Intensive Care Med
https://doi.org/10.1007/s00134-023-07154-0

ORIGINAL

Towards model-informed precision dosing of piperacillin: multicenter systematic external evaluation of pharmacokinetic models in critically ill adults with a focus on Bayesian forecasting

Sebastian Greppmair[1], Alexander Brinkmann[2], Anka Roehr[3], Otto Frey[3], Stefan Hagel[4], Christoph Dorn[5], Amélie Marsot[6], Ibrahim El-Haffaf[6], Michael Zoller[1], Thomas Saller[1], Johannes Zander[7], Lea Marie Schatz[8], Christina Scharf[1], Josef Briegel[1], Iris K. Minichmayr[9], Sebastian G. Wicha[10] and Uwe Liebchen[1*]

© 2023 The Author(s)

Abstract

Purpose: Inadequate piperacillin (PIP) exposure in intensive care unit (ICU) patients threatens therapeutic success. Model-informed precision dosing (MIPD) might be promising to individualize dosing; however, the transferability of published models to external populations is uncertain. This study aimed to externally evaluate the available PIP population pharmacokinetic (PopPK) models.

Methods: A multicenter dataset of 561 ICU patients (11 centers/3654 concentrations) was used for the evaluation of 24 identified models. Model performance was investigated for a priori (A) predictions, i.e., considering dosing records and patient characteristics only, and for Bayesian forecasting, i.e., additionally including the first (B1) or first and second (B2) therapeutic drug monitoring (TDM) samples per patient. Median relative prediction error (MPE) [%] and median absolute relative prediction error (MAPE) [%] were calculated to quantify accuracy and precision.

Results: The evaluation revealed a large inter-model variability (A: MPE − 135.6–78.3% and MAPE 35.7–135.6%). Integration of TDM data improved all model predictions (B1/B2 relative improvement vs. A: $|MPE|_{median\_all\_models}$ 45.1/67.5%; $MAPE_{median\_all\_models}$ 29/39%). The model by Kim et al. was identified to be most appropriate for the total dataset (A/B1/B2: MPE − 9.8/− 5.9/− 0.9%; MAPE 37/27.3/23.7%), Udy et al. performed best in patients receiving intermittent infusion, and Klastrup et al. best predicted patients receiving continuous infusion. Additional evaluations stratified by sex and renal replacement therapy revealed further promising models.

Conclusion: The predictive performance of published PIP models in ICU patients varied considerably, highlighting the relevance of appropriate model selection for MIPD. Our differentiated external evaluation identified specific models suitable for clinical use, especially in combination with TDM.

Keywords: Intensive care medicine, Model-informed precision dosing, Piperacillin, Pharmacokinetics/pharmacodynamics, Therapeutic drug monitoring

*Correspondence: uwe.liebchen@med.uni-muenchen.de
[1] Department of Anaesthesiology, University Hospital, LMU Munich, Marchioninistrasse 15, 81377 Munich, Germany
Full author information is available at the end of the article

Springer

Introduction

Infections increase the length of stay and mortality of patients in the intensive care unit (ICU) [1, 2]. Timely administration of antibiotics and dosing that results in effective exposure are of paramount importance to treatment outcomes [3–5]. However, critically ill patients commonly exhibit pathophysiological changes, altering the volume of distribution (V_d) and the elimination of antibiotics, making dosing and pharmacodynamic (PD) target attainment challenging [6–8]. While excessive drug/multidrug concentrations might cause life-threating toxicity, subtherapeutic concentrations increase the risk of treatment failure and emerging antibacterial resistance [9, 10]. To avoid these scenarios, individualized dosing approaches are urgently needed [11].

Piperacillin (PIP) is the most frequently prescribed β-lactam antibiotic in German ICUs [12]. Its time-dependent antibacterial activity is characterized by the time (T) during which free (*f*) concentrations exceed the minimum inhibitory concentration (MIC) of the pathogen ($fT_{>MIC}$) [13]. Standard doses of PIP have been reported to result in poor target attainment [14, 15]. Dose adjustments using traditional therapeutic drug monitoring (TDM) have therefore been investigated. Hagel et al. showed increased, but still poor attainment of target concentrations during continuous infusion (CI) in 249 critically ill patients (37.3 vs. 14.6%, odds ratio (OR) 4.5, 95% confidence interval 2.9–6.9, $p < 0.001$) [16]. Along with others, the authors subsequently proposed to investigate the additional benefit of using dosing software [16, 17]. Model-informed precision dosing (MIPD) is a promising predictive technology that supports dosing and combination with TDM results. It is available through dosing software and, thus, convenient to use [18–20]. Prior to widespread bedside use, a thorough evaluation of the population pharmacokinetic (PopPK) model(s) underlying the dosing software is strongly recommended by the United States (US) Food and Drug Administration [21], and by an expert panel defining research priorities towards antibiotic precision dosing [11]. Similarly, Cotta et al. called for "externally validated and clinically appropriate PopPK models" in their presentation of "ideal characteristics of MIPD software in the ICU" [22].

Numerous PopPK models have been developed for PIP, most of which are based on small monocentric studies [23] and may not be generalizable to other populations. A recent evaluation of six PIP models in 30 ICU patients receiving CI demonstrated large inter-model variability regarding predictability [24]. The transferability of these results to other populations is uncertain due to the limited number of patients and the monocentric setting. Furthermore, a clinically oriented model assessment in conjunction with TDM (Bayesian forecasting) was lacking [25]. The aim of the present study was to evaluate the predictive performance of available PIP PopPK models with and without TDM using an external multicenter dataset to facilitate model selection for MIPD in critically ill patients.

Take-home message

This multicenter external evaluation in critically ill patients provides a selection of different pharmacokinetic models for piperacillin as promising candidates for successful clinical use following implementation in an open-access dosing software. The selected models should be used in combination with timely therapeutic drug monitoring (Bayesian forecasting) to achieve most reliable and precise optimization of individual antibiotic therapy.

Methods

Evaluation dataset

Clinical data of 561 ICU patients treated with PIP (3654 samples) were available from four previous studies including eleven different German centers [16, 26–28]. Details are described in electronic supplementary material (ESM_Main, page 2) and summarized in Table 1.

Evaluation of population pharmacokinetic models for piperacillin

The systematic literature review in PubMed, including search terms and a detailed flowchart of PopPK model screening, as well as the software used for (i) model reconstruction, (ii) prediction of concentrations, and (iii) output processing, is presented in ESM (ESM_Main, page 3–5, Fig. S1). An overview of the twenty-four identified models and underlying studies is provided (ESM_Main Tables S1–2). The report was guided by the TRIPOD checklist [31]. If models were based on free PIP concentrations ($PIP_{unbound}$), 70% of the total PIP concentrations (PIP_{total}) available in the evaluation dataset were assumed for model assessment [24, 32]. Three prediction scenarios were examined for each PK model, two of which investigated a Bayesian approach, i.e., the combination of PopPK model and TDM results based on the Bayes' theorem:

(1) A priori (A): prediction of all PIP concentrations based on dosing history and patient covariates only ($n_{predicted\_samples\_A} = 3654$)

Table 1 Demographic and clinical characteristics of patients in the evaluation dataset

	Total ($n_{patients}$ = 561) ($n_{samples}$ = 3654) Multicenter	Study 1 ($n_{patients}$ = 207) ($n_{samples}$ = 1064) 9 centers [16]	Study 2 ($n_{patients}$ = 282) ($n_{samples}$ = 731) Heidenheim [26]	Study 3 ($n_{patients}$ = 12) ($n_{samples}$ = 24) Regensburg [27]	Study 4 ($n_{patients}$ = 60) ($n_{samples}$ = 1835) Munich [28]
Patient characteristics[a]					
Sex					
Male	370 (66)	145 (70)	176 (62)	6 (50)	43 (72)
Female	191 (34)	62 (30)	106 (38)	6 (50)	17 (28)
Age [years]	67 [16–94]	66 [19–90]	70 [16–94]	52 [19–81]	64 [23–82]
Height [cm]	173 [100–205]	175 [153–195]	170 [100–205]	170 [158–190]	173 [150–198]
Weight [kg][b]	80 [25–203]	80 [43–193]	77 [25–203]	80 [48–105]	80 [50–150]
ABW [kg][b]	72 [26–124]	74 [47–114]	70 [26–124]	69 [53–90]	71 [48–112]
BMI[b]	26.1 [11.1–70.2]	26 [14.9–66.8]	26.2 [11.1–70.2]	25.8 [16.6–36.7]	26.6 [17.3–40.7]
eGFR [mL/min][b]	59 [10–313]	64 [10–313]	56 [10–261]	67 [11–151]	60 [14–167]
Serum creatinine [mg/dL][b]	1.3 [0.3–8.2]	1.1 [0.3–5.8]	1.3 [0.4–8.2]	1.6 [0.9–4]	1.4 [0.5–5.1]
Albumin [g/dL][b]	2.5 [1–4.1]	2.5 [1–4.1][c]	NA	2.2 [1.8–2.8]	2.7 [2.1–3.4]
CRP [mg/dL][b]	13.4 [0.4–40.5]	13.9 [0.5–40.5]	NA	9.4 [0.4–29.8]	11.3 [1.3–37.3]
Sepsis	496 (88)	207 (100)	226 (80)	3 (25)	60 (100)
RRT	96 (17)	23 (11)	55 (20)	2 (17)	16 (27)
ECMO	6 (8)	NA	NA	3 (25)	3 (5)
TZP administration and TDM					
Type of infusion	CI/II	CI	CI	CI	II
Number of TDM samples/patient[a]	7 [2–50]	6 [2–11]	3 [2–6]	2 [2–2]	33 [14–50]
Analytical method		HPLC/LC–MS/MS	HPLC	HPLC–UV	UHPLC–MS/MS
Total plasma PIP conc. [mg/L][a]	75.5 [0.1–812.8]	69.3 [0.3–725]	54 [10–300]	66 [5.1–467]	111.1 [0.1–812.8]

[a] Values presented per study: *n* (%) for categorical variables; median [minimum–maximum] for continuous variables

[b] Median value per patient (as multiple observations were obtained)

[c] 52 patients with missing albumin values (replaced by median of study population)

ABW adjusted body weight [29], *IBW* ideal body weight (Devine formula) [29], *BMI* body mass index, *eGFR* estimated glomerular filtration rate (Cockcroft-Gault) [30], *CRP* C-reactive protein, *RRT* renal replacement therapy, *ECMO* extracorporeal membrane oxygenation, *TZP* piperacillin/tazobactam, *PIP* piperacillin, *TDM* therapeutic drug monitoring, *CI* continuous infusion, *II* intermittent infusion, *(U)HPLC(–UV)* (ultra) high-performance liquid chromatography (–ultraviolet detection), *MS/MS* tandem mass spectrometry, *conc.* Concentration, *NA* not available

(2) Bayesian 1 (B1): prediction of concentrations considering the first TDM sample for each patient ($n_{predicted\_samples\_B1} = 3093$)
(3) Bayesian 2 (B2): prediction of concentrations considering the first and second TDM sample for each patient ($n_{predicted\_samples\_B2} = 2532$).

By considering TDM data, model parameter variability is taken into account to refine the initial PopPK parameters for the individual patient. The integrated samples for Bayesian forecasting were different, however, the large number and wide range of concentrations increased the comparability of predictions.

Model evaluation included statistical and graphical methods, employing prediction- and simulation-based diagnostics as recommended [25]. First, numerical comparisons between predicted and observed PIP plasma concentrations (c_{pred} and c_{obs}) served to quantify model-specific relative prediction errors (rPE) based on the total evaluation dataset and additionally for each individual study. The median relative prediction error (MPE) and median absolute relative prediction error (MAPE) were calculated to reflect accuracy and precision.

$$rPE_{ij} = \frac{c_{pred,i,j} - c_{obs,i,j}}{(c_{pred,i,j} + c_{obs,i,j})/2}, \quad (1)$$

$_i$: individual patient, $_j$: measurement.

$$MPE[\%] = median(\{rPE_{1,1}, \ldots, rPE_{i,j}\}) * 100, \quad (2)$$

$$MAPE[\%] = median(\{|rPE_{1,1}|, |\ldots|, |rPE_{i,j}|\}) * 100. \quad (3)$$

To facilitate clinical interpretation, an illustration based on the rPE (A/B1/B2) translated to PIP concentrations was performed:

$$c_{rPEi,j} = c_{target} - \left(c_{target} * rPE_{i,j}\right). \quad (4)$$

The target concentration (c_{target}) was exemplarily defined as 64 mg/L, which is related to the 4xMIC-epidemiological cutoff (ECOFF) for *Pseudomonas aeruginosa* according to the European Committee on Antimicrobial Susceptibility Testing (EUCAST). Density plots illustrate the range of the expected concentrations; i.e., the deviations from c_{target} due to model-specific inaccuracy and imprecision are simultaneously reflected according to the identified rPE. A symmetric target range of 32–96 mg/L (2-6xMIC-ECOFF) was considered clinically acceptable and attainment rates (%) were quantified.

Furthermore, the evaluation was stratified by CI and II to account for different PIP infusion regimens (studies 1–3: CI, $n_{samples\_CI} = 1819$; study 4: II, $n_{samples\_II} = 1835$).

Apart from numerical evaluation, c_{pred} and c_{obs} were compared graphically using goodness-of-fit (GOF) plots and prediction-corrected visual predictive checks (pcVPC) ($n_{simulations} = 1000$) [33].

The models that best represented the total evaluation dataset numerically and graphically were used to further investigate factors influencing the predictive performance:

(1) To explore the precision of Bayesian predictions as a function of time after integration of TDM data, the absolute rPE (B1/B2) vs. time after TDM (TaTDM) was assessed by descriptive statistics and the Kendall rank correlation coefficient (τ). This analysis was performed using studies 1 and 2, as these provided observation periods > 7 days (489 patients, CI, $n_{predicted\_samples\_B1/B2} = 1306/817$).
(2) To quantify the differences between (i) males vs. females, and (ii) non-renal replacement therapy (non-RRT) vs. RRT patients, point-biserial correlations ($r_{pb\_corrected}$) were calculated between the rPE (A) and each binomial group.

Besides, all 24 models underwent independent subgroup evaluations using the external dataset stratified by sex and RRT, and the above-described numerical analysis and GOF plots.

Results

Predictive performance of population pharmacokinetic models for piperacillin

The predictive performance markedly varied between all models (total dataset, A/B1/B2: MPE − 135.6–78.3/− 77.3–72.3/− 55.8–65.3%; MAPE 35.7–135.6/26.2–77.7/23.8–67.3%). Overall, data of each study population were predicted similarly to the pooled dataset, except that centers with CI performed slightly better in the model evaluation than that with II (see below). For most models, a common predictive trend (e.g., MPE < / > 0) and comparable performance per study site was observed (ESM_Main Tables S3–5). Figure 1 illustrates the accuracy and precision of the models based on the total dataset. Differences in performance were most evident for population predictions (A). Underestimation (MPE < 0; $n_{models\_A/B1/B2}$: 17/15/11) was more frequent than overestimation (MPE > 0; $n_{models\_A/B1/B2}$: 7/9/13). Consideration of TDM samples improved predictions for all models (ESM_Main Table S5). Models developed merely based on ICU patients did not generally appear superior to models without underlying ICU populations. However, taking together minimum inaccuracy (MPE) and imprecision (MAPE), the models by Kim (2022) [34], Klastrup [35] and Udy et al. [36], all originating from ICU patient data, performed best (A/B1/B2: |MPE| < 13/7.7/7.5%; MAPE < 37.5/28.0/28.4%). Graphical GOF plots agreed with these numerical results.

Based on the rPE of the three best-predicting models, Fig. 2 illustrates the range of expected concentrations targeting 64 mg/L (see ESM_Main Fig. S2 for all model results). A symmetric range with a median concentration near c_{target} indicates high predictive performance, with improved model precision and accuracy associated with more concentrations approaching 64 mg/L. Over- or underestimation is reflected by the frequency of out-of-range concentrations. Overall, higher target range attainment rates were observed for Bayesian approaches, and the model by Kim et al. showed the best result for the pooled dataset (A/B1/B2: 62.7/77.6/80.6%).

The model evaluation revealed overall better predictions for data collected during CI compared to II (A: $|MPE|_{median\_all\_models}$: 27.8 vs. 35.7%; $MAPE_{median\_all\_models}$ 46.2 vs. 50.9%). Integration of a second TDM sample (B2) substantially improved the predictions for II, but not for CI (ESM_Main Table S5). The model by Klastrup et al. showed the best numerical accuracy and precision for CI but was inferior to Kim and Udy et al. regarding II. With respect to the 95% confidence intervals of MPE and MAPE, however, the performance of the latter models partially overlapped (ESM_Main Fig. S4). GOF plots and pcVPCs confirmed the numerical evaluation results (ESM_Main Tables S3–5, Figs. S3–6).

A significant, albeit small positive correlation between the precision of Bayesian predictions and TaTDM was detected, i.e., higher imprecision given longer time (Kim, Klastrup, Udy et al.: $\tau_{B1} = 0.08, 0.11, 0.13$; $\tau_{B2} = 0.10, 0.10, 0.10$; $p_{B1/B2} < 0.001$). Comparison of MAPE values vs.

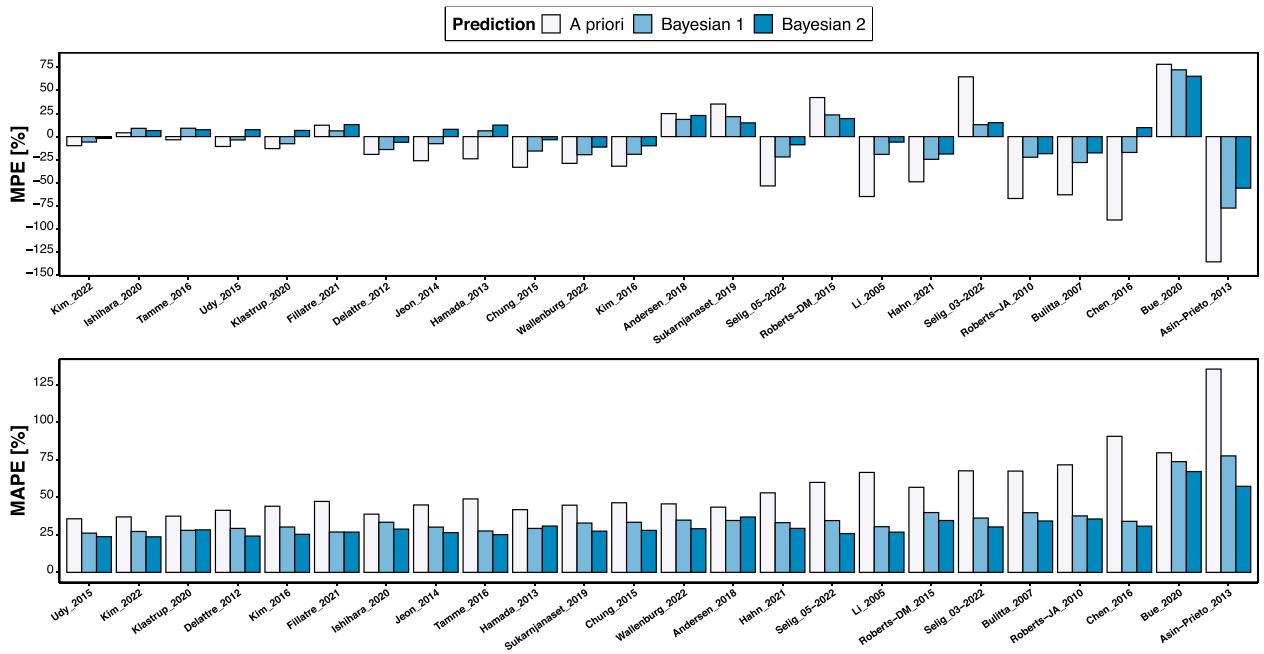


Fig. 1 Predictive performance of the evaluated population pharmacokinetic models. Median relative prediction error (MPE) [%] as a measure of accuracy and median absolute relative prediction error (MAPE) [%] reflecting precision of the predictions, comparing the observed plasma piperacillin (PIP) concentrations for each model; A priori (A): Prediction of all concentrations based on dosing history and patient covariates only; Bayesian 1 (B1): Prediction of concentrations considering the first therapeutic drug monitoring (TDM) sample for each patient; Bayesian 2 (B2): Prediction of concentrations considering the first and second TDM sample for each patient; models were ordered (from left to right) according to the model-specific sum (A + B1 + B2) of (absolute) MPE and MAPE values, respectively

TaTDM further indicated decreasing precision (e.g., Udy et al. B1: $MAPE_{24h}/MAPE_{168h}$: 17.4/29.9%; B2: $MAPE_{24h}/MAPE_{168h}$: 22.3/40.2%, Fig. 3). Moreover, higher imprecision and underestimation were observed in (i) females compared to males (Kim/Klastrup/Udy et al.: $r_{pb\_corrected} = -0.01/-0.09/-0.08$; $p = 0.79/< 0.001/< 0.01$) and (ii) patients undergoing RRT (Kim/Klastrup/Udy et al.: $r_{pb\_corrected} = -0.35/-0.26/-0.17$; $p < 0.001$), as detailed in ESM_Main Fig. S7, Table S6.

Stratified subgroup evaluations of all 24 models are presented in ESM_Subgroup Figs. S1–5, Tables S1–3. Separation by sex revealed overall higher accuracy and precision for men compared to women (A: $|MPE|_{median\_all\_models}$: 28.7 vs. 35.8%; $MAPE_{median\_all\_models}$ 46.7 vs. 49.3%); however, the difference was not apparent in Bayesian predictions. Notably, the model by Kim et al. performed best in women (A/B1/B2: MPE_{female} −11.6/−6.4/0.8%; $MAPE_{female}$ 39.3/29/24.5%). Stratification by non-RRT/RRT patients disclosed more models underestimating PIP concentrations in the RRT group ($n_{A/B1/B2}$: 19/19/14), and predictions appeared worse overall (A: $|MPE|_{median\_all\_models}$: 28.7 vs. 47.7%; $MAPE_{median\_all\_models}$ 47.6 vs. 53.7%). While the model by Kim et al. was favorable in non-RRT patients (A/B1/B2: MPE −1/−2.4/1.6%; MAPE 34.7/27.2/23.8%), the model by Roberts et al. (2015) [37] produced the most accurate and precise population predictions in RRT patients (A: MPE 0.8%; MAPE 33.3%). When considering TDM data, adequate predictions for RRT patients were achieved with several models (e.g., Tamme et al. [38], B1/B2: MPE 0.6/−2.9%; MAPE 22.1/20.9%). Further subgroup analyses (e.g., female + RRT) are available in ESM (ESM_Subgroup Figs. S3–5, Tables S1–3).

Taken together all evaluation results, a flowchart showing the best-performing models for each subgroup and clinical target attainment within our external data is presented (ESM_Subgroup Figs. S6–7). The candidate models were then implemented in the open-access TDMx dosing software (www.TDMx.eu), which allows to explore PIP MIPD scenarios relevant for clinical practice (Fig. 4).

Discussion

This study investigates the external validity of 24 PopPK models for PIP in a multicenter cohort of 561 ICU patients and highlights MIPD candidates appropriate for specific situations and patient groups (e.g., CI/II administration, male/female, RRT/non-RRT), considering a priori and Bayesian predictions.

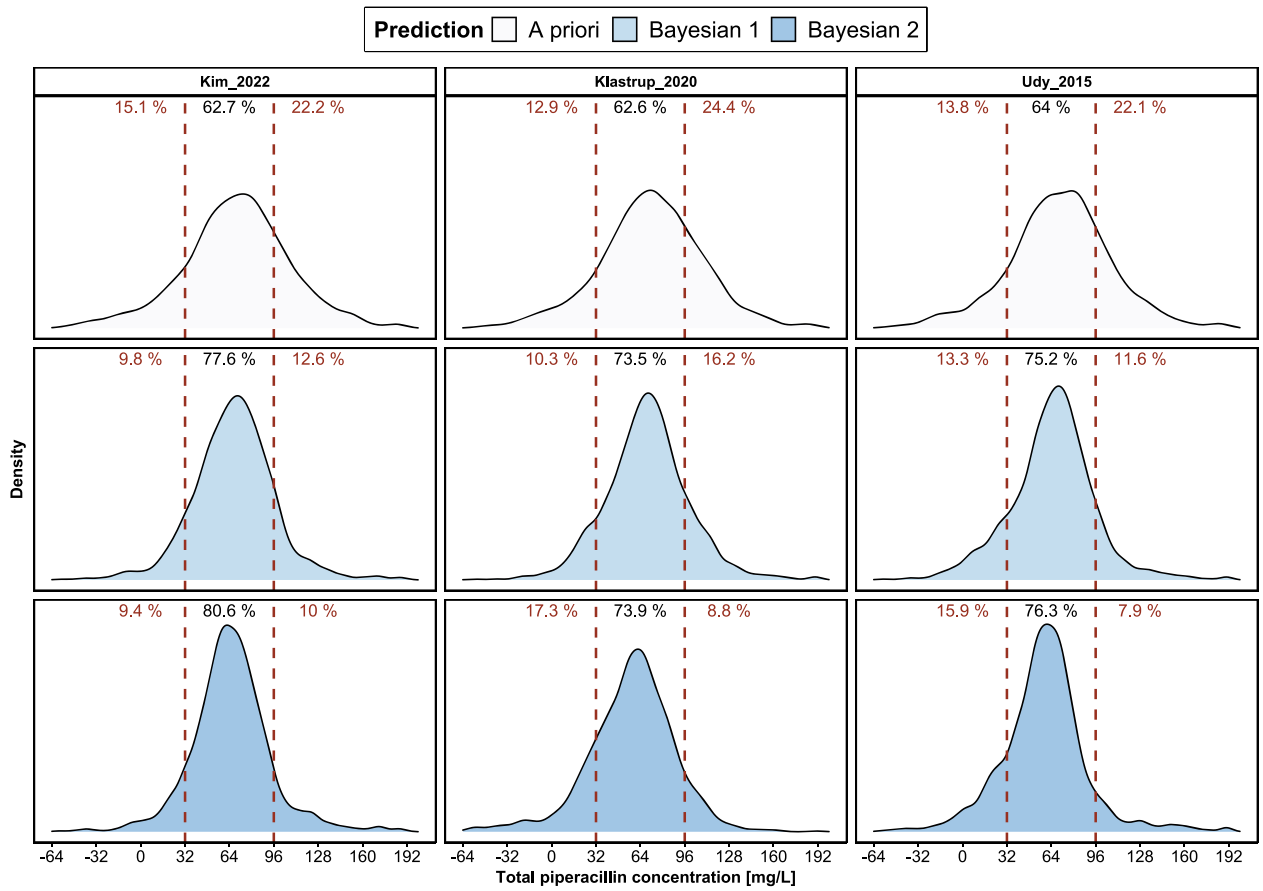


Fig. 2 Expected target attainment of the population pharmacokinetic models by Kim, Klastrup and Udy et al. by translating the relative prediction errors for the total evaluation dataset to piperacillin plasma concentrations. A priori (A): Prediction of all concentrations based on dosing history and patient covariates only; Bayesian 1 (B1): Prediction of concentrations considering the first therapeutic drug monitoring (TDM) sample for each patient; Bayesian 2 (B2): Prediction of concentrations considering the first and second TDM sample for each patient; red dashed lines represent the lower and upper limits of the target concentration range; percentages indicate the proportion of predicted concentrations within (black color) and outside (red color) the target range

Acceptable values for model accuracy (e.g., |MPE| < 20%) and imprecision (e.g., MAPE < 30%), originating from a propofol PK model evaluation by Miyabe-Nishiwaki and their feasibility for models of time-dependent antibiotics have been debated [24, 39]. However, alternative thresholds to assess the suitability of a model are lacking. The highest predictive performance within our external dataset was observed for the models by Kim, Klastrup, and Udy et al., but none of these achieved the above-mentioned thresholds for a priori predictions (A). When considering TDM measurements, however, several models met the defined thresholds, highly suggesting the combination of model-based therapy and TDM. Similar to our work, Chai et al. recently assessed the accuracy of the ID-ODS dosing software using external data from 75 critically ill patients, concluding overall improved predictive performance/dosing using Bayesian forecasting [40]. While our study showed no substantial benefit of considering two vs. one TDM sample for predicting CI data, two samples were superior to one for II data. However, Bayesian precision has been shown to decrease over time, suggesting the consideration of timely TDM samples during CI. Fixed time intervals cannot be recommended based on our data and require further investigation. Continuous drug monitoring using biosensors with direct feedback of real-time antibiotic concentrations to PK models represents a future scenario, enabling fully automated closed-loop techniques between drug sensor, dosing software, and self-adaptive infusion pumps [41–43].

Our evaluation results are largely consistent with a previous study conducted in France using external monocentric data from 30 critically ill patients [24]. Despite methodological differences, the authors also recommended the models by Udy and Klastrup et al. as

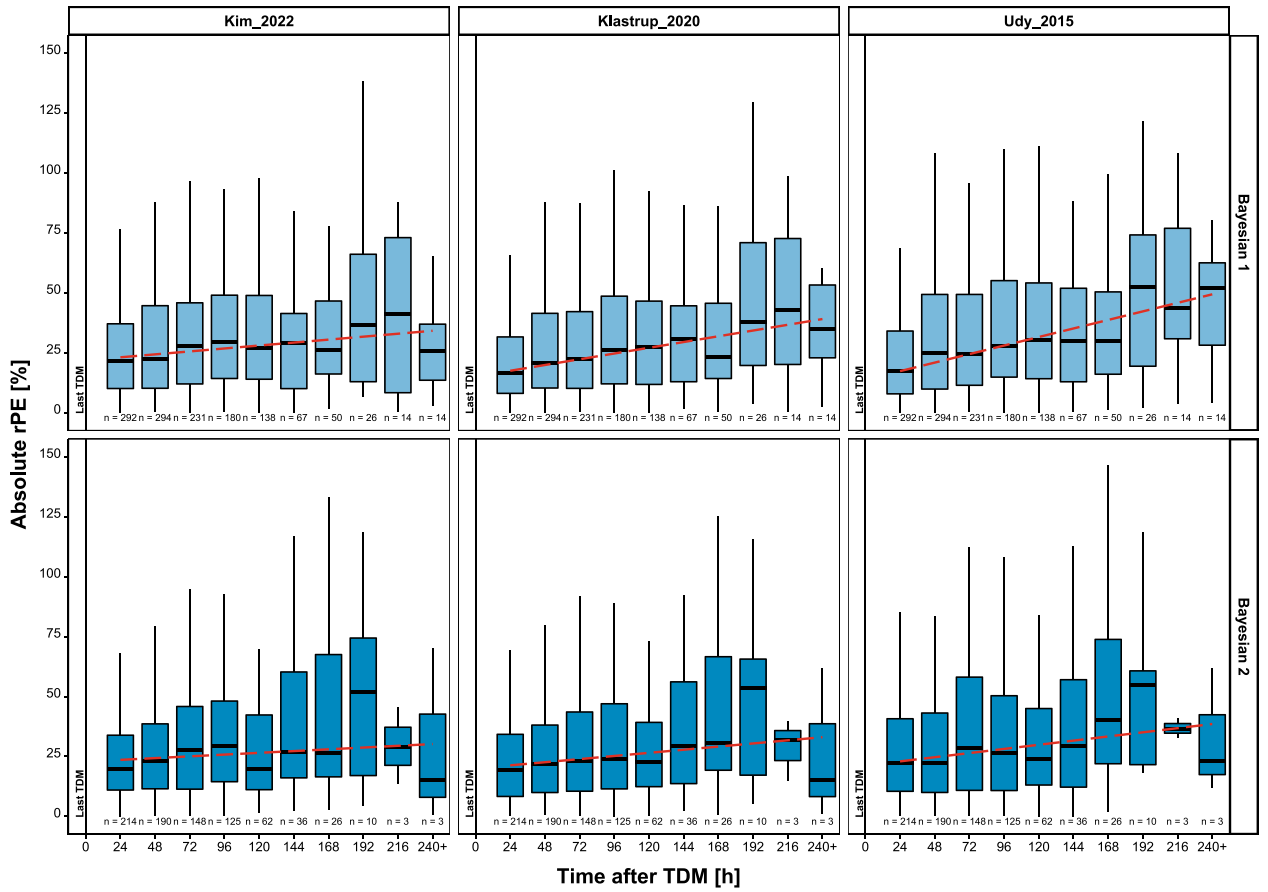
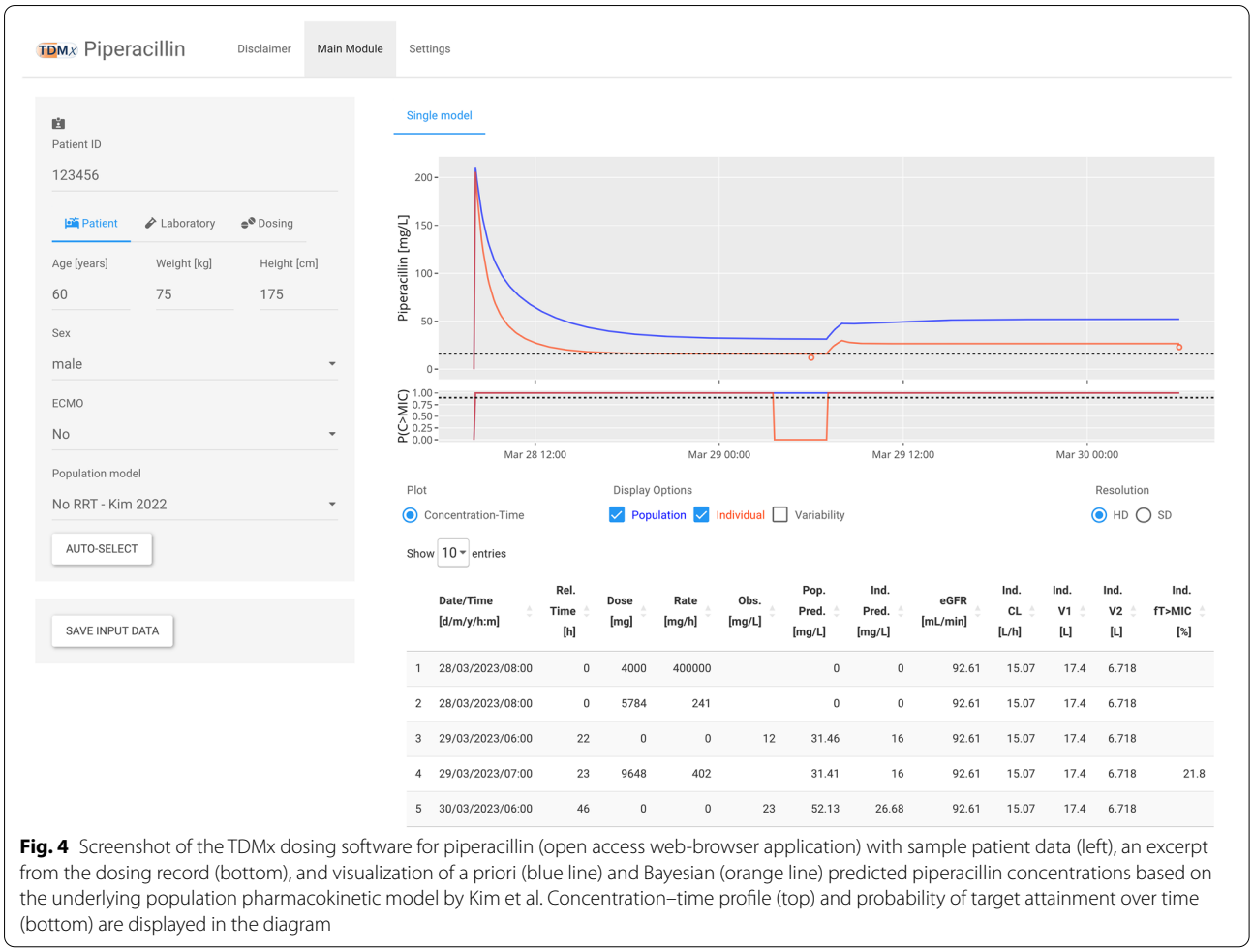


Fig. 3 Precision of the models by Kim, Klastrup and Udy et al. as a function of time since the last therapeutic drug monitoring. Absolute relative prediction error (Absolute rPE) [%] of Bayesian predictions versus time interval [h] since the last considered therapeutic drug monitoring (TDM) sample; values based on the pooled evaluation dataset of studies 1 and 2; Bayesian 1 (B1): Prediction of concentrations considering the first TDM sample for each patient; Bayesian 2 (B2): Prediction of concentrations considering the first and second TDM sample for each patient; n: number of binned concentrations for each boxplot (binning: {24, 48, 72…} + −12 h); boxplots represent the 25th, 50th and 75th percentiles, whiskers the 1.5-fold interquartile range; red dashed lines: linear regression through binned median absolute prediction errors (MAPEs)

potential candidates for MIPD (Kim et al. was not published at that time), making our results likely transferable to other centers/countries. In our evaluation, the model by Udy et al. best predicted data collected during II and Klastrup et al. best predicted CI data, which seems plausible since the models were built upon II and CI data, respectively. Both models showed similar performance to that of Kim et al., which revealed most appropriate for the pooled dataset (CI + II), also due to less gender bias and superior performance in non-RRT patients.

Our analyses demonstrated that PIP predictions were overall more accurate and precise for males, presumably as women were underrepresented in all datasets underlying the 24 investigated models (ESM_Main Table S1). The PK parameters estimated (e.g., V_d, CL) may, thus, have been driven by men. A smaller hydrophilic V_d due to a relatively lower muscle mass and a higher proportion of adipose tissue may explain higher/accumulating concentrations in women. Moreover, renal clearance is physiologically about 10% lower than in men [44]. Although some models accounted for sex by integrating renal function using the Cockcroft-Gault (CG) equation [30], this did not appear sufficient to avoid differences in predictions. RRT patients showed considerably worse predictions than non-RRT patients (A). Some models were built upon data excluding RRT patients (e.g., Klastrup and Udy et al.) or were derived from predominantly non-RRT patients (e.g., Kim et al.: 79%), which may explain lowered predictive performance. In contrast, some models were developed exclusively based on RRT patients (ESM_Main Table S1); however, only the model by Roberts et al. (2015) performed well regarding a priori predictions for RRT patients. RRT is associated with highly variable patient PKs due to multiple influences like residual diuresis,



	Date/Time [d/m/y/h:m]	Rel. Time [h]	Dose [mg]	Rate [mg/h]	Obs. [mg/L]	Pop. Pred. [mg/L]	Ind. Pred. [mg/L]	eGFR [mL/min]	Ind. CL [L/h]	Ind. V1 [L]	Ind. V2 [L]	Ind. fT>MIC [%]
1	28/03/2023/08:00	0	4000	400000		0	0	92.61	15.07	17.4	6.718	
2	28/03/2023/08:00	0	5784	241		0	0	92.61	15.07	17.4	6.718	
3	29/03/2023/06:00	22	0	0	12	31.46	16	92.61	15.07	17.4	6.718	
4	29/03/2023/07:00	23	9648	402		31.41	16	92.61	15.07	17.4	6.718	21.8
5	30/03/2023/06:00	46	0	0	23	52.13	26.68	92.61	15.07	17.4	6.718	

Fig. 4 Screenshot of the TDMx dosing software for piperacillin (open access web-browser application) with sample patient data (left), an excerpt from the dosing record (bottom), and visualization of a priori (blue line) and Bayesian (orange line) predicted piperacillin concentrations based on the underlying population pharmacokinetic model by Kim et al. Concentration–time profile (top) and probability of target attainment over time (bottom) are displayed in the diagram

type of RRT, dialysis membrane, duration, and intensity settings [41]. Unfortunately, RRT-related variables were not covered by the evaluation dataset, which precluded a detailed exploration of predictions in RRT patients.

A recent randomized controlled trial by Ewoldt et al. investigated a potential benefit of MIPD in 388 critically ill patients receiving ciprofloxacin and β-lactams [45]. Such investigations are urgently needed to link an impact of novel dosing strategies to clinical outcomes. Surprisingly, the authors did not find increased target attainment when using a commercial MIPD software. The applied model for PIP (Andersen et al. [46]) was developed based on clinical data from 22 septic non-ICU patients and published evidence of its transferability to external populations seems lacking. In fact, our study revealed overestimation and mediocre performance of the respective model in the investigated critically ill patients (A/B1/B2: MPE 24.9/18.6/22.8%; MAPE 43.5/34.6/36.9%). Thus, it can only be speculated whether model selection had a relevant influence on the study results, also because the trial included several β-lactams, whereas our study included only PIP [22, 47, 48]. However, both studies combined highlight the key role of externally evaluating PopPK models prior to clinical implementation.

Some limitations of this study shall be acknowledged. First, external evaluation depends on the quality of the evaluation dataset and clinical collection may be error prone, potentially distorting model predictions [49]. To limit this shortcoming, our combined dataset was carefully reviewed and each model assessment was conducted identically to enhance inter-model comparability. Second, although this was a multicenter study, all patients were admitted to national ICUs (Germany), and the evaluation for II dosing is based on a single monocentric trial, limiting extrapolation to other settings. Third, we only included parametric models and cannot exclude similar or better performance of models built with less common modeling approaches. Fourth, our evaluation considered PIP concentrations only; however, recent studies support a holistic PIP/tazobactam stewardship, particularly in severe cases of sepsis. Considering tazobactam may

improve target attainment and antibacterial effectiveness for both drugs simultaneously, and future evaluation of a combined model would, thus, be desirable [24, 50]. Fifth, the clinical dataset covered PIP_{total}, whereas some models were developed based on $PIP_{unbound}$. We adjusted the available concentrations according to the protein binding stated by the manufacturer (30%) [24]. However, the unbound fraction may vary in critically ill patients, and possible skewing of the results can therefore not be excluded [27, 32, 51]. Sixth, our study exemplarily followed a recommended PIP c_{target} of 64 mg/L, but for β-lactams "the optimal PK/PD target remains debated" according to Novy et al. [47]. A range of 32–96 mg/L was considered acceptable. Lower concentrations are at risk of being ineffective due to variable protein binding and insufficient penetration at the infection site, while higher concentrations are associated with significantly increased 28-day mortality [16]. Seventh, some minor covariate adjustments were necessary due to unavailable covariates (e.g., eGFR instead of measured CL_{CR}) (ESM_Main Table S1). As several characteristics are included in the CG equation [30], a multivariate positive effect on model performance has been discussed [24]. Last, external evaluations (and model selection) include a subjective component, yet we followed a systematic approach with defined criteria to make the evaluation as objective as possible.

In conclusion, the predictive performance of published PopPK models for PIP in critically ill patients varied remarkably. Selecting an appropriate model is essential for high-quality MIPD. Models should be combined with TDM as soon as possible to improve predictability. Studies on the prospective application of MIPD of PIP at the bedside using Bayesian dosing software like TDMx is desired to assess a clinical benefit for patients.

Finally, the authors would like to emphasize that this evaluation is not intended to discredit any PK model, but rather to extrapolate potential candidates for cross-center clinical application in critically ill patients, to move from conventional dosing of PIP toward a more individualized approach.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1007/s00134-023-07154-0.

Author details

[1] Department of Anaesthesiology, University Hospital, LMU Munich, Marchioninistrasse 15, 81377 Munich, Germany. [2] Department of Anaesthesiology and Intensive Care Medicine, General Hospital of Heidenheim, 89522 Heidenheim, Germany. [3] Department of Pharmacy, General Hospital of Heidenheim, 89522 Heidenheim, Germany. [4] Institute for Infectious Diseases and Infection Control, University Hospital, Friedrich-Schiller-University Jena, 07747 Jena, Germany. [5] Institute of Pharmacy, University of Regensburg, 93053 Regensburg, Germany. [6] Faculty of Pharmacy, University of Montréal, Pavillon Jean-Coutu, 2940 Chemin de Polytechnique, Montréal, QC H3T 1J4, Canada. [7] Laboratory Dr. Brunner, Laboratory Medical Care Center Konstanz GmbH, 78464 Constance, Germany. [8] Department of Pharmaceutical and Medical Chemistry, Clinical Pharmacy, University of Muenster, 48149 Muenster, Germany. [9] Department of Clinical Pharmacology, Medical University Vienna, 1090 Vienna, Austria. [10] Department of Clinical Pharmacy, Institute of Pharmacy, University of Hamburg, 20146 Hamburg, Germany.

Author contributions

SG, AB, SH, CL, CD, LS, SGW, JB and UL contributed to the study conception and design. AB, AR, OF, SH, CD, CS, MZ and JZ acquired and provided clinical data. SG and UL performed the literature search, model reconstruction, statistical data analyses and prepared the original manuscript draft. All authors participated in the interpretation of the evaluation results and critically revised the manuscript. All authors have final responsibility and accountability for the accuracy and integrity of any part of the work.

Funding

Open Access funding enabled and organized by Projekt DEAL. This work was funded by the German Research Foundation (DFG, Project No.: 504995322). UL acknowledges the funding of research by the Munich Clinician-Scientist Program. The funders had no role in the design of this study, in the collection, analysis, and interpretation of data, in the writing of the report, and in the decision to submit the paper for publication.

Data availability

The clinical data used and analyzed in the current external model evaluation were derived from four previously published studies. Access to each dataset for research purposes may be available upon reasonable request.

Declarations

Conflicts of interest

AB reports lecture fees and support for attending meetings from Fresenius Medical Care (FMC), MSD Sharp & Dohme GmbH, Pfizer Pharma GmbH, Verein zur Förderung der GHE e.V. (Hannover), Labor Limbach and participation on an Advisory Board at Stuttgarter Intensivkongress (SIK), Anästhesie und Intensivmedizin (Journal), Antibiotics (Journal). OF declares honoraria for lectures and presentations of MSD Germany and Fresenius Kabi Germany. SH reports grants from the Federal Ministry of Education and Research (BMBF), lecture fees from Pfizer, MSD, InfectoPharm, Advanz and Philips, and support for attending meetings from Pfizer and Advanz. MZ received funding of research from CytoSorbents Europe GmbH and was part of the Advisory Board at Gilead Ambisome. TS declares funding by the Munich Clinician-Scientist Program (Faculty of Medicine, LMU Munich) and serves as a speaker for the program. CS reports speaker honoraria from CytoSorbents Europe GmbH. UL reports consulting fees from CytoSorbents Europe GmbH. SGW reports consulting fees from Merck KGaA and Medicines for Malaria Venture, research grants from Boehringer Ingelheim, as well as speaker honoraria from GSK. SGW is the founder and lead developer of the TDMx project (TDMx.eu).

Ethics approval

Ethical approval for this study was waived by the local institutional review board (No.: 21-1162 KB).

Consent for publication

All authors read and approved the final manuscript for submission.

Open Access

This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright

holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc/4.0/.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 5 April 2023 Accepted: 27 June 2023
Published online: 13 July 2023

References

1. Evans L, Rhodes A, Alhazzani W et al (2021) Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. Crit Care Med 49:e1063. https://doi.org/10.1097/CCM.0000000000005337
2. Engel C, Brunkhorst FM, Bone H-G et al (2007) Epidemiology of sepsis in Germany: results from a national prospective multicenter study. Intensive Care Med 33:606–618. https://doi.org/10.1007/s00134-006-0517-7
3. Seymour CW, Gesten F, Prescott HC et al (2017) Time to treatment and mortality during mandated emergency care for sepsis. N Engl J Med 376:2235–2244. https://doi.org/10.1056/NEJMoa1703058
4. Kumar A, Roberts D, Wood KE et al (2006) Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. Crit Care Med 34:1589–1596. https://doi.org/10.1097/01.CCM.0000217961.75225.E9
5. Zhang D, Micek ST, Kollef MH (2015) Time to appropriate antibiotic therapy is an independent determinant of postinfection ICU and hospital lengths of stay in patients with sepsis. Crit Care Med 43:2133–2140. https://doi.org/10.1097/CCM.0000000000001140
6. Abdul-Aziz MH, Alffenaar J-WC, Bassetti M et al (2020) Antimicrobial therapeutic drug monitoring in critically ill adult patients: a Position Paper. Intensive Care Med 46:1127–1153. https://doi.org/10.1007/s00134-020-06050-1
7. Roberts JA, Lipman J (2009) Pharmacokinetic issues for antibiotics in the critically ill patient. Crit Care Med 37:840–851. https://doi.org/10.1097/CCM.0b013e3181961bff
8. Gonçalves-Pereira J, Póvoa P (2011) Antibiotics in critically ill patients: a systematic review of the pharmacokinetics of β-lactams. Crit Care 15:R206. https://doi.org/10.1186/cc10441
9. Scharf C, Liebchen U, Paal M et al (2020) The higher the better? Defining the optimal beta-lactam target for critically ill patients to reach infection resolution and improve outcome. J Intensive Care 8:86. https://doi.org/10.1186/s40560-020-00504-w
10. Cusumano JA, Klinker KP, Huttner A et al (2020) Towards precision medicine: therapeutic drug monitoring-guided dosing of vancomycin and β-lactam antibiotics to maximize effectiveness and minimize toxicity. Am J Health Syst Pharm 77:1104–1112. https://doi.org/10.1093/ajhp/zxaa128
11. Bulman ZP, Wicha SG, Nielsen EI et al (2022) Research priorities towards precision antibiotic therapy to improve patient care. Lancet Microbe 3:e795–e802. https://doi.org/10.1016/S2666-5247(22)00121-5
12. Liebchen U, Paal M, Scharf C et al (2020) The ONTAI study - a survey on antimicrobial dosing and the practice of therapeutic drug monitoring in German intensive care units. J Crit Care 60:260–266. https://doi.org/10.1016/j.jcrc.2020.08.027
13. Tannous E, Lipman S, Tonna A et al (2020) Time above the MIC of piperacillin-tazobactam as a predictor of outcome in pseudomonas aeruginosa bacteremia. Antimicrob Agents Chemother 64:e02571-e2619. https://doi.org/10.1128/AAC.02571-19
14. Roberts JA, Paul SK, Akova M et al (2014) DALI: defining antibiotic levels in intensive care unit patients: are current β-lactam antibiotic doses sufficient for critically Ill patients? Clin Infect Dis 58:1072–1083. https://doi.org/10.1093/cid/ciu027
15. Scharf C, Paal M, Schroeder I et al (2020) Therapeutic drug monitoring of meropenem and piperacillin in critical illness-experience and recommendations from one year in routine clinical practice. Antibiotics (Basel) 9:131. https://doi.org/10.3390/antibiotics9030131
16. Hagel S, Bach F, Brenner T et al (2022) Effect of therapeutic drug monitoring-based dose optimization of piperacillin/tazobactam on sepsis-related organ dysfunction in patients with sepsis: a randomized controlled trial. Intensive Care Med 48:311–321. https://doi.org/10.1007/s00134-021-06609-6
17. Gijsen M, Dreesen E, Wauters J et al (2022) The TARGET trial as a plea for model-informed precision dosing of piperacillin/tazobactam in patients with sepsis. Intensive Care Med. https://doi.org/10.1007/s00134-022-06679-0
18. Wicha SG, Märtson A-G, Nielsen EI 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
19. Jager NGL, Chai MG, van Hest RM et al (2022) Precision dosing software to optimize antimicrobial dosing: a systematic search and follow-up survey of available programs. Clin Microbiol Infect 28:1211–1224. https://doi.org/10.1016/j.cmi.2022.03.041
20. Tängdén T, Ramos Martín V, Felton TW et al (2017) The role of infection models and PK/ PD modelling for optimising care of critically ill patients with severe infections. Intensive Care Med 43:1021–1032. https://doi.org/10.1007/s00134-017-4780-6
21. US FDA (2022) Population pharmacokinetics: guidance for industry. In: Clinical pharmacology. https://www.fda.gov/media/128793/download. Accessed 31 Jan 2023
22. Cotta MO, Lipman J, De Waele J (2023) Advancing precision-based antimicrobial dosing in critically ill patients. Intensive Care Med. https://doi.org/10.1007/s00134-022-06969-7
23. El-Haffaf I, Caissy J-A, Marsot A (2021) Piperacillin-tazobactam in intensive care units: a review of population pharmacokinetic analyses. Clin Pharmacokinet 60:855–875. https://doi.org/10.1007/s40262-021-01013-1
24. El-Haffaf I, Guilhaumou R, Velly L, Marsot A (2022) Using a validated population pharmacokinetic model for dosing recommendations of continuous infusion piperacillin for critically Ill adult patients. Clin Pharmacokinet. https://doi.org/10.1007/s40262-022-01118-1
25. Cheng Y, Wang C-Y, Li Z-R et al (2021) Can population pharmacokinetics of antibiotics be extrapolated? Implications of external evaluations. Clin Pharmacokinet 60:53–68. https://doi.org/10.1007/s40262-020-00937-4
26. Richter DC, Frey O, Röhr A et al (2019) Therapeutic drug monitoring-guided continuous infusion of piperacillin/tazobactam significantly improves pharmacokinetic target attainment in critically ill patients: a retrospective analysis of four years of clinical experience. Infection 47:1001–1011. https://doi.org/10.1007/s15010-019-01352-z
27. Schießer S, Hitzenbichler F, Kees MG et al (2021) Measurement of free plasma concentrations of beta-lactam antibiotics: an applicability study in intensive care unit patients. Ther Drug Monit 43:264–270. https://doi.org/10.1097/FTD.0000000000000827
28. Zander J, Döbbeler G, Nagel D et al (2016) Piperacillin concentration in relation to therapeutic range in critically ill patients—a prospective observational study. Crit Care 20:79. https://doi.org/10.1186/s13054-016-1255-z
29. Kirkpatrick DB (1999) Pharmacokinetics of gentamicin in 957 patients with varying renal function dosed once daily. Br J Clin Pharmacol 47(6):637–643. https://doi.org/10.1046/j.1365-2125.1999.00938.x
30. Cockcroft DW, Gault MH (1976) Prediction of creatinine clearance from serum creatinine. Nephron 16:31–41. https://doi.org/10.1159/000180580
31. Collins GS, Reitsma JB, Altman DG, Moons KG (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. BMC Med 13:1. https://doi.org/10.1186/s12916-014-0241-z
32. El-Haffaf I, Guilhaumou R, Velly L, Marsot A (2022) Impact of piperacillin unbound fraction variability on dosing recommendations in critically ill patients. Br J Clin Pharmacol. https://doi.org/10.1111/bcp.15619
33. Bergstrand M, Hooker AC, Wallin JE, Karlsson MO (2011) Prediction-corrected visual predictive checks for diagnosing nonlinear mixed-effects models. AAPS J 13:143–151. https://doi.org/10.1208/s12248-011-9255-z
34. Kim YK, Kim HS, Park S et al (2022) Population pharmacokinetics of piperacillin/tazobactam in critically ill Korean patients and the effects of extracorporeal membrane oxygenation. J Antimicrob Chemother 77:1353–1364. https://doi.org/10.1093/jac/dkac059
35. Klastrup V, Thorsted A, Storgaard M et al (2020) Population pharmacokinetics of piperacillin following continuous infusion in critically Ill patients and impact of renal function on target attainment. Antimicrob Agents Chemother 64:e02556-e2619. https://doi.org/10.1128/AAC.02556-19

36. Udy AA, Lipman J, Jarrett P et al (2015) Are standard doses of piperacillin sufficient for critically ill patients with augmented creatinine clearance? Crit Care 19:28. https://doi.org/10.1186/s13054-015-0750-y
37. Roberts DM, Liu X, Roberts JA et al (2015) A multicenter study on the effect of continuous hemodiafiltration intensity on antibiotic pharmacokinetics. Crit Care 19:84. https://doi.org/10.1186/s13054-015-0818-8
38. Tamme K, Oselin K, Kipper K et al (2016) Pharmacokinetics and pharmacodynamics of piperacillin/tazobactam during high volume haemodiafiltration in patients with septic shock. Acta Anaesthesiol Scand 60:230–240. https://doi.org/10.1111/aas.12629
39. Miyabe-Nishiwaki T, Masui K, Kaneko A et al (2013) Evaluation of the predictive performance of a pharmacokinetic model for propofol in Japanese macaques (*Macaca fuscata* fuscata). J Vet Pharmacol Ther 36:169–173. https://doi.org/10.1111/j.1365-2885.2012.01404.x
40. Chai MG, Roberts JA, Farkas A, Cotta MO (2023) Accuracy of a precision dosing software program for predicting antibiotic concentrations in critically ill patients. J Antimicrob Chemother 78:354–358. https://doi.org/10.1093/jac/dkac392
41. Roberts JA, Bellomo R, Cotta MO et al (2022) Machines that help machines to help patients: optimising antimicrobial dosing in patients receiving extracorporeal membrane oxygenation and renal replacement therapy using dosing software. Intensive Care Med 48:1338–1351. https://doi.org/10.1007/s00134-022-06847-2
42. Bian S, Zhu B, Rong G, Sawan M (2021) Towards wearable and implantable continuous drug monitoring: a review. J Pharm Anal 11:1–14. https://doi.org/10.1016/j.jpha.2020.08.001
43. Ates HC, Roberts JA, Lipman J et al (2020) On-site therapeutic drug monitoring. Trends Biotechnol 38:1262–1277. https://doi.org/10.1016/j.tibtech.2020.03.001
44. Schwartz JB (2007) The current state of knowledge on age, sex, and their interactions on clinical pharmacology. Clin Pharmacol Ther 82:87–96. https://doi.org/10.1038/sj.clpt.6100226
45. Ewoldt TMJ, Abdulla A, Rietdijk WJR et al (2022) Model-informed precision dosing of beta-lactam antibiotics and ciprofloxacin in critically ill patients: a multicentre randomised clinical trial. Intensive Care Med. https://doi.org/10.1007/s00134-022-06921-9
46. Andersen MG, Thorsted A, Storgaard M et al (2018) Population pharmacokinetics of piperacillin in sepsis patients: should alternative dosing strategies be considered? Antimicrob Agents Chemother 62:e02306-e2317. https://doi.org/10.1128/AAC.02306-17
47. Novy E, Martinière H, Roger C (2023) The current status and future perspectives of beta-lactam therapeutic drug monitoring in critically Ill patients. Antibiotics (Basel) 12:681. https://doi.org/10.3390/antibiotics12040681
48. Liebchen U, Briegel J, Brinkmann A et al (2023) Individualised dosing of antibiotics in ICU patients: timing, target and model selection matter. Intensive Care Med. https://doi.org/10.1007/s00134-023-06990-4
49. Alihodzic D, Broeker A, Baehr M et al (2020) Impact of inaccurate documentation of sampling and infusion time in model-informed precision dosing. Front Pharmacol. https://doi.org/10.3389/fphar.2020.00172
50. Wallenburg E, Ter Heine R, Schouten JA et al (2022) An integral pharmacokinetic analysis of piperacillin and tazobactam in plasma and urine in critically Ill patients. Clin Pharmacokinet. https://doi.org/10.1007/s40262-022-01113-6
51. Colman S, Stove V, De Waele JJ, Verstraete AG (2019) Measuring unbound versus total piperacillin concentrations in plasma of critically Ill patients: methodological issues and relevance. Ther Drug Monit 41:325–330. https://doi.org/10.1097/FTD.0000000000000602