 regarding their ability to achieve therapeutic levels. Second, the majority of study subjects
271 were older adults, which may have influenced the model’s predictive performance in
272 younger people. Third, our ML model was derived from a single-center, observational study,
273 potentially limiting external generalizability. In addition, it should be noted that the above
274 results are at least partially dependent on the current testing dataset, and thus further
275 evaluation in other populations is required. Lastly and most importantly, the dosing
276 regimens recorded in the dataset were designed to achieve target trough concentrations
277 rather than target AUC/MIC values, raising the concern that our model tends to predict
278 trough-guided dosing (that is, adjusting doses by considering trough concentrations only).

279 This is thought to increase the occurrence of vancomycin-associated AKI compared to
280 AUC-guided dosing, and recently published guidelines do not recommend this approach in

281 patients with serious, invasive MRSA infections (1, 26). A dataset of AUC-guided dosing
 282 regimens is needed to build a prediction model for AUC-guided dosing.
 283 Collectively, we developed a novel ML model to predict individually tailored vancomycin
 284 initial dosing regimens, and this model had a high rate of achieving therapeutic
 285 concentration windows. Our predictive model will aid decision making for initial
 286 vancomycin dosing and contribute to early attainment of a therapeutic range, which is
 287 crucial for clinical and microbiological success in treating serious infections due to MRSA.

288 **Acknowledgments**

289

290 This work was supported by JSPS KAKENHI (Grant Numbers JP 20H03428 and
291 JP19H03532).

292

293 **Conflicts of Interest**

294

295 The authors declare no conflicts of interests.

References

1. Rybak MJ, Le J, Lodise TP, Levine DP, Bradley JS, Liu C, Mueller BA, Pai MP, Wong-Beringer A, Rotschafer JC, Rodvold KA, Maples HD, Lomaestro BM. 2020. Therapeutic monitoring of vancomycin for serious methicillin-resistant *Staphylococcus aureus* infections: A revised consensus guideline and review by the American Society of Health-System Pharmacists, the Infectious Diseases Society of America, the Pediatric Am J Heal Pharm. Oxford University Press.
2. Suzuki Y, Kawasaki K, Sato Y, Tokimatsu I, Itoh H, Hiramatsu K, Takeyama M, Kadota JI. 2012. Is peak concentration needed in therapeutic drug monitoring of vancomycin? a pharmacokinetic-pharmacodynamic analysis in patients with methicillin-resistant *Staphylococcus aureus* pneumonia. *Chemotherapy* 58:308–312.
3. Sakoulas G, Gold HS, Cohen RA, Venkataraman L, Moellering RC, Eliopoulos GM. 2006. Effects of prolonged vancomycin administration on methicillin-resistant *Staphylococcus aureus* (MRSA) in a patient with recurrent bacteraemia. *J Antimicrob Chemother* 57:699–704.
4. Rybak M, Lomaestro B, Rotschafer JC, Moellering R, Craig W, Billeter M,

313 Dalovisio JR, Levine DP, Reilly C. 2009. Therapeutic monitoring of vancomycin in
314 adult patients: A consensus review of the American Society of Health-System
315 Pharmacists, the Infectious Diseases Society of America, and the Society of
316 Infectious Diseases Pharmacists. Am J Heal Pharm. American Society of
317 Health-Systems Pharmacy.

318 5. 2005. Guidelines for the management of adults with hospital-acquired,
319 ventilator-associated, and healthcare-associated pneumonia. Am J Respir Crit Care
320 Med.

321 6. Imai S, Takekuma Y, Miyai T, Sugawara M. 2020. A new algorithm optimized for
322 initial dose settings of vancomycin using machine learning. Biol Pharm Bull 43:188–
323 193.

324 7. Oda K, Katanoda T, Hashiguchi Y, Kondo S, Narita Y, Iwamura K, Nosaka K, Jono
325 H, Saito H. 2020. Development and evaluation of a vancomycin dosing nomogram
326 to achieve the target area under the concentration-time curve. A retrospective study.
327 J Infect Chemother 26:444–450.

328 8. Marqués-Miñana MR, Saadeddin A, Peris JE. 2010. Population pharmacokinetic
329 analysis of vancomycin in neonates. A new proposal of initial dosage guideline. Br J

330 Clin Pharmacol 70:713–720.

331 9. Thalakada R, Legal M, Lau TTY, Luey T, Batterink J, Ensom MHH. 2012.

332 Development and validation of a novel vancomycin dosing nomogram for achieving

333 high-target trough levels at 2 Canadian teaching hospitals. Can J Hosp Pharm

334 65:180–187.

335 10. Broeker A, Nardecchia M, Klinker KP, Derendorf H, Day RO, Marriott DJ, Carland

336 JE, Stocker SL, Wicha SG. 2019. Towards precision dosing of vancomycin: a

337 systematic evaluation of pharmacometric models for Bayesian forecasting. Clin

338 Microbiol Infect 25:1286.e1-1286.e7.

339 11. Kufel WD, Seabury RW, Mogle BT, Beccari M V., Probst LA, Steele JM. 2019.

340 Readiness to implement vancomycin monitoring based on area under the

341 concentration-time curve: A cross-sectional survey of a national health consortium.

342 Am J Heal Pharm 76:889–894.

343 12. Komoto A, Maiguma T, Teshima D, Sugiyama T, Haruki Y. 2018. Effects of

344 pharmacist intervention in Vancomycin treatment for patients with bacteremia due to

345 Methicillin-resistant Staphylococcus aureus. PLoS One 13.

346 13. Momattin H, Zogheib M, Homoud A, Al-Tawfiq JA. 2015. Safety and outcome of

347 pharmacy-led vancomycin dosing and monitoring. *Chemotherapy* 61:3–7.

348 14. Shah P, Kendall F, Khozin S, Goosen R, Hu J, Laramie J, Ringel M, Schork N. 2019.

349 Artificial intelligence and machine learning in clinical development: a translational

350 perspective. *npj Digit Med* 2.

351 15. Huang X, Yu Z, Wei X, Shi J, Wang Y, Wang Z, Chen J, Bu S, Li L, Gao F, Zhang J,

352 Xu A. 2021. Prediction of vancomycin dose on high-dimensional data using machine

353 learning techniques. *Expert Rev Clin Pharmacol* 1–11.

354 16. Crass RL, Dunn R, Hong J, Krop LC, Pai MP. 2018. Dosing vancomycin in the

355 super obese: Less is more. *J Antimicrob Chemother.* Oxford University Press.

356 17. Pavlov YL. 2001. Random forests. *Mach Learn* 45:5–32.

357 18. Probst P, Boulesteix AL, Bischl B. 2019. Tunability: Importance of hyperparameters

358 of machine learning algorithms. *J Mach Learn Res* 20:1–32.

359 19. Kohavi R. 1995. A Study of Cross-Validation and Bootstrap for Accuracy

360 Estimation and Model Selection. *Proc 14th Int Jt Conf Artif Intell* 338–345.

361 20. Lodise TP, Lomaestro B, Graves J, Drusano GL. 2008. Larger vancomycin doses (at

362 least four grams per day) are associated with an increased incidence of

363 nephrotoxicity. *Antimicrob Agents Chemother* 52:1330–1336.

- 364 21. Neely MN, Kato L, Youn G, Kraler L, Bayard D, Van Guilder M, Schumitzky A,
365 Yamada W, Jones B, Minejima E. 2018. Prospective trial on the use of trough
366 concentration versus area under the curve to determine therapeutic vancomycin
367 dosing. *Antimicrob Agents Chemother* 62.
- 368 22. Vandecasteele SJ, De Vriese AS. 2010. Recent changes in vancomycin use in renal
369 failure. *Kidney Int*.
- 370 23. Li J, Udy AA, Kirkpatrick CMJ, Lipman J, Roberts JA. 2012. Improving
371 vancomycin prescription in critical illness through a drug use evaluation process: A
372 weight-based dosing intervention study. *Int J Antimicrob Agents* 39:69–72.
- 373 24. Sebban M, Nock R, Chauchat JH, Rakotomalala R. 2000. Impact of Learning Set
374 Quality and Size. *Int J Comput Syst Signal* 1:85–105.
- 375 25. D'souza RN, Huang PY, Yeh FC. 2020. Structural Analysis and Optimization of
376 Convolutional Neural Networks with a Small Sample Size. *Sci Rep* 10.
- 377 26. Oda K, Jono H, Nosaka K, Saito H. 2020. Reduced nephrotoxicity with vancomycin
378 therapeutic drug monitoring guided by area under the concentration–time curve
379 against a trough 15–20 µg/mL concentration. *Int J Antimicrob Agents* 56.

Figure legends

Fig. 1 Schematic diagram of RF classification. (A) Training samples and features are randomly sampled; then, classification (decision) trees are constructed with each set of samples and features. The $n_estimator$ is defined as the number of trees. (B) Each decision tree classifies the test sample. The class of the test sample is predicted by majority vote (in this case, the test sample is assigned to class B)

Fig. 2 Schematic diagram of steps in the optimization of hyperparameters. Step 1: The dataset is first split into a training set and a testing set in a ratio of 84:22. Step 2: The training set is further split into five partitions, that is, 67:17. Then four partitions and the remaining partition were used as the training and testing sets, respectively, and prediction accuracy was evaluated. The same procedure was repeated for each subset of the dataset, and the mean accuracy score was recorded. Step 3: Step 2 was repeated with different sets of hyperparameters ($n_estimators$ and max_depth), and mean accuracy scores were recorded in each iteration. The set of hyperparameters with the highest mean accuracy scores in Step 2 was obtained as the optimized parameters and used for model construction.

399

400 **Fig. 3** Feature importance in the ML-based prediction model.

Step 1. Splitting the dataset

Step 2. Cross-validation

Repeat for each parameter

Step 3. Optimization of hyperparameters

In this case, $n_estimator = 40$ and $max_depth = 8$ are the optimized hyperparameters.

Fig. 2 Schematic diagram of steps in the optimization of hyperparameters. Step 1: The dataset is first split into a training set and a testing set in a ratio of 84:22. Step 2: The training set is further split into five partitions, that is, 67:17. Then four partitions and the remaining partition were used as the training and testing sets, respectively, and prediction accuracy was evaluated. The same procedure was repeated for each subset of the dataset, and the mean accuracy score was recorded. Step 3: Step 2 was repeated with different sets of hyperparameters ($n_estimators$ and max_depth), and mean accuracy scores were recorded in each iteration. The set of hyperparameters with the highest mean accuracy scores in Step 2 was obtained as the optimized parameters and used for model construction.

Fig. 3 Feature importance in the ML-based prediction model.

TABLE 1 Clinical characteristics of the study subjects

Characteristics	Total population (n = 106)	Training group (n = 84)	Testing group (n = 22)	P value (training vs testing)
Age (year) [median (IQR)]	70.0 (20.3)	68.5 (21.5)	73.0 (15.8)	0.227 ^a
Gender				
Male [n (%)]	69 (65.1)	57 (67.9)	12 (54.5)	0.244 ^b
Female [n (%)]	37 (34.9)	27 (32.1)	10 (45.5)	
BW (kg) [mean $\pm$ SD]	56.3 $\pm$ 13.8	55.8 $\pm$ 13.9	58.3 $\pm$ 13.5	0.391 ^a
BMI [mean $\pm$ SD]	21.6 $\pm$ 5.1	21.2 $\pm$ 4.7	23.1 $\pm$ 6.2	0.160 ^a
CL _{CR} (mL/min) [mean $\pm$ SD]	85.5 $\pm$ 59.2	88.5 $\pm$ 63.5	74.1 $\pm$ 37.8	0.654 ^a
Target attainment at initial TDM [achieved/total (%)]	54/76 (71.1%)	42/61 (68.9%)	12/15 (80.0%)	0.532 ^c

Abbreviations: IQR, interquartile range; BW, body weight; SD, standard deviation; BMI, body mass index; CL_{CR}, creatinine clearance;
TDM, therapeutic drug monitoring.

^aMann-Whitney U test, ^bChi-square test, ^cFisher's exact test

TABLE 2 Model-recommended dosing regimens and estimated trough concentrations in the testing dataset

Case no.	Gender	BW (kg)	BMI	Age (year)	CL _{CR} (mL/min)	eGFR (mL/min)	Loading dose (mg)		Maintenance dose			Trough concentration (mg/L)			AKI
							Expert	Current ML model	Expert (mg/day)	Current ML model (mg/day)	Previous ML model [mg/day (mg/kg/day)] ^a	Expert (actual concentration)	Current ML model ^b	Previous ML model ^{b,c}	
1	Male	46.0	21.3	76	60.1	67.3	1000	1000	1500	1500	1297 (28.2)	10.77	10.77	9.31	No
2	Female	52.6	20.5	70	71.3	64.5	1000	1500	2000	1500	1483 (28.2)	12.76	9.57	9.46	No
3	Female	62.0	26.0	77	69.9	60.2	1500	1000	2000	2000	1748 (28.2)	11.92	11.92	10.42	No
4	Male	58.3	20.0	38	103.2	84.7	1250	1500	2000	2000	ND	NA	ND	ND	No
5	Female	49.3	24.7	88	32.9	34.2	1000	1000	500	500	ND	NA	ND	ND	No
6	Male	54.2	20.7	82	21.4	22.7	1000	1000	500	1000	ND	19.80	ND	ND	AKI
7	Male	58.9	22.2	80	75.4	83.2	1000	1500	2000	2000	1630 (27.7)	11.70	11.70	9.54	No
8	Male	47.5	16.4	78	56.0	69.7	1000	1000	1000	1000	1340 (28.2)	19.00	19.00	25.45	No
9	Male	58.8	22.5	82	20.3	20.3	1000	1000	400	500	ND	NA	ND	ND	No
10	Female	48.3	19.7	79	30.5	29.8	1000	1000	750	750	ND	16.91	16.91	ND	No
11	Male	65.2	21.4	65	97.0	89.5	1500	1500	2000	2000	ND	NA	ND	ND	No
12	Male	69.5	24.5	71	96.5	88.7	1500	1500	2000	2000	2537 (36.5)	7.19	7.19	9.12	No
13	Male	55.7	21.4	47	71.2	58.1	1500	1500	2000	1500	ND	NA	ND	ND	No
14	Female	103.0	46.4	43	190.2	92.1	2000	2000	3000	2000	3729 (36.2)	17.60	11.73	21.87	No
15	Female	71.0	29.2	72	74.0	55.3	1000	1000	2000	2000	1420 (20.0)	15.42	15.42	10.95	No
16	Male	66.0	23.7	74	68.0	64.5	1500	1000	2000	2000	1861 (28.2)	13.52	13.52	12.58	No
17	Female	37.1	17.2	75	73.0	83.6	1000	1000	1500	1500	ND	16.37	16.37	ND	No
18	Female	45.5	17.1	63	137.9	137.7	1000	1250	750	1500	ND	NA	ND	ND	No
19	Female	63.5	26.1	61	56.4	39.4	1500	1000	1500	2000	ND	25.27	33.69	ND	No
20	Male	48.2	18.8	68	68.9	72.9	1250	750	2000	1500	1759 (36.5)	13.70	10.28	12.05	No
21	Female	70.3	28.1	46	90.7	56.1	1500	1500	2000	2000	1863 (26.5)	33.00	33.00	30.74	No
22	Male	51.0	19.4	77	65.6	75.1	1000	1000	1500	1500	1413 (27.7)	10.03	10.03	9.45	No

Abbreviations: BW, body weight; BMI, body mass index; CL_{CR}, creatinine clearance; eGFR, estimated glomerular filtration rate; ML, machine learning; ND, data not determined; NA, data not available; AKI, acute kidney injury.

^aDose per body weight (mg/kg/day) is also described, because a previous ML model predicted doses in this format.

^bConcentrations were not determined if patients developed AKI before TDM.

^cPatients with a BW ≥ 40 kg and an eGFR ≥ 50 mL/min were included.

TABLE 3 Relative ML-recommended doses

Ratio of ML- to expert-recommended dose	Loading dose (n = 22)	Maintenance dose (n = 22)
> 150% [n (%)]	0 (0.0)	2 (9.1)
> 125% and ≤ 150% [n (%)]	2 (9.1)	1 (4.5)
> 100% and ≤ 125% [n (%)]	2 (9.1)	1 (4.5)
100% (identical)	14 (63.6)	14 (63.6)
≥ 75% and < 100% [n (%)]	0 (0.0)	3 (13.6)
≥ 50% and < 75% [n (%)]	4 (18.2)	1 (4.5)
< 50% [n (%)]	0 (0.0)	0 (0.0)

Abbreviation: ML, machine learning.

TABLE 4 Rates of achieving targeted trough levels with the expert- and current ML model–driven dosing regimens

Trough concentration	Expert (n = 15)	Current ML model (n = 15)	P value (Fisher's exact test)
< 10 mg/L [n (%)]	1 (6.7)	2 (13.3)	1.0
10–20 mg/L [n (%)]	12 (80.0)	11 (73.3)	
> 20 mg/L [n (%)]	2 (13.3)	2 (13.3)	

Abbreviation: ML, machine learning.

TABLE 5 Rates of achieving targeted trough levels with the current and previous ML model–driven dosing regimens

Trough concentration	Expert (n = 12)	Current ML model (n = 12)	Previous ML model * (n = 12)
< 10 mg/L [n (%)]	1 (8.3)	2 (16.7)	5 (41.7)
10–20 mg/L [n (%)]	10 (83.3)	9 (75.0)	4 (33.3)
> 20 mg/L [n (%)]	1 (8.3)	1 (8.3)	3 (25.0)
P value (Fisher's exact test, vs. expert)		1.0	0.0361
P value (Fisher's exact test, current vs. previous ML)		0.0995	

Abbreviation: ML, machine learning.

* Imai S et al. 2020 (6)