Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Priori Predictive Performance of Four Model-Informed Precision Dosing Software Programs for Vancomycin in Hospitalized Patients: A Single-Center Retrospective Study

Qiao Wang
Nanjing Medical University
JiaXing Wei
Nanjing University
XinYan Bai
Nanjing University of Chinese Medicine
YuMeng Li
Nanjing University of Chinese Medicine
YunJuan Zhao
Nanjing University of Chinese Medicine
HaiPing Wu
Nanjing University
GuoHua Zhou
Nanjing Medical University
XueJiao Liu
liuxuejiao0615@163.com
Nanjing University

Research Article

Keywords: vancomycin, population pharmacokinetics, model-informed precision dosing, a priori prediction, prediction error

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11553704/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Purpose

We aimed to evaluate the *a priori* predictive performance of four model-informed precision dosing (MIPD) software programs and identify factors associated with prediction error in inpatients receiving vancomycin.

Methods

We retrospectively included patients who received intravenous vancomycin and underwent therapeutic drug monitoring (TDM). C_{min}, C_{max}, and AUC_{0-24} were predicted using SmartDose, PharmVAN, NextDose, and IOA. Performance was evaluated by prediction error (PE), absolute prediction error (APE), F30 (proportion of APE ≤ 30%), root mean square error (RMSE), and classification agreement rate (CAR). For the best-performing program, multivariable generalized linear modeling (GLM) identified factors associated with AUC_{0-24} APE.

Results

IOA demonstrated the most favorable overall performance. Its median APE and F30 were 29.8% and 51.0% for C_{min}; 15.3% and 77.6% for C_{max}; and 20.9% and 70.9% for AUC_{0-24}, respectively. IOA also had the lowest RMSEs (7.74 mg/L, 10.86 mg/L, and 202.68 mg·h/L) and the largest relative radar-chart area (58.9%). Acute kidney injury (AKI; Exp(*B*) = 2.103, *P* = 0.024) and low daily dose (< 2 g/day; Exp(*B*) = 1.498, *P* = 0.030) were independently associated with greater AUC_{0-24} APE. Patients with AKI had lower F30 (20.0% vs 73.7%) and higher RMSE (507.16 vs 171.64 mg·h/L) than those without.

Conclusions

IOA showed relatively favorable *a priori* performance, but errors persisted, particularly in patients with AKI. *A priori* predictions should be interpreted cautiously in unstable renal function, and early TDM incorporation is recommended for Bayesian updating.

Introduction

Vancomycin is widely used to treat severe Gram-positive bacterial infections, particularly those caused by methicillin-resistant *Staphylococcus aureus* (MRSA) [1]. Its narrow therapeutic window and substantial inter- and intraindividual pharmacokinetic variability, driven by renal function, disease severity, and physiological changes, complicate the selection of empirical doses that provide safe and effective exposure [2]. Insufficient exposure may increase the risk of treatment failure, whereas excessive exposure increases the risk of toxicity, particularly acute kidney injury. The 2020 consensus guideline recommends an area under the concentration-time curve to minimum inhibitory concentration ratio (AUC_{0-24}/MIC)-guided approach for serious MRSA infections, with a target AUC_{0-24} of 400–600 mg·h/L when the MIC is 1 mg/L, and supports Bayesian estimation based on population pharmacokinetic (PopPK) models [3]. Nevertheless, target attainment remains suboptimal in real-world practice, with one study reporting an AUC_{0-24}/MIC target attainment rate of only 27.9% [4, 5]. Therefore, improving *a priori* prediction of vancomycin exposure before therapeutic drug monitoring (TDM) data become available remains an important challenge in individualized dosing.

With the increasing adoption of model-informed precision dosing, several PopPK-based Bayesian software programs have been developed for vancomycin dose optimization. Before measured concentrations become available for Bayesian updating, predictions rely primarily on the population prior and available patient covariates; therefore *a priori* performance depends on the external applicability of the underlying PopPK model [6]. External evaluations have demonstrated substantial variability in the predictive performance of vancomycin PopPK models across independent patient populations, highlighting the importance of compatibility between models and patient populations [7]. However, comparative evaluations of Bayesian dosing software programs remain limited. Bai et al. recently compared three Bayesian software programs (SmartDose, Pharmado, and PrecisePK) in Chinese ICU patients by evaluating their ability to predict vancomycin concentrations under *a priori* and *a posteriori* conditions. Using relative bias and relative root mean square error (rRMSE) as evaluation metrics, they assessed the accuracy and precision of predicted drug concentrations before and after Bayesian updating with measured concentrations. They reported acceptable bias but limited precision in *a priori* concentration prediction, with improved predictive performance after observed concentrations were incorporated [8]. However, their study was limited to critically ill patients and primarily focused on concentration prediction accuracy. The comparative performance of different Bayesian software programs across broader hospitalized populations and clinically relevant exposure metrics, particularly AUC_{0-24}-based assessment, remains insufficiently characterized.

Accordingly, this study compared the *a priori* predictive performance of four widely used Bayesian vancomycin dosing software programs - SmartDose, PharmVAN, NextDose, and IOA - selected based on their broad application, high accessibility, and availability for clinical or research use. Using real-world TDM data from Chinese patients, the predictive performance of these programs was evaluated for C_{min}, C_{max}, and AUC_{0-24} in

terms of bias, accuracy, prediction acceptability, and classification agreement. We also examined variation in AUC_{0-24} prediction errors across clinical characteristics and applied generalized linear modeling (GLM) to identify factors independently associated with absolute prediction error. These findings may help inform the selection and refinement of Bayesian dosing software for initial individualized vancomycin therapy in clinical practice.

Methods

Study Design and Patients

This was a single-center retrospective observational study conducted at Jinling Clinical Medical College of Nanjing Medical University. During the study period, model-informed precision dosing (MIPD) was not routinely used to guide initial vancomycin dosing. Vancomycin regimens were prescribed by treating physicians based on patients' clinical characteristics and subsequently adjusted according to therapeutic drug monitoring (TDM) results. Therefore, all software-based predictions evaluated in this study were generated retrospectively and did not influence actual clinical dosing decisions.

Consecutive intravenous vancomycin treatment courses among hospitalized patients who underwent TDM between August 1, 2025, and April 30, 2026, were screened. Courses were eligible if they met the following criteria: (1) patients aged ≥ 18 years; (2) intermittent intravenous vancomycin administration for ≥ 2 days; (3) availability of at least one paired peak-trough concentration measurement; (4) sufficient dosing and sampling information for pharmacokinetic evaluation, including administration, infusion, and blood sampling times. Courses were excluded if they involved (1) outpatient treatment; (2) continuous intravenous infusion; (3) renal replacement therapy or extracorporeal membrane oxygenation; (4) vancomycin concentrations outside the validated quantification range of the clinical assay (3-100 mg/L); (5) incomplete clinical, dosing, or sampling information required for pharmacokinetic calculations and software-based predictions. The screening process is illustrated in Supplementary Figure S1. The study was approved by the Ethics Committee of Jinling Clinical Medical College of Nanjing Medical University (approval no. 2026DZKY-148-01).

Clinical Data Collection and Variable Definitions

Demographic characteristics, renal function and other laboratory parameters, hospital department, vancomycin dosing regimens, dosing and sampling times, serum vancomycin concentrations, and concomitant medications were collected. Body mass index (BMI) was calculated from height and weight, and estimated glomerular filtration rate (eGFR) was calculated using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation [9].

Acute kidney injury (AKI) was defined according to the serum creatinine criteria in the Kidney Disease: Improving Global Outcomes (KDIGO) Clinical Practice Guidelines for AKI [10], as an increase in serum creatinine by ≥ 0.3 mg/dL (≥ 26.5 μmol/L) within 48 h or to ≥ 1.5 times baseline within 7 days. The urine output criterion was not applied because complete and reliable serial urine output data were unavailable. Patients were categorized according to their actual daily vancomycin dose into low-dose (< 2 g/day), standard-dose (2 g/day), or high-dose (> 2 g/day) groups. Variables with > 30% missing data were excluded from subsequent analyses, whereas those with ≤ 30% missingness were handled using multiple imputation with predictive mean matching (PMM).

Calculation of Reference AUC

The AUC_{0-24} was estimated from paired peak-trough concentration measurements and served as the reference AUC_{0-24} for evaluating the predictive performance of each software program. The reference AUC_{0-24} was calculated using a modified trapezoidal method, as previously described [11]. The equations were as follows:

$$K_e = \frac{\ln(C_2/C_1)}{(t_1 - t_2)} \quad \text{equation}(1)$$

$$C_{max} = C_2 / e^{-K_e \times (t_2 - t_{max})} \quad \text{equation}(2)$$

$$C_{min} = C_1 / e^{-K_e \times (t_{min} - t_1)} \quad \text{equation}(3)$$

$$AUC_{0-24} = \left(\frac{t_{inf} \times (C_{max} + C_{min})}{2} + \frac{(C_{max} - C_{min})}{K_e}\right) \times \text{nequation}(4)$$

where C_1 and C_2 represent the measured trough and peak concentrations, respectively; t_1 and t_2 represent the corresponding sampling times (in hours); K_e is the elimination rate constant; C_{max} and C_{min} represent the extrapolated peak and trough concentrations at steady state, respectively;

t_{max} represents the end of the current infusion, whereas t_{min} represents the start of the next infusion; t_{inf} is the infusion duration; and n is the number of vancomycin doses administered within 24 h.

C_{max} and C_{min} were obtained by backward and forward extrapolation using Eqs. (1)–(3), respectively, and were then entered into Eq. (4) to calculate the reference AUC_{0-24} for each paired measurement.

Four Model-Informed Precision Dosing Software Programs

Four web-based MIPD software programs - SmartDose (http://SmartDose.cn/, Shanghai Jiao Tong University, China), PharmVAN (http://www.pharmado.net/, Department of Pharmacy, Peking University Third Hospital, China), NextDose (https://www.NextDose.org/, University of Gothenburg, Sweden), and IOA (https://huashan.99rongle.com/companyLogin, Huashan Hospital, Fudan University, China) in this study. Detailed input requirements and operating procedures for each program are provided in Supplementary Appendix 1 (Figures A1-A4).

All software programs were used in the *a priori* mode, with only patient demographic characteristics, dosing information, and clinical covariates entered into the models. No measured vancomycin concentrations were incorporated for Bayesian updating.

Predicted C_{min}, C_{max}, and AUC_{0-24} were obtained from each program. For NextDose, AUC_{0-24} was calculated by extracting all concentration-time points displayed on the predicted 24-hour concentration-time curve and applying the trapezoidal rule. Software-predicted values were then compared with the corresponding measured concentrations or reference AUC_{0-24}.

Evaluation of Predictive Performance

The *a priori* predictive performance of each software program was evaluated using prediction error (PE), absolute prediction error (APE), root mean square error (RMSE), F30, and classification agreement rate (CAR). PE and APE were calculated using Eqs. (5) and (6), respectively, with APE defined as the absolute value of PE. RMSE and F30 were calculated using Eqs. (7) and (8), respectively.

$$\mathrm{PE_i}(\%) := \frac{\mathrm{C_{Pred_i}} - \mathrm{C_{Obs_i}}}{\mathrm{C_{Obs_i}}} \times 100\% \quad \text{equation (5)}$$

$$\mathrm{APE_i}(\%) := |\mathrm{PE_i}| = \left| \frac{\mathrm{C_{Pred_i}} - \mathrm{C_{Obs_i}}}{\mathrm{C_{Obs_i}}} \right| \times 100\% \quad \text{equation (6)}$$

$$\mathrm{RMSE_i} = \sqrt{\frac{1}{\mathrm{N}} \sum_{\mathrm{i=1}}^{\mathrm{N}} \left(\mathrm{C_{Pred_i}} - \mathrm{C_{Obs_i}}\right)^2} \quad \text{equation (7)}$$

$$\mathrm{F_{30}} = \frac{\mathrm{N_{APE \le 30}}}{\mathrm{N}} \times 100\% \quad \text{equation (8)}$$

where C_{predi} and C_{obsi} represent the *i*th predicted value and the corresponding observed value, respectively. Based on previously established criteria for evaluating the predictive performance of vancomycin PopPK models, an APE ≤ 30% was considered an acceptable prediction, and F30 was defined as the proportion of predictions with an APE ≤ 30% [12, 13].

To assess agreement between predicted and observed values across exposure categories, each measure was classified into three categories. AUC_{0-24} values of < 400, 400–600, and > 600 mg·h/L were classified as low, target, and high exposure, respectively. C_{min} values of < 10, 10–20, and > 20 mg/L and C_{max} values of < 30, 30–40, and > 40 mg/L were classified as categories 1, 2, and 3, respectively. Classification agreement was defined as assignment of the predicted and observed values to the same category. The CAR was calculated as the proportion of paired measurements with concordant predicted and observed classifications. Cohen's κ coefficient was used to assess agreement beyond chance. The C_{min} and C_{max} categories were used solely to evaluate classification agreement and were not considered guideline-recommended independent targets for dose adjustment.

Visualization of Overall Predictive Performance

Raincloud plots were used to visualize the distributions of predicted values, observed values, and prediction errors by combining individual data points, box plots, and probability density distributions.

To provide a multidimensional assessment of *a priori* predictive performance, four standardized metrics - F30, prediction accuracy score (PA), CAR, and normalized RMSE score (NRMSE) - were calculated for C_{min}, C_{max}, and AUC_{0-24}, yielding 12 performance dimensions.

PA was calculated from the median APE using Eq. (9), whereas NRMSE was standardized by the RMSE relative to the mean of the corresponding observed values using Eq. (10).

$$\mathrm{PA}\ \%\ =100\%-\mathrm{median}\ \mathrm{APE}\ \ \ \

Table 1
Demographic Characteristics of Included Patients

Characteristics	Value
Male*, n (%)	108 (63.16)
Age*, median (Q1, Q3), years	57 (45.5, 68)
Weight*, median (Q1, Q3), kg	65 (60, 70)
Height*, median (Q1, Q3), cm	168 (162, 173)
BMI*, median (Q1, Q3), kg/m²	23.67 (21.72, 26.03)
Serum creatinine, median (Q1, Q3), μmol/L	51.65 (38.45, 68.3)
eGFR, median (Q1, Q3), mL/min/1.73 m²	109.1 (96.03, 125.75)
Serum albumin, median (Q1, Q3), g/L	32.8 (29.8, 35.9)
Total protein, median (Q1, Q3), g/L	60.6 (56.7, 64.3)
Urea, median (Q1, Q3), mmol/L	5.55 (4.32, 8.1)
AST, median (Q1, Q3), U/L	23.5 (16.25, 37.75)
ALT, median (Q1, Q3), U/L	33 (18, 56.75)
Renal function status, n (%)	
AKI	10 (5.1)
Normal renal function	128 (65.3)
Augmented renal clearance (ARC)	37 (18.88)
Renal insufficiency	21 (10.71)
Vancomycin trough concentration, n (%)	
< 10mg/L	83 (42.35)
10–20 mg/L	85 (43.37)
> 20mg/L	28 (14.28)
Vancomycin peak concentration, n (%)	
< 30 mg/L	103 (52.55)
30–40 mg/L	55 (28.06)
> 40mg/L	38 (19.39)
Vancomycin 24-h AUC	
< 400 mg•h/L	71 (36.22)
400–600 mg•h/L	90 (45.92)
> 600 mg•h/L	35 (17.86)
Daily vancomycin dose, n (%)	
Low dose (< 2 g/day)	29 (14.79)
Standard dose (2 g/day)	157 (80.1)
High dose (> 2 g/day)	10 (5.1)
Sampling department, n (%)	
Neurosurgery	115 (58.67)
ICU	51 (26.02)

Variables marked with "*" were summarized at the patient level (N = 171); all other variables were summarized based on 196 paired peak-trough concentration datasets

Characteristics	Value
Other departments	30 (15.31)
Concomitant medications, n (%)	
Diuretics	62 (31.63)
Vasoactive drugs	13 (6.63)
Variables marked with "*" were summarized at the patient level (N = 171); all other variables were summarized based on 196 paired peak-trough concentration datasets	

A Priori Predictive Performance for Vancomycin C_{min}

The distributions of observed and software-predicted C_{min} values are shown in Fig. 1a. The Wilcoxon signed-rank test revealed significant differences between predicted and observed C_{min} values for SmartDose, PharmVAN, and IOA (all $P < 0.05$), whereas no significant difference was observed for NextDose ($P = 0.217$). The corresponding PE distributions are presented in Fig. 1b. PE values were predominantly below zero for all four software programs, indicating an overall tendency toward underprediction, with more pronounced negative bias for SmartDose and PharmVAN and smaller overall bias for NextDose and IOA. The overall PE distributions differed significantly among the four software programs (Friedman test, $P < 0.001$). Detailed PE results and pairwise comparisons among software programs are presented in Supplementary Table S3.

The APE distributions are shown in Fig. 1c. The overall APE distributions differed significantly among the four software programs (Friedman test, $P < 0.001$). Bonferroni-adjusted pairwise comparisons showed that APE was significantly lower for IOA than for SmartDose and PharmVAN (both adjusted $P < 0.05$), whereas no significant difference was observed between IOA and NextDose (adjusted $P = 1.000$; Supplementary Table S4). RMSE values with 95% confidence intervals are presented in Fig. 1d. IOA had the lowest RMSE at 7.74 mg/L, compared with 8.98 mg/L for SmartDose, 9.65 mg/L for NextDose, and 11.10 mg/L for PharmVAN.

F30 results are shown in Fig. 1e. IOA achieved the highest F30 at 51.0%, followed by NextDose (45.4%), PharmVAN (37.2%), and SmartDose (30.6%). Pairwise McNemar tests showed that F30 was significantly higher for IOA than for SmartDose (adjusted P < 0.001) and PharmVAN (adjusted $P = 0.012$), whereas the difference between IOA and NextDose was not statistically significant ($P = 1.000$; Supplementary Table S6).

The distributions of concordant, underpredicted, and overpredicted classifications are presented in Fig. 1f. IOA showed the highest proportion of concordant classifications (67.9%), followed by NextDose and PharmVAN (both 59.2%) and SmartDose (53.6%). Cohen's κ was highest for IOA ($\kappa = 0.452$), followed by NextDose ($\kappa = 0.326$), PharmVAN ($\kappa = 0.301$), and SmartDose ($\kappa = 0.210$), with detailed agreement results presented in Supplementary Table S7.

A Priori Predictive Performance for Vancomycin C_{max}

The distributions of observed and software-predicted C_{max} values are shown in Fig. 2a. The Wilcoxon signed-rank test indicated significant differences between predicted and observed C_{max} values for all four software programs (all $P < 0.001$). The corresponding PE distributions are presented in Fig. 2b. Based on the direction of PE, PharmVAN and NextDose showed an overall tendency toward overprediction, whereas SmartDose and IOA tended toward underprediction. The overall PE distributions differed significantly among the four software programs (Friedman test, $P < 0.001$). Detailed PE statistics and Bonferroni-adjusted pairwise comparisons are provided in Supplementary Table S3.

The APE distributions are shown in Fig. 2c. The overall APE distributions differed significantly among the four software programs (Friedman test, $P < 0.001$). Bonferroni-adjusted pairwise comparisons showed that APE was significantly lower for IOA than for SmartDose, PharmVAN, and NextDose (all adjusted $P < 0.05$; Supplementary Table S4). RMSE values with 95% confidence intervals are presented in Fig. 2d. IOA had the lowest RMSE at 10.86 mg/L, followed by NextDose (14.48 mg/L), PharmVAN (15.16 mg/L), and SmartDose (16.83 mg/L).

F30 results are shown in Fig. 2e. IOA achieved the highest F30 at 77.6%, followed by PharmVAN (53.6%), NextDose (52.6%), and SmartDose (33.2%). Pairwise McNemar tests showed significant differences in F30 for all software comparisons (all adjusted $P < 0.05$) except PharmVAN versus NextDose (adjusted $P > 0.05$; Supplementary Table S6).

The distributions of concordant, underpredicted, and overpredicted classifications are shown in Fig. 2f. IOA showed the highest proportion of concordant classifications (57.1%), followed by SmartDose (52.0%), PharmVAN (49.5%), and NextDose (40.8%). Cohen's κ was highest for IOA ($\kappa = 0.245$), followed by PharmVAN ($\kappa = 0.239$), NextDose ($\kappa = 0.145$), and SmartDose ($\kappa = 0.080$), indicating generally limited agreement in C_{max} classification across the four software programs, with detailed agreement results presented in Supplementary Table S7.

A Priori Predictive Performance for Vancomycin AUC_{0-24}

The distributions of reference and software-predicted AUC_{0-24} values are shown in Fig. 3a. The Wilcoxon signed-rank test indicated significant differences between predicted and reference AUC_{0-24} values for all four software programs (all $P < 0.05$). The corresponding PE distributions are presented in Fig. 3b. Based on the direction of PE, SmartDose, PharmVAN, and IOA showed an overall tendency toward underprediction, whereas NextDose exhibited a slight tendency toward overprediction. The overall PE distributions differed significantly among the four software programs (Friedman test, $P < 0.001$). Detailed PE statistics and Bonferroni-adjusted pairwise comparisons are provided in Supplementary Table S3.

The APE distributions are shown in Fig. 3c. The overall APE distributions differed significantly among the four software programs (Friedman test, $P < 0.001$). Bonferroni-adjusted pairwise comparisons showed that APE was significantly lower for IOA than for SmartDose and PharmVAN (both adjusted P < 0.05), whereas no significant difference was observed between IOA and NextDose (adjusted $P = 0.679$; Supplementary Table S4). RMSE values with 95% confidence intervals are presented in Fig. 3d. IOA had the lowest RMSE at 202.68 mg·h/L, followed by NextDose (251.40 mg·h/L), SmartDose (275.15 mg·h/L), and PharmVAN (302.54 mg·h/L).

F30 results are shown in Fig. 3e. IOA achieved the highest F30 at 70.9%, followed by NextDose (59.7%), PharmVAN (47.5%), and SmartDose (33.7%). Pairwise McNemar tests showed that F30 was significantly higher for IOA than for SmartDose and PharmVAN (both adjusted $P < 0.001$), whereas the difference between IOA and NextDose was not statistically significant (adjusted $P = 0.110$; Supplementary Table S6).

The distributions of concordant, underpredicted, and overpredicted classifications are shown in Fig. 3f. Using 400–600 mg·h/L as the reference threshold, IOA achieved the highest proportion of concordant classifications (58.2%), followed by NextDose and PharmVAN (both 55.1%) and SmartDose (45.4%). Cohen's κ was highest for IOA ($\kappa = 0.339$), followed by NextDose ($\kappa = 0.323$), PharmVAN ($\kappa = 0.299$), and SmartDose ($\kappa = 0.144$), with detailed agreement results presented in Supplementary Table S7.

Overall Evaluation of A Priori Predictive Performance

For a multidimensional comparison of *a priori* predictive performance, four standardized metrics - F30, prediction accuracy score (PA), classification agreement rate (CAR), and normalized RMSE score (NRMSE) - were evaluated for C_{min}, C_{max}, and AUC_{0-24}, yielding a total of 12 dimensions for radar-chart analysis (Fig. 4).

The four software programs showed different performance profiles across the dimensions evaluated. IOA achieved relatively high standardized scores across most dimensions and had the largest relative radar-chart coverage area (58.9%), followed by NextDose (41.1%), PharmVAN (35.1%), and SmartDose (28.4%). Overall, the multidimensional descriptive assessment showed relatively favorable *a priori* predictive performance for IOA, consistent with the general trends observed in the individual evaluations of C_{min}, C_{max}, and AUC_{0-24}.

Factors Associated With IOA AUC_{0-24} Prediction Error and Subgroup Analyses

Given the relatively favorable overall predictive performance of IOA in the multidimensional evaluation, a multivariable generalized linear model was used to investigate clinical factors associated with its AUC_{0-24} prediction error (Fig. 5a). After adjustment for other covariates, AKI ($Exp(B) = 2.103$, $P = 0.024$) and a low daily vancomycin dose (< 2 g/day; $Exp(B) = 1.498$, $P = 0.030$) were independently associated with greater AUC_{0-24} APE.

Subgroup analyses showed substantially poorer AUC_{0-24} predictive performance of IOA in patients with AKI (Fig. 5b-d). Compared with the non-AKI group, the AKI group had a significantly lower F30 (20.0% vs. 73.7%, $P < 0.01$) and a significantly higher RMSE [507.16 mg·h/L (95% CI, 353.14-661.11) vs. 171.64 mg·h/L (95% CI, 135.68–208.60), $P < 0.001$]. However, no significant difference in CAR was observed between the AKI and non-AKI groups (60.0% vs. 58.1%, $P = 1.000$), and the distributions across the three classification outcomes (concordant, underpredicted, and overpredicted) also did not differ significantly ($P = 0.517$).

When stratified by daily vancomycin dose, RMSE values were 216.44, 205.42, and 82.26 mg·h/L in the low-dose (< 2 g/day), standard-dose (2 g/day), and high-dose (> 2 g/day) groups, respectively, with no statistically significant difference among the groups ($P = 0.200$; Fig. 5e-g). CAR likewise did not differ significantly among the three dose groups ($P = 0.261$). However, the distributions across the three classification outcomes differed significantly ($P = 0.021$). Underprediction was more frequent in the low-dose subgroup, whereas no underprediction was observed in the high-dose group, in which overprediction predominated.

Discussion

Using real-world TDM data, this study compared the *a priori* predictive performance of four Bayesian dosing software programs for vancomycin C_{min}, C_{max}, and AUC_{0-24} before individual concentration measurements were incorporated. Performance varied substantially across programs including prediction bias, accuracy, prediction acceptability, and classification agreement. Overall, IOA showed relatively favorable predictive performance, with lower prediction errors and higher prediction acceptability across most evaluated metrics, as well as the largest relative radar-chart coverage area (58.9%). However, the advantages of IOA were not statistically significant for all metrics or pairwise comparisons, and prediction precision and classification agreement remained limited for some pharmacokinetic endpoints. Further analyses showed that AKI and a low daily vancomycin dose were independently associated with greater AUC_{0-24} APE, suggesting that *a priori* predictive performance may

deteriorate in specific clinical subgroups. Collectively, these findings indicate that currently available Bayesian dosing software can support initial individualized vancomycin dosing; however, their predictive performance before individual TDM data become available remains limited by model characteristics and patient-level pharmacokinetic heterogeneity.

At the *a priori* prediction stage, when measured drug concentrations are unavailable for Bayesian updating, predictions rely primarily on the underlying PopPK prior model and patient-specific covariates that account for inter- and intra-individual pharmacokinetic variability [14]. Previous studies have shown substantial variation in the external *a priori* predictive performance of vancomycin PopPK models [15, 16]. For instance, Broeker et al [17]. evaluated 31 published models and reported wide variation in both relative bias and relative RMSE, highlighting the importance of model-population compatibility. This variation does not necessarily mean that one modeling strategy is intrinsically better; rather, it largely reflects differences in model structure, covariate selection [18], and the compatibility between the model-development population and the target population [19–21]. Therefore, the relatively favorable performance of IOA observed in this study may primarily reflect compatibility between its underlying model characteristics and the Chinese adult patients in this cohort. However, because detailed information regarding IOA's model parameterization and development datasets is not publicly available, the reasons for its relatively favorable performance remain unclear [22]. The limitations of relying on population priors and static clinical covariates become particularly evident in patients with highly variable pharmacokinetics, such as those with unstable renal function [23].

This limitation may be particularly relevant in patients with dynamically changing renal function. In the present study, IOA showed a substantially higher AUC_{0-24} RMSE and a significantly lower F30 in patients with AKI, and AKI remained independently associated with greater absolute prediction error in the multivariable analysis. These findings are pharmacokinetically plausible because vancomycin undergoes predominantly renal elimination, whereas renal function in patients with AKI changes over time and is not at steady state. Consequently, eGFR or creatinine clearance calculated from a single serum creatinine measurement may not adequately capture rapid changes in actual renal function. Previous studies have shown that incorporating time-varying renal function covariates can improve vancomycin prediction bias and precision in patients with declining renal function [24]. Moreover, incorporating measured concentrations through Bayesian estimation may further improve prediction in patients with unstable renal function [25]. An evaluation of Bayesian dosing software also demonstrated improved vancomycin prediction when serum creatinine measured around the sampling time was used instead of the initial serum creatinine value [26]. Thus, the greater *a priori* prediction error observed in patients with AKI may primarily reflect the limitations of static baseline covariates in characterizing dynamic renal function and time-varying PK processes, rather than a limitation of a specific software program or PopPK model alone.

In addition to renal function, the dose-stratified analyses suggested that *a priori* prediction errors may not be entirely random. Although the overall difference in AUC_{0-24} RMSE among daily-dose groups was not statistically significant ($P = 0.200$), the direction of prediction bias varied across dose strata. Underprediction was more frequent in the low-dose group, whereas no underprediction was observed in the high-dose group, and the distribution of the three classification outcomes differed significantly among dose groups ($P = 0.021$). A low daily vancomycin dose (< 2 g/day) also remained independently associated with greater absolute prediction error in the multivariable analysis. These findings do not establish that dose itself causes prediction bias, given that clinical dose selection depends on patient-specific factors, including renal function, age, body weight, disease severity, and other patient characteristics. PopPK covariate models are designed to explain predictable variability associated with patient characteristics; however, substantial unexplained interindividual and residual variability may remain [27, 28]. Our findings therefore suggest that residual errors in *a priori* prediction may contain structured information related to clinical and treatment characteristics. If replicated in larger and external populations, such patterns may help improve initial exposure prediction.

Although the primary objective of this study was to characterize the *a priori* predictive performance of vancomycin dosing software before the availability of individual concentration measurements, we also conducted an exploratory analysis in 17 patients with available TDM data to evaluate *a posteriori* predictive performance after Bayesian updating. Bayesian updating integrates individual concentration measurements with population pharmacokinetic models to refine patient-specific parameter estimates and is an essential component of model-informed precision dosing workflows [29, 30]. Detailed results of the *a posteriori* evaluation are provided in Supplementary Figures S2-S7 and Supplementary Tables S2-S6. Overall, incorporating individual concentration data improved predictive performance across the evaluated software programs, although the extent of improvement varied by prediction targets and software platforms. For AUC_{0-24} prediction, the primary therapeutic monitoring target for vancomycin, F30 increased across all software programs after Bayesian updating, with IOA achieving the highest post-update acceptability. Among the evaluated programs, SmartDose demonstrated the greatest relative improvement after Bayesian updating. These findings suggest that Bayesian updating can enhance individualized exposure prediction, while differences among software platforms may persist throughout the MIPD workflow. Because this analysis included only 17 patients, these findings should be interpreted as exploratory and require validation in larger prospective cohorts.

In current practice, initial *a priori* predictions should be interpreted cautiously in high-risk patients (e.g., those with AKI), and TDM data should be incorporated as early as feasible for Bayesian updating [31, 32]. This study has several limitations. First, the single-center retrospective design cannot rule out unmeasured confounding, and multicenter validation is needed. Second, excluding patients with vancomycin concentrations outside the institutional analytical range (3-100 mg/L) may introduce selection bias. Third, the small sample size (n = 17) in the *a posteriori* subgroup means that those findings remain exploratory. Finally, incomplete access to the underlying PopPK models limited our ability to explain the observed performance differences.

In conclusion, although IOA showed relatively favorable *a priori* predictive performance within this cohort, clinically relevant errors remained, particularly in patients with acute kidney injury. These findings support cautious interpretation of *a priori* predictions and early use of TDM data for Bayesian updating. Recent studies demonstrate that hybrid models integrating machine learning with population pharmacokinetics can substantially improve vancomycin prediction accuracy before TDM data become available [33]. Therefore, our future work will focus on developing and validating such data-driven hybrid models to further optimize individualized precision dosing.

Declarations

Clinical trial number

not applicable.

Ethical approval:

Ethical approval was obtained from the Ethics Committee of Jinling Clinical Medical College of Nanjing Medical University (approval no. 2026DZKY-148-01). The research was performed in accordance with the ethical standards of the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Consent to participate:

Not applicable. As this was a retrospective study using routinely collected clinical data obtained from the hospital electronic system, informed consent was not obtained. Patient confidentiality was strictly maintained throughout the data collection and analysis process.

Consent for Publication:

Not applicable. This manuscript does not involve any identifiable clinical photographs of patients or any identifiable patient data.

Funding:

No funding was received for conducting this study or assisting with the preparation of this manuscript.

Author Contribution

Conceptualization, methodology, and study design were contributed by QW, GHZ, and XJL. Data collection and curation were performed by QW, JXW, YML. Data validation was conducted by XYB,YJZ. Formal analysis, data cleaning, and visualization were carried out by QW. The first draft of the manuscript was written by QW. Methodological guidance was provided by JXW. Project administration, supervision, and critical revision of the manuscript were contributed by HPW, GHZ, and XJL. All authors read and approved the final manuscript.

Data Availability

Aggregate data supporting the conclusions of this study are included within the article and its supplementary materials. Individual-level patient data are not publicly available due to ethical and privacy considerations but may be provided upon reasonable request and with appropriate approvals.

Code availability:

Not applicable.

References

1. Sisodia K, Panda A, Haridas H, Rynga D, Kumar D, Gaind R (2026) Vancomycin MIC creep and the emergence of oxacillin-susceptible MRSA: A three-year surveillance study from a tertiary care centre. New Microbes New Infect 71:101742. https://doi.org/10.1016/j.nmni.2026.101742
2. Rybak MJ (2006) The pharmacokinetic and pharmacodynamic properties of vancomycin. Clin Infect Dis 42(Suppl 1):S35–39. https://doi.org/10.1086/491712

3. Rybak MJ, Le J, Lodise TP, Levine DP, Bradley JS, Liu C et al (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 infectious diseases society, and the society of infectious diseases pharmacists. Am J Health-System Pharm 77:835–864. https://doi.org/10.1093/ajhp/zxaa036
4. Bradley N, Ng K (2023) Evaluation of real-world vancomycin dosing and attainment of therapeutic drug monitoring targets. Pharm (Basel) 11:95. https://doi.org/10.3390/pharmacy11030095
5. Dogan CZ, Kara E, Pinar A, Demirkan K, Metan G (2025) Therapeutic drug monitoring in patients treated with vancomycin: A single center, prospective, observational, real-world study. Eur J Clin Microbiol Infect Dis 44:2159–2165. https://doi.org/10.1007/s10096-025-05182-w
6. Oommen T, Thommandram A, Palanica A, Fossat Y (2022) A free open-source bayesian vancomycin dosing app for adults: Design and evaluation study. JMIR Form Res 6:e30577. https://doi.org/10.2196/30577
7. Blouin M, Métras MÉ, Gaudreault C, Dubé MH, Boulanger MC, Cloutier K et al (2024) External evaluation of neonatal vancomycin population pharmacokinetic models: Moving from first-order equations to bayesian-guided therapeutic monitoring. Pharmacotherapy 44:907–919. https://doi.org/10.1002/phar.4623
8. Bai G, Qi H, Huang Y, Zhang J, Zhao H, Wen R et al (2025) Predictive performance of bayesian dosing software for vancomycin in intensive care unit patients. Ther Drug Monit 47:594–602. https://doi.org/10.1097/FTD.0000000000001310
9. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, Feldman HI et al (2009) A new equation to estimate glomerular filtration rate. Ann Intern Med 150:604–612. https://doi.org/10.7326/0003-4819-150-9-200905050-00006
10. KDIGO 2026 Clinical Practice Guideline for Acute Kidney Injury (AKI) and Acute Kidney Disease (AKD) Public Review Draft [Internet]. Kidney Disease: Improving Global Outcomes (KDIGO); 2026 Mar. https://kdigo.org/wp-content/uploads/2026/03/KDIGO-2026-AKI-AKD-Guideline-Public-Review-Draft-March-2026.pdf. Accessed 28 June 2026
11. Pai MP, Neely M, Rodvold KA, Lodise TP (2014) Innovative approaches to optimizing the delivery of vancomycin in individual patients. Adv Drug Deliv Rev 77:50–57. https://doi.org/10.1016/j.addr.2014.05.016
12. Lv C, Lu J, Jing L, Liu T-T, Chen M, Zhang R et al (2021) Systematic external evaluation of reported population pharmacokinetic models of vancomycin in Chinese children and adolescents. J Clin Pharm Ther 46:820–831. https://doi.org/10.1111/jcpt.13363
13. Duong A, El Gamal A, Bilodeau V, Huot J, Delorme C, Poudrette J et al (2024) Vancomycin: An analysis and evaluation of eight population pharmacokinetic models for clinical application in general adult population. Pharmacotherapy 44:425–434. https://doi.org/10.1002/phar.2941
14. Marsot A, Boulamery A, Bruguerolle B, Simon N, Vancomycin (2012) A review of population pharmacokinetic analyses. Clin Pharmacokinet 51:1–13. https://doi.org/10.2165/11596390-000000000-00000
15. Wicha SG, Märtson A-G, Nielsen EI, Koch BCP, Friberg LE, Alffenaar J-W et al (2021) From therapeutic drug monitoring to model-informed precision dosing for antibiotics. Clin Pharmacol Ther 109:928–941. https://doi.org/10.1002/cpt.2202
16. Heus A, Uster DW, Grootaert V, Vermeulen N, Somers A, In't Veld DH et al (2022) Model-informed precision dosing of vancomycin via continuous infusion: A clinical fit-for-purpose evaluation of published PK models. Int J Antimicrob Agents 59:106579. https://doi.org/10.1016/j.ijantimicag.2022.106579
17. Broeker A, Nardecchia M, Klinker KP, Derendorf H, Day RO, Marriott DJ et al (2019) Towards precision dosing of vancomycin: A systematic evaluation of pharmacometric models for bayesian forecasting. Clin Microbiol Infect 25:1286. .e1-1286.e7 https://doi.org/10.1016/j.cmi.2019.02.029
18. Liu Y-X, Wen H, Niu W-J, Li J-J, Li Z-L, Jiao Z (2021) External evaluation of vancomycin population pharmacokinetic models at two clinical centers. Front Pharmacol 12:623907. https://doi.org/10.3389/fphar.2021.623907
19. Holford N, O'Hanlon CJ, Allegaert K, Anderson B, Falcão A, Simon N et al (2024) A physiological approach to renal clearance: from premature neonates to adults. Br J Clin Pharmacol 90:1066–1080. https://doi.org/10.1111/bcp.15978
20. He X-R, Liu Z-H, Ji S-M, Liu T-T, Li L, Zhou T-Y et al (2014) [population pharmacokinetics of vancomycin and prediction of pharmacodynamics in the Chinese people]. Yao Xue Xue Bao 49:1528–1535
21. Yasuhara M, Iga T, Zenda H, Okumura K, Oguma T, Yano Y et al (1998) Population pharmacokinetics of vancomycin in japanese pediatric patients. Ther Drug Monit 20:612–618. https://doi.org/10.1097/00007691-199812000-00005
22. Individualized medication guidance platform for anti infective Drugs [Internet]. [cited 2026 Sept 3]. https://huashan.99rongle.com/huashan-jiliang/hmschushi. Accessed 3 Sept 2026
23. Guo T, van Hest RM, Roggeveen LF, Fleuren LM, Thoral PJ, Bosman RJ et al (2019) External evaluation of population pharmacokinetic models of vancomycin in large cohorts of intensive care unit patients. Antimicrob Agents Chemother 63:e02543–e02518. https://doi.org/10.1128/AAC.02543-18
24. Yamamoto T, Terakawa H, Hisaka A, Suzuki H (2012) Bayesian estimation of pharmacokinetic parameters of vancomycin in patients with decreasing renal function. J Pharm Sci 101:2968–2975. https://doi.org/10.1002/jps.23183
25. Jaisue S, Pongsakul C, D'Argenio DZ, Sermsappasuk P (2020) Population pharmacokinetic modeling of vancomycin in thai patients with heterogeneous and unstable renal function. Ther Drug Monit 42:856–865. https://doi.org/10.1097/FTD.0000000000000801

26. Yu Y-X, Lu J, Lu H, Li L, Li J-J, Shi L et al (2022) Predictive performance of reported vancomycin population pharmacokinetic model in patients with different renal function status, especially those with augmented renal clearance. Eur J Hosp Pharm 29:e6–14. https://doi.org/10.1136/ejhpharm-2020-002477
27. Mould DR, Upton RN (2013) Basic concepts in population modeling, simulation, and model-based drug development-part 2: Introduction to pharmacokinetic modeling methods. CPT Pharmacometrics Syst Pharmacol 2:e38. https://doi.org/10.1038/psp.2013.14
28. Sanghavi K, Ribbing J, Rogers JA, Ahmed MA, Karlsson MO, Holford N et al (2024) Covariate modeling in pharmacometrics: General points for consideration. CPT Pharmacometrics Syst Pharmacol 13:710–728. https://doi.org/10.1002/psp4.13115
29. Guo T, van Hest RM, Zwep LB, Roggeveen LF, Fleuren LM, Bosman RJ et al (2020) Optimizing predictive performance of bayesian forecasting for vancomycin concentration in intensive care patients. Pharm Res 37:171. https://doi.org/10.1007/s11095-020-02908-7
30. Chen A, Gupta A, Do DH, Nazer LH (2022) Bayesian method application: Integrating mathematical modeling into clinical pharmacy through vancomycin therapeutic monitoring. Pharmacol Res Perspect 10:e01026. https://doi.org/10.1002/prp2.1026
31. Fochesato A, Brooks L, Bazgir O, Pierrillas PB, Jamois C, Lu J et al (2026) Building hybrid pharmacometric-machine learning models in oncology drug development: Current state and recommendations. CPT Pharmacometrics Syst Pharmacol 15:e70113. https://doi.org/10.1002/psp4.70113
32. Matsumoto K, Oda K, Shoji K, Hanai Y, Takahashi Y, Fujii S et al (2022) Clinical practice guidelines for therapeutic drug monitoring of vancomycin in the framework of model-informed precision dosing: A consensus review by the japanese society of chemotherapy and the japanese society of therapeutic drug monitoring. Pharmaceutics 14:489. https://doi.org/10.3390/pharmaceutics14030489
33. Chen K, Wang C, Wei Y, Ma S, Huang W, Dong Y et al (2025) Machine learning and population pharmacokinetics: A hybrid approach for optimizing vancomycin therapy in sepsis patients. Microbiol Spectr 13:e0049925. https://doi.org/10.1128/spectrum.00499-25

Figures

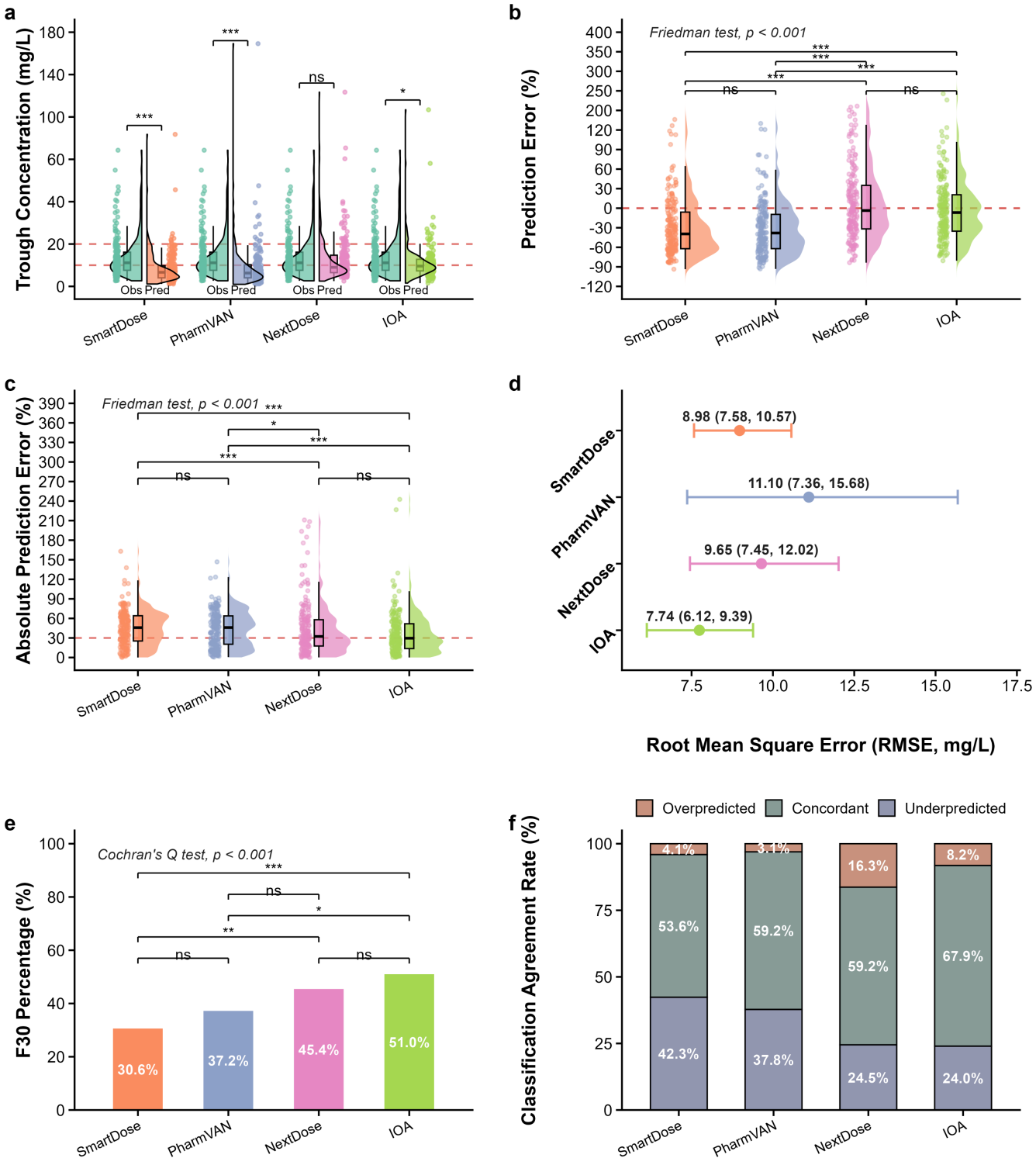


Figure 1

A priori predictive performance of four model-informed precision dosing software programs for vancomycin trough concentration (Cmin). a Distribution of observed and software-predicted Cmin values. The red dashed lines indicate the prespecified classification thresholds of 10 and 20 mg/L. b Distribution of prediction error (PE); the horizontal dashed line indicates PE = 0%. c Distribution of absolute prediction error (APE); the red dashed line indicates the 30% threshold used to define an acceptable prediction. d Root mean square error (RMSE) for each software program; points represent RMSE estimates, and horizontal error bars indicate 95% confidence intervals. e F30, defined as the proportion of predictions with an APE ≤ 30%. f Distribution of predicted classification outcomes according to prespecified Cmin categories (< 10, 10-20, and > 20 mg/L). In panel a, "Obs" and "Pred" denote the observed and software-predicted values, respectively. Statistical significance is denoted as $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$; ns, not significant

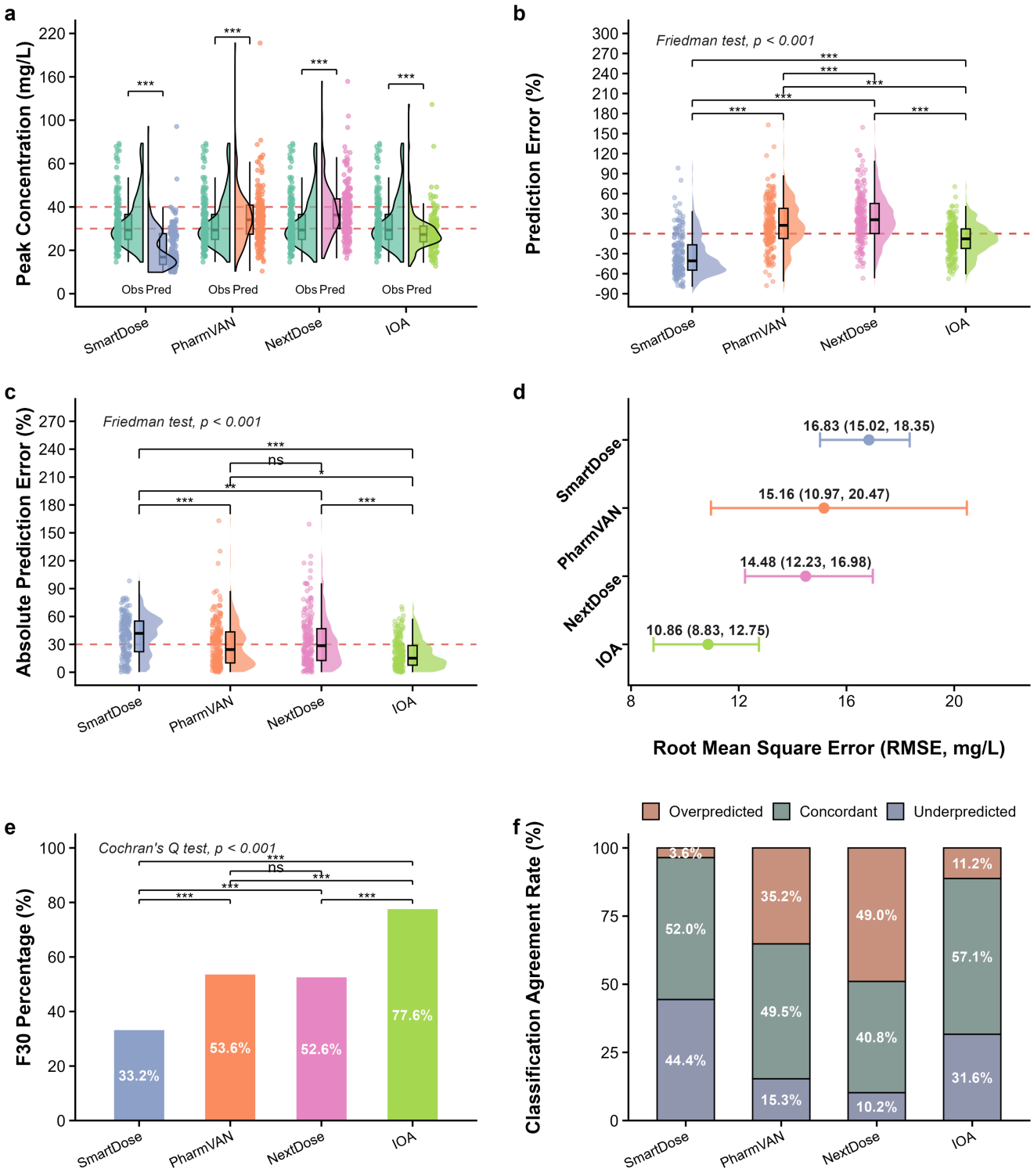


Figure 2

***A priori* predictive performance of four model-informed precision dosing software programs for vancomycin peak concentration (C_{max}).** a Distribution of observed and software-predicted C_{max} values. The red dashed lines indicate the prespecified classification thresholds of 30 and 40 mg/L. b Distribution of prediction error (PE); the horizontal dashed line indicates PE = 0%. c Distribution of absolute prediction error (APE); the red dashed line indicates the 30% threshold used to define an acceptable prediction. d Root mean square error (RMSE) for each software program; points represent RMSE estimates, and horizontal error bars indicate 95% confidence intervals. e F30, defined as the proportion of predictions with an APE ≤ 30%. f Distribution of classification outcomes according to the prespecified C_{max} categories (< 30, 30-40, and > 40 mg/L). In panel a, "Obs" and "Pred" denote the observed and software-predicted values, respectively. Statistical significance is denoted as $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$; ns, not significant

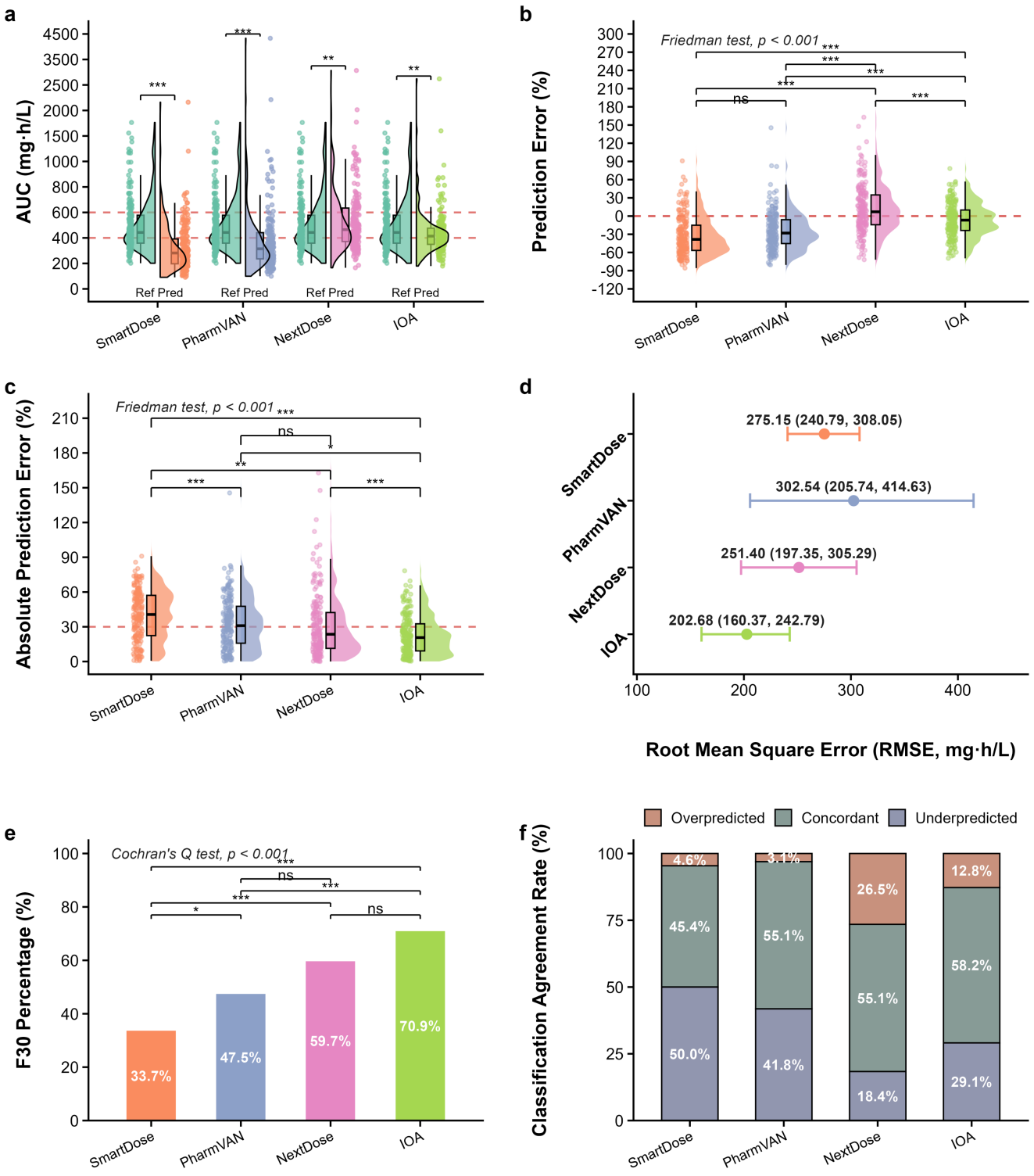


Figure 3

***A priori* predictive performance of four model-informed precision dosing software programs for vancomycin $AUC_{0\text{-}24}$.** a Distribution of reference and software-predicted $AUC_{0\text{-}24}$ values. The red dashed lines indicate the exposure thresholds of 400 and 600 mg·h/L. b Distribution of prediction error (PE); the horizontal dashed line indicates PE = 0%. c Distribution of absolute prediction error (APE); the red dashed line indicates the 30% threshold used to define an acceptable prediction. d Root mean square error (RMSE) for each software program; points represent RMSE estimates, and horizontal error bars indicate 95% confidence intervals. e F30, defined as the proportion of predictions with an APE ≤ 30%. f Distribution of exposure-classification outcomes based on $AUC_{0\text{-}24}$ categories of < 400, 400-600, and > 600 mg·h/L.In panel a, "Ref" and "Pred" denote the reference and software-predicted values, respectively. Statistical significance is denoted as $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$; ns, not significant

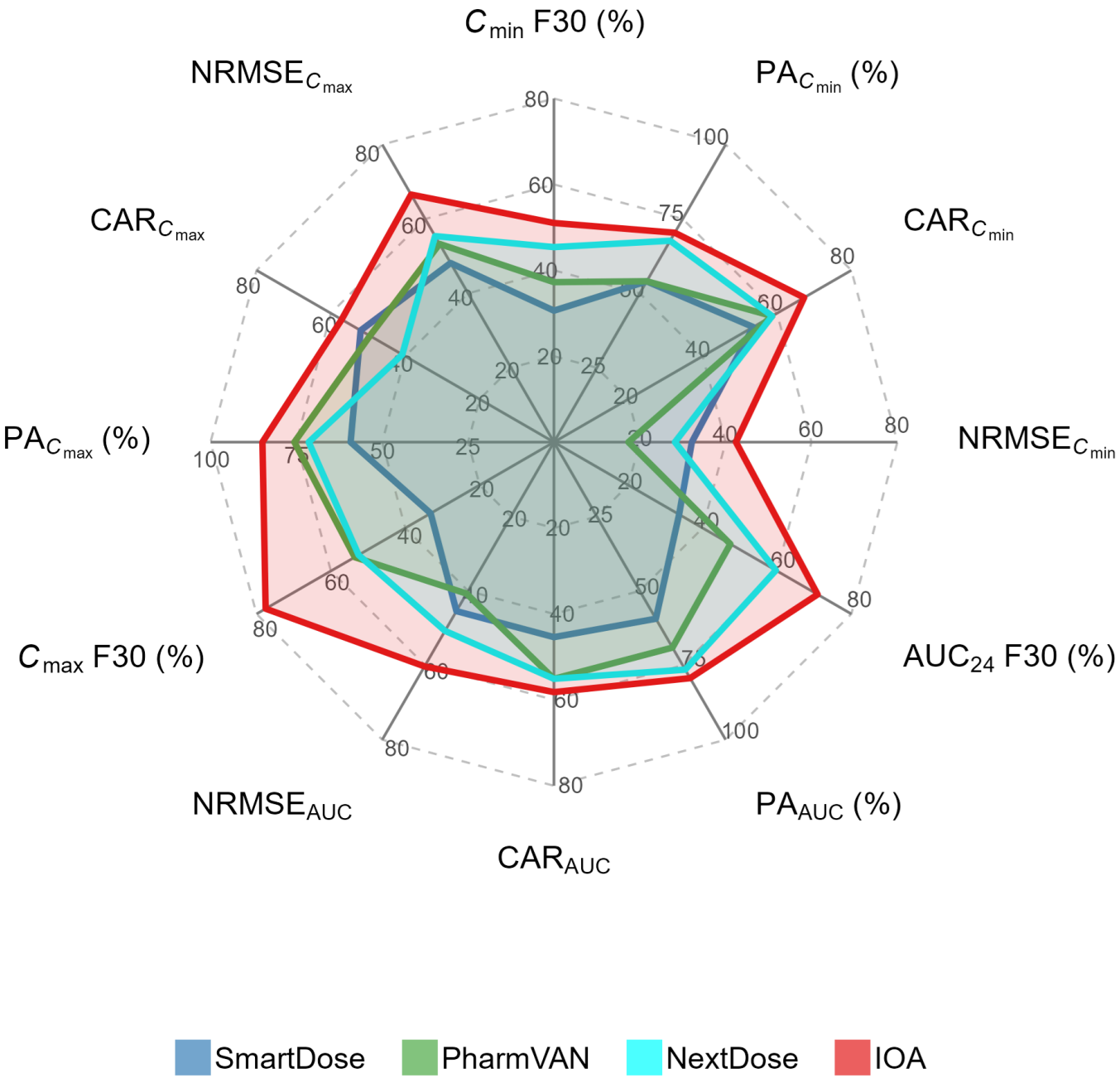


Figure 4

Multidimensional comparison of the *a priori* predictive performance of the four model-informed precision dosing software programs. Radar-chart analysis was based on 12 standardized performance dimensions derived from four metrics - F30, prediction accuracy score (PA), classification agreement rate (CAR), and normalized root mean square error score (NRMSE) - for C_{min}, C_{max}, and $AUC_{0\text{-}24}$. All metrics were standardized on a 0-100 scale, with higher scores indicating better performance. PA was derived from the median absolute prediction error (APE), whereas NRMSE was derived from RMSE normalized to the corresponding mean observed or reference value. The relative radar-chart coverage area was used solely as a descriptive multidimensional summary and was not subjected to inferential statistical analysis. The relative radar-chart coverage area was largest for IOA (58.9%), followed by NextDose (41.1%), PharmVAN (35.1%), SmartDose (28.4%)

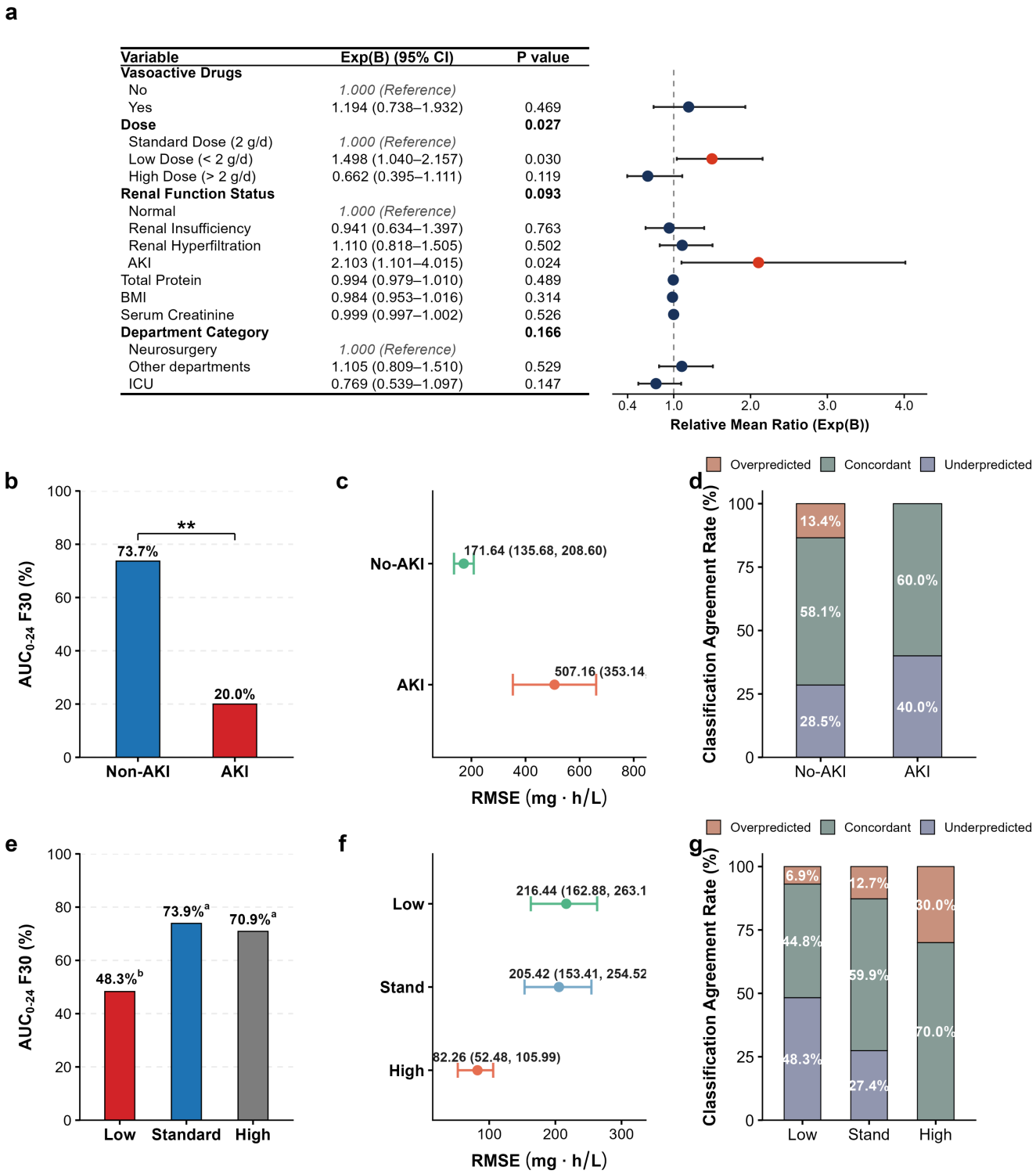


Figure 5

Factors associated with IOA AUC_{0-24} prediction error and subgroup analyses according to acute kidney injury and daily vancomycin dose. a Multivariable generalized linear model of factors associated with AUC_{0-24} absolute prediction error (APE), expressed as Relative Mean Ratio [Exp(*B*)] with 95% CIs (red points indicate $P < 0.05$). **b-d** Predictive performance stratified by AKI status, showing **b** F30 (APE < 30%), **c** RMSE, and **d** classification agreement rate distribution. **e-g** Predictive performance stratified by daily vancomycin dose (< 2, 2, and > 2 g/day), showing **e** F30 (different superscript letters indicate significant pairwise differences), **f** RMSE, and **g** classification agreement rate distribution. Exposure classification categories: < 400, 400-600, and > 600 mg·h/L. $P < 0.01$ for F30 between AKI and NO-AKI groups

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ESM1TablesS1S9.docx
- ESM2FiguresS1S7.docx
- ESM3FigureA1A4.docx