

ARTICLE

Bayesian Data Assimilation to Support Informed Decision Making in Individualized Chemotherapy

Corinna Maier^{1,2}, Niklas Hartung¹, Jana de Wiljes^{1,3}, Charlotte Kloft⁴ and Wilhelm Huisinga^{1,*}

An essential component of therapeutic drug/biomarker monitoring (TDM) is to combine patient data with prior knowledge for model-based predictions of therapy outcomes. Current Bayesian forecasting tools typically rely only on the most probable model parameters (maximum *a posteriori* (MAP) estimate). This MAP-based approach, however, does neither necessarily predict the most probable outcome nor does it quantify the risks of treatment inefficacy or toxicity. Bayesian data assimilation (DA) methods overcome these limitations by providing a comprehensive uncertainty quantification. We compare DA methods with MAP-based approaches and show how probabilistic statements about key markers related to chemotherapy-induced neutropenia can be leveraged for more informative decision support in individualized chemotherapy. Sequential Bayesian DA proved to be most computationally efficient for handling interoccasion variability and integrating TDM data. For new digital monitoring devices enabling more frequent data collection, these features will be of critical importance to improve patient care decisions in various therapeutic areas.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

☑ Typically, maximum *a posteriori* (MAP) estimation is used to combine prior knowledge with patient-specific monitoring data for model-based predictions to support individualized therapy (e.g., in chemotherapy).

WHAT QUESTION DID THIS STUDY ADDRESS?

☑ The study addresses major limitations of MAP-based prediction for efficient and reliable clinical decision support: (i) it only provides a point estimate without associated uncertainties, (ii) it does not transform into the most probable observation (e.g., nadir concentration), and (iii) it processes data in a batch.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

☑ Bayesian data assimilation (DA) methods are compared that allow a comprehensive uncertainty quantification by

approximating the full posterior distribution. We demonstrate the advantages of uncertainty quantification in therapeutic drug/biomarker monitoring for chemotherapy-induced neutropenia. Sequential DA methods were found to be particularly suitable for long-run analyses as they process data sequentially.

HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTIC?

☑ Comprehensive uncertainty quantification and recursive data processing enable reliable, efficient, and individual decision support during ongoing treatment and will become increasingly relevant with the development of novel digital monitoring devices.

In the presence of a narrow therapeutic window and large interpatient variability, therapeutic drug/biomarker monitoring (TDM) is indicated for safe and efficacious therapies. With the help of Bayesian forecasting tools, patient-specific data are combined with prior knowledge from previous clinical studies and a drug-specific model to enable model-informed precision dosing (MIPD).¹ Typically, only the most probable individual parameter values (i.e., the maximum *a posteriori* (MAP) estimates, are used to predict the individual therapy outcome without quantifying associated uncertainties.² Thus, relevant risks associated with a dosing regimen

selection (e.g., treatment inefficacy or unacceptable toxicity), are not determined hindering well-founded therapeutic decision making.

Quantifying associated uncertainties is at the heart of making more informed decisions, not only in clinical applications. In this article, we thoroughly compare in a TDM context different Bayesian data assimilation (DA) methods that either (i) estimate the full posterior distribution (termed full Bayesian approaches) to quantify uncertainties, or (ii) estimate only its mode (MAP estimation) without uncertainty quantification. The full Bayesian approaches comprise not only methods that

¹Institute of Mathematics, University of Potsdam, Potsdam, Germany; ²Graduate Research Training Program PharMetrX: Pharmacometrics & Computational Disease Modelling, Freie Universität Berlin and University of Potsdam, Berlin, Potsdam, Germany; ³Department of Mathematics and Statistics, University of Reading, Whiteknights, UK; ⁴Department of Clinical Pharmacy and Biochemistry, Institute of Pharmacy, Freie Universität Berlin, Berlin, Germany. *Correspondence: Wilhelm Huisinga (huisinga@uni-potsdam.de)

process patient data collected over time in a batch (i.e., all at once), like Markov Chain Monte Carlo (MCMC), sampling importance resampling (SIR), and a normal approximation (NAP) to the posterior at the MAP estimate, but also particle filters (PFs) that allow for efficient sequential data processing. PFs are well-established in areas of application in which real-time predictions based on online/monitoring data are required, as in navigation, meteorology, and tracking.³⁻⁵

In the context of chemotherapy-induced neutropenia—the most frequent dose-limiting side effect for cytotoxic anticancer drugs with substantial decrease of neutrophil granulocytes and, thus, potentially life-threatening fever and infections⁶—we demonstrate the clear benefits of uncertainty quantification compared with purely MAP-based predictions (e.g., ref. 7) using the gold-standard model for neutropenia.⁸ Further, we compare the full Bayesian approaches regarding quality of uncertainty quantification and computational runtime for multiple cycle chemotherapy.⁹ Although MCMC, SIR, and PF all provide a reliable uncertainty quantification, the efficient data processing of the sequential approach will be clearly beneficial in a continuous monitoring context, where digital healthcare devices (e.g., wearables) allow patients to measure and report individual marker concentrations online.

METHODS

First, the statistical framework of MIPD in TDM is introduced, which is used throughout the different methods described below. Then, the considered clinical application scenarios are described along with the prior knowledge from literature.

Statistical framework

TDM in the context of MIPD builds on prior knowledge in the form of a structural, observational, covariate, and statistical model. In the sequel, TDM data are considered for a single individual and, therefore, there is no running index for individuals. The structural and observational models are given as:

$$\frac{dx}{dt}(t) = f(x(t); \theta, u), \quad x(t_0) = x_0(\theta) \quad (1)$$

$$h(t) = h(x(t), \theta) \quad (2)$$

with state vector $x = x(t)$ (including drug/biomarker concentrations), individual parameter values θ (e.g., volumes and clearances), and rates of change $f(x; \theta, u)$ of all state variables for a given input u (e.g., dose). Because typically only a part of the state variables is observed, the function h maps x to the observed quantities $h(x, \theta)$ (e.g., plasma drug or neutrophil concentration), including potential state-space transformations (e.g., log-transformed output). The initial conditions x_0 are defined by the pretreatment levels (e.g., baseline values). The covariate and statistical model link the patient-specific covariates “cov” and observations $(t_j, y_j)_{j=1, \dots, n}$ to the model predictions $h_j(\theta) = h(x(t_j), \theta)$, accounting for measurement errors and possibly model misspecification.

$$Y_{j|\theta=\theta} \sim \rho(\cdot | \theta; h_j(\theta), \Sigma), \quad j = 1, \dots, n \text{ (independent)} \quad (3)$$

$$\Theta \sim \rho_{\Theta}(\cdot; \theta^{\text{TV}}(\text{cov}), \Omega), \quad (4)$$

where $\theta^{\text{TV}}(\text{cov})$ denotes the typical hyperparameter values (TV) that might depend on covariates. The dot (“.”) in a probability distribution serves as placeholder for its argument. Often, $Y_{j|\theta=\theta} = h_j(\theta) + \varepsilon_j$ with $\varepsilon_j \sim_{iid} \mathcal{N}(0, \Sigma)$. Prior knowledge about the parameters is provided by population analyses of clinical studies, in which nonlinear mixed effect (NLME) approaches are used to estimate the functional relationship $\text{cov} \mapsto \theta^{\text{TV}}(\text{cov})$, and the parameters Σ and Ω .

The challenge in MIPD is to infer information on the individual parameter values θ of a patient based on his/her covariate values and measurements. Here, a Bayesian approach is highly beneficial: the unexplained interindividual variability in the population model (Eq. 4) defines the prior uncertainty about the individual parameter values. In this context, the hyperparameters (i.e., all parameters after the semicolon in Eq. 4) are assumed to be known (fixed). As a consequence, we drop them as well as the subscripts in the notation in the sequel. As a result, the likelihood at the individual level reads $\rho(\cdot | \theta) = \rho(\cdot | \theta; h_j(\theta), \Sigma)$ and the prior $\rho(\cdot) = \rho_{\Theta}(\cdot; \theta^{\text{TV}}(\text{cov}), \Omega)$. Then, assimilating measurements $y_{1:n} = (y_1, \dots, y_n)^T$ into the model based on Bayes’s formula:

$$\rho(\theta | y_{1:n}) = \frac{\rho(y_{1:n} | \theta) \cdot \rho(\theta)}{\rho(y_{1:n})} \quad (5)$$

allows to learn about individual parameter values from the data. The remaining uncertainty of parameter values is encoded in the posterior $\rho(\cdot | y_{1:n})$. Note that $\rho(y_{1:n} | \theta) = \rho(y_1 | \theta) \cdot \dots \cdot \rho(y_n | \theta)$ due to independence in Eq. 3. The denominator $\rho(y_{1:n})$ in Eq. 5 serves as a normalization factor, denoting the probability of the data. In contrast to MAP estimation, which summarizes the posterior by its mode, Bayesian DA approaches rely on sample approximations of the posterior

$$\rho(\theta | y_{1:n}) \approx \sum_{s=1}^S w_n^{(s)} \delta_{\theta_n^{(s)}}(\theta) \quad (6)$$

based on a sample

$$\mathbb{S}_n := \left\{ \left(\theta_n^{(s)}, w_n^{(s)} \right), \quad s = 1, \dots, S \right\},$$

with sample parameters $\theta_n^{(s)}$ of the posterior (Section 2.5 in ref. 10), weights $w_n^{(s)}$, which sum to one (i.e., $\sum_{s=1}^S w_n^{(s)} = 1$ and point masses $\delta_{\theta_n^{(s)}}(\theta)$ at $\theta_n^{(s)}$). If $w_n^{(s)} = 1/S$ for all s , the sample is called unweighted. Based on the posterior sample, we may also approximate quantities of interest in the observable space by solving Eqs. 1 and 2 for all elements in \mathbb{S}_n . This also serves as the basis for credible intervals (Cris); applying subsequently the residual error model Eq. 3, the basis for prediction intervals.

Because direct sampling from the posterior is, in general, not possible, alternative approaches (described below) need to be used to generate a sample of the posterior. For a detailed description, see **Section S3–Section S7**.

MAP estimation

MAP estimation approximates the mode of the posterior distribution (i.e., the most probable parameter values given patient-specific measurements $y_{1:n}$):

$$\hat{\theta}_n^{\text{MAP}} = \arg \max_{\theta} \rho(\theta | y_{1:n}). \quad (7)$$

The MAP estimate $\hat{\theta}_n^{\text{MAP}}$ is a one-point summary of the posterior distribution, without quantification of the associated uncertainty.

NAP

To overcome the one-point summary limitation of the MAP estimate, the posterior $p(\cdot | y_{1:n})$ may be approximated locally by a normal distribution located at the MAP estimate (Section 4.1 in ref. 11).

$$p(\cdot | y_{1:n}) \approx \mathcal{N} \left(\hat{\theta}_n^{\text{MAP}}, \mathcal{I}^{-1} \left(\hat{\theta}_n^{\text{MAP}} \right) \right), \quad (8)$$

where \mathcal{I} denotes the total observed Fisher information matrix (Section 2.5 in ref. 12).

$$\mathcal{I}(\theta) := \mathcal{I}^{\text{post}}(\theta) = -\frac{d^2}{d\theta^2} \log \rho(\theta | y_{1:n})$$

of the posterior. The uncertainty in the parameters is then propagated to the observables by first sampling from the normal distribution in Eq. 8 and then solving the structural model for each sample.¹³ Alternatively to this sampling-based approach, the Delta method (Section 5.5 in ref. 14) could be used, see also **Section S4** and **Figure S2**.

SIR

The SIR algorithm is a full Bayesian approach. It generates an unweighted sample \mathbb{S}_n from the posterior $p(\theta | y_{1:n})$ based on a sample from a so-called importance distribution G , from which samples can easily be generated (Section 10.4 in ref. 11). SIR proceeds in three steps. S-Step: Sampling from the importance distribution G resulting in a sample $\tilde{\mathbb{S}}_n$. We considered the prior as importance distribution, assuming that the patient under consideration is sufficiently well-represented by the clinical patient populations given by the prior. I-Step: Each sample point $\tilde{\theta}_n^{(s)} \in \tilde{\mathbb{S}}_n$ is assigned an importance weight $\tilde{w}_n^{(s)} = p(y_{1:n} | \tilde{\theta}_n^{(s)})$ given by the likelihood (in case G is the prior). R-Step: After normalization of the weights $w_n^{(s)} = \tilde{w}_n^{(s)} / \sum_s \tilde{w}_n^{(s)}$, a resampling is performed: S unweighted samples $\theta_n^{(s)}$ are drawn from $\tilde{\theta}_n^{(s)}$ according to weights $w_n^{(s)}$.

Note that once a new data point y_{n+1} becomes available, the SIR algorithm does not simply update the present sample points $\theta_n^{(s)}$ in \mathbb{S}_n , but re-performs all three steps based on the updated posterior $p(\cdot | y_{1:n+1})$ to determine \mathbb{S}_{n+1} .

MCMC

A popular alternative to SIR in Bayesian inference are MCMC methods with a wide range of different algorithms.¹⁵ MCMC generates an unweighted sample \mathbb{S}_n from the

posterior by means of a Markov chain (Section 11 in ref. 11). MCMC comprises two steps to generate sample points $\theta_n^{(s)}$: a proposal step (generating a potential new sample point θ^*) and an acceptance step (accepting or rejecting θ^* as a new sample point). The challenge in MCMC is to design application-specific proposal distributions.

In TDM, MCMC was previously considered with the prior as fixed proposal distribution (independence sampler) for sparse patient monitoring data.¹⁶ We observed, however, large rejection rates with increasing number n of data points, because the posterior becomes narrower, see **Section S6**. To counteract large rejection rates, we used an adaptive Metropolis-Hastings sampler with log-normally distributed proposal distribution (see **Section S6** and **Figure S4** for details).

PF

In contrast to SIR and MCMC, which process data in a batch, PF constitutes a sequential approach to DA, see^{3,17,18} for a detailed introduction. Given a weighted sample \mathbb{S}_n of the posterior $p(\cdot | y_{1:n})$ and a new data point y_{n+1} , PF generates a weighted sample \mathbb{S}_{n+1} of the posterior $p(\cdot | y_{1:n+1})$ by updating \mathbb{S}_n using a sequential version of Bayes's formula, as follows:

$$p(\theta | y_{1:n+1}) \propto p(y_{n+1} | \theta) \cdot p(\theta | y_{1:n}). \quad (9)$$

For $n=0$, the distribution $p(\theta | y_{1:0})$ is identical to the prior $p(\theta)$ (Section 3 in ref. 10). Note that $p(y_{n+1} | \theta) = p(y_{n+1} | y_{1:n}, \theta)$ due to independence in Eq. 3. Analogously to the I-step in SIR, the weights $w_n^{(s)}$ are updated proportional to the (local) likelihood: $w_{n+1}^{(s)} \propto p(y_{n+1} | \theta) \cdot w_n^{(s)}$ involving, however, only the new data point y_{n+1} . As in SIR and MCMC, evaluation of the likelihood involves solving the structural model (1). Because the structural model is deterministic, one may either solve Eq. 1 with initial condition $x_0(\theta)$ for the total timespan $[t_0, t_{n+1}]$, or with initial condition $x_n(\theta)$ for the incremental timespan $[t_n, t_{n+1}]$. The latter approach requires to store for each sample point $(\theta_n^{(s)}, w_n^{(s)})$ also the corresponding state $x_n^{(s)}$ at time t_n , because typically the structural model cannot be solved analytically. The incremental approach makes use of the Markov property that the future state is independent of the past when the present state is known.

The resulting triple $(\theta_n^{(s)}, w_n^{(s)}, x_n^{(s)})$ is called a particle. In our setting, the ensemble of particles can be interpreted as the state of a population of virtual individuals at time t_n , whose "diversity" represents the uncertainty about the state/parameters of the patient at time t_n , given the individual measurements $y_{1:n}$. The posterior $p(\cdot | y_{1:n})$ obtained by n sequential update steps in Eq. 9 is mathematically identical to the posterior obtained in Eq. 5 by assimilating all data $y_{1:n}$ in a batch (Section 3.3.3 in ref. 19). However, the sequential update is much more efficient as it involves a reduced integration time span.

In contrast to SIR, PF does not perform resampling by default. Only if too many samples carry an almost negligible weight and the total weight is limited to only a few samples (weight degeneracy), a resampling is performed. We used a criterion based on the effective sample size

$$S_{\text{eff}}(t_n) = \frac{1}{\sum_{s=1}^S (w_n^{(s)})^2}$$

to decide whether to resample. Starting initially with uniform weights $w_0^{(s)} = 1/S$ with $S_{\text{eff}}(t_0) = S$, resampling was carried out once $S_{\text{eff}} < S/2$ (effective ensemble size half of the initial ensemble size). If resampling is performed, it is followed by a so-called rejuvenation step³ to prevent sample impoverishment by fixation to limited parameter values:

$$\theta_n^{(s)} = \tilde{\theta}_n^{(s)} + \xi_n^{(s)}, \quad \text{with } \xi_n^{(s)} \sim_{iid} \mathcal{N}\left(0, \tau \cdot \left| \tilde{\theta}_n^{(s)} \right|\right).$$

with rejuvenation parameter τ , where $\tilde{\theta}_n^{(s)}$ denotes the re-sampled parameters. These two steps, resampling and rejuvenation, ensure that the weighted sample S_n adequately represents areas of posterior probability, see also **Figure S5**.

Biomarker data during chemotherapy for single/multiple cycle simulation studies

Two simulation studies (see below) were performed in MATLAB R2017b/2018b, see supplementary information files SCode, to analyze the approaches regarding their suitability to support MIPD.

For the single cycle study with docetaxel (100 mg/m², 1 hour infusion), we used the NLME model in ref. 20. It is based on the well-known pharmacodynamic model in ref. 8 (**Figure S6**) and describes the effect of a single dose of the anticancer drug docetaxel based on monitoring neutrophil counts. Important model parameters are the drug effect parameter “Slope” and the pretreatment baseline neutrophil concentration “Circ0.” For inference, neutrophil concentrations were considered on a log-scale at time points $t=0,3, \dots, 21$ days postdose, see **Section S8** for full details. This simulation study aims to demonstrate the limitations of MAP estimation for a model frequently used in MIPD for TDM.^{7,21} Because

recursive data processing and decision-support gain in relevance for long-term monitoring, we performed a simulation study for multiple cycle therapy with paclitaxel using the NLME model in ref. 9. It describes the effect of the anticancer drug paclitaxel (200 mg/m², 3-hour infusion) over six cycles of 3 weeks each, corresponding to treatment arm A of the CEPAC/TDM study.²² In ref. 9, an aggravation of neutropenia over subsequent treatment cycles is accounted for by bone marrow exhaustion⁹ (**Figure S6**). The model includes interoccasion variability (IOV) on pharmacokinetic parameters describing the variability between cycles within one patient. Therefore, the parameter values comprise the interindividual parameters (θ_{iIV}) and a parameter for each occasion (θ_{iOV}). As a consequence, the size of θ increases with every occasion/cycle, $\theta = (\theta_{iIV}, \theta_{iOV}^1, \dots, \theta_{iOV}^{n_c})$, where n_c denotes the number of cycles, see **Section S9** for details. Neutrophil counts were assumed to be monitored every third day. We were interested in a setting where data become available sequentially (one-by-one). To this end, neutrophil count data were simulated for a virtual patient using Eqs. 3 and 4 and the corresponding model. Then the individual parameter values were inferred based on the simulated neutrophil count data available up to a certain time point, using the same model. For the statistical analysis, this procedure was repeated for $N = 100$ virtual patients (with covariate characteristics mirroring the real study population underlying the NLME model).

Key characteristics for decision support in cytotoxic chemotherapy

We investigated different characteristics $T(\cdot)$ of the neutropenia time-course related to risk and recovery.⁷ Depending on the nadir (i.e., minimal neutrophil concentration), different grades of neutropenia are distinguished, see **Figure 1**.

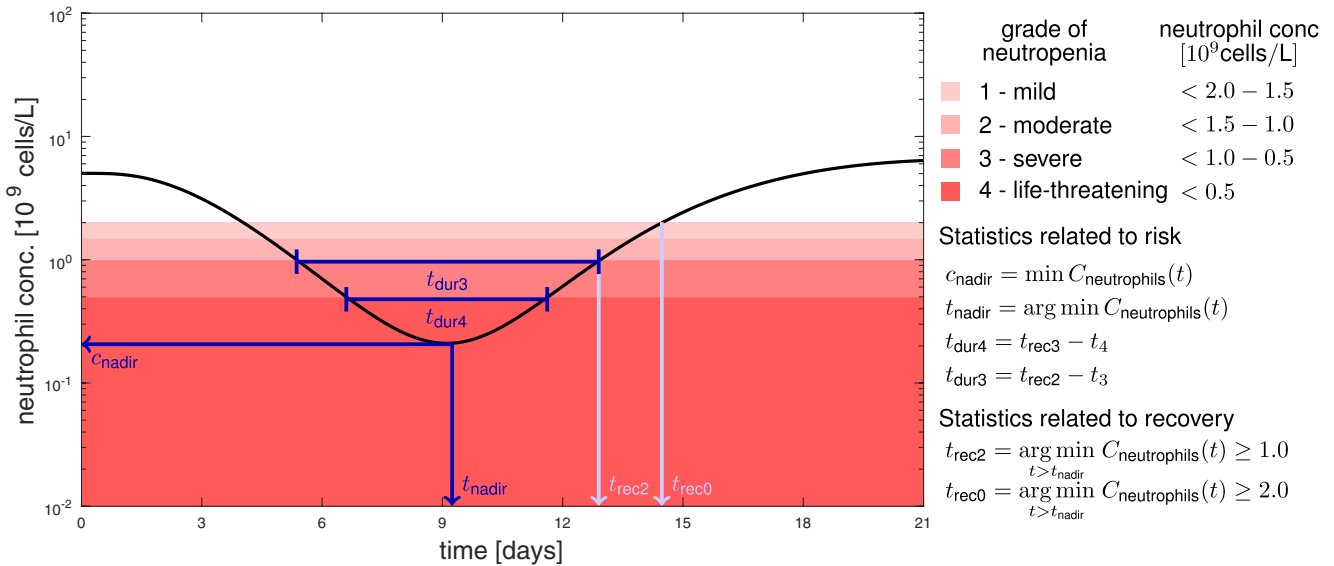


Figure 1 Key characteristics for decision making in cytotoxic chemotherapy related to risk (dark blue) and recovery (light blue) of neutropenia. Neutropenia grades are defined according to the Common Terminology Criteria for Adverse Events.³¹ Note that the shades of red are related to the increasing toxicity; however, grade 0 (white) over the whole cycle is associated with ineffective treatment. As key statistics for decision support we consider the lowest neutrophil concentration (c_{nadir}), the time at which the nadir is reached (t_{nadir}), the duration of neutropenia grade 3 and grade 4 (t_{dur3} and t_{dur4} , respectively), as well as the times until recovery to neutropenia grade 2 and 0 (t_{rec2} and t_{rec0} , respectively).

Neutropenia grade 4 is dose-limiting as this severe reduction in neutrophils exposes patients to life-threatening infections. On the contrary, neutropenia grade 0 is also undesired as it is associated with a worse overall treatment outcome.²³ The time t_{nadir} at which the nadir is reached is important for time management of intervention. We considered the patient out of risk at time $t_{\text{rec}2}$ when neutropenia grade 2 is reached post nadir. For the initiation of the next treatment cycle, the recovery time $t_{\text{rec}0}$ to grade 0 is important. In addition, risk is also related to the duration $t_{\text{dur}3}$ and $t_{\text{dur}4}$ of an individual being in grade 3 and 4 neutropenia, respectively.

Workflow in Bayesian forecasting

In full Bayesian forecasting, uncertainty is quantified on the parameter level and subsequently propagated to the observable level, possibly summarized for some key quantities of interest, see **Figure 2**. Prior to observing patient-specific data, the parameter uncertainty is characterized by the prior (cmp. Eq. 4). It allows to make *a priori* predictions of the neutropenia time course and its uncertainty in form of a $(1-\alpha)$ -confidence interval. In addition, *a priori* predictions for quantities of interest can be derived (e.g., the neutropenia grade (**Figure 2**, left column)). Once patient-specific data are assimilated into the Bayesian model, the remaining uncertainty on the parameter values is characterized by the posterior, allowing to also update the uncertainty in the observable space (CrIs) and the quantities of interest (**Figure 2**, middle column).

Forward uncertainty propagation corresponds to transforming a probability distribution (prior or posterior) under a (possibly nonlinear) mapping $T(\cdot)$, resulting in a transformed quantity $\psi = T(\theta)$. For illustration, we assume the one-dimensional case with strictly increasing T and $\theta = T^{-1}(\psi)$. Then the posterior in terms of ψ is given by Section 1 in ref. 24.

$$p_{\Psi}(\psi|y_{1:n}) = p_{\Theta}(\theta|y_{1:n}) \cdot \frac{dT^{-1}(\psi)}{d\psi}, \quad (10)$$

which is approximated in sampling-based approaches (cf. Eq. 6) by

$$\hat{p}_{\Psi}(\psi|y_{1:n}) = \sum_{s=1}^S w_n^{(s)} \delta_{\psi^{(s)}}(\psi),$$

with $\psi^{(s)} = T(\theta^{(s)})$. This allows the computation of any desired summary statistic (e.g., posterior expectation or quantiles). MAP estimation, in contrast, characterizes the posterior by a single value and allows only to make a single MAP-based prediction by mapping the MAP estimate $\hat{\theta}^{\text{MAP}}$ to the quantity of interest $T(\hat{\theta}^{\text{MAP}})$, lacking crucial information on its uncertainty (**Figure 2**, right column). Importantly, for nonlinear T , this does *not* result in the most probable outcome, due to the Jacobian factor $\frac{d\theta}{d\psi} = \frac{dT^{-1}(\psi)}{d\psi}$ in Eq. 10^{25,26}: The most probable outcome is defined as the outcome with maximum posterior probability.

$$\hat{\psi}_n^{\text{MAP}} = \arg \max_{\psi} p_{\Psi}(\psi|y_{1:n}), \quad (11)$$

which satisfies (assuming for illustration that T is strictly increasing).

$$\begin{aligned} 0 &= \frac{d}{d\psi} p_{\Psi}(\psi|y_{1:n}) \stackrel{\text{Eq. 10}}{=} \frac{d}{d\psi} \left[p_{\Theta}(T^{-1}(\psi)|y_{1:n}) \cdot \frac{dT^{-1}(\psi)}{d\psi} \right] \\ &= \frac{d}{d\theta} p_{\Theta}(T^{-1}(\psi)|y_{1:n}) \cdot \left(\frac{dT^{-1}(\psi)}{d\psi} \right)^2 \\ &\quad + p_{\Theta}(T^{-1}(\psi)|y_{1:n}) \cdot \frac{d^2 T^{-1}(\psi)}{d\psi^2}. \end{aligned} \quad (12)$$

For the transformed MAP estimate, $\psi = T(\hat{\theta}^{\text{MAP}})$, the first term in the right hand side of Eq. 12 is zero, because its first factor vanishes by definition. The second term, however, is non-zero, because both its factors are non-zero for nonlinear T . Therefore, the transformed MAP estimate does not satisfy the condition for the mode of the transformed posterior probability and, hence, $T(\hat{\theta}^{\text{MAP}}) \neq \hat{\psi}^{\text{MAP}}$.

Method comparison

For all sampling-based methods (NAP, SIR, MCMC, and PF) we used a sample of size $S = 10^3$. Because the posterior is analytically intractable, an extensive sample of size $S = 10^6$ was used as reference, called “full Bayes (reference)” in the sequel, which was generated by SIR and cross-checked with MCMC (see **Figure S7**, because these approaches are exact in the limit $S \rightarrow \infty$). As a statistical measure for the quality of uncertainty quantification, we considered the Hellinger distance

$$H(\hat{P}, P^{\text{ref}}) = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^b \left(\sqrt{\hat{p}_i} - \sqrt{p_i^{\text{ref}}} \right)^2}, \quad (13)$$

which measures the difference between the discrete sampling-based *a posteriori* probability distribution $\hat{P} = (\hat{p}_1, \dots, \hat{p}_b)$ and the reference solution $P^{\text{ref}} = (p_1^{\text{ref}}, \dots, p_b^{\text{ref}})$ generated with SIR $S = 10^6$ for b fixed bins.

RESULTS

First, we show the limitations of MAP estimation for MIPD and how full Bayesian approaches can overcome these limitations (using SIR with $S = 10^6$ as reference). Next, we compare different full Bayesian approaches with reduced sample sizes regarding accuracy and computational efficiency.

Unfavorable properties of MAP-based predictions

The first example of decision support in individualized chemotherapy uses the most frequently used model of neutropenia.⁸ The MAP estimate $\hat{\theta}_n^{\text{MAP}}$ is derived from the parameter posterior $p(\cdot|y_{1:n})$ given experimental data $y_{1:n} = (y_1, \dots, y_n)^T$, see Eq. 7. In the context of TDM, it is used to predict the future time course $x(t; \hat{\theta}_n^{\text{MAP}})$ of the patient and thereon based observables. In mathematical terms, $\hat{\theta}_n^{\text{MAP}}$ is mapped to some quantity of interest $T(\hat{\theta}_n^{\text{MAP}})$ (e.g., the nadir concentration). As pharmacometric models are generally nonlinear, this does, however, not result in the most probable outcome (see also paragraph preceding Eq. 12 in the Methods). This is due to the fact that first determining the MAP estimate and then

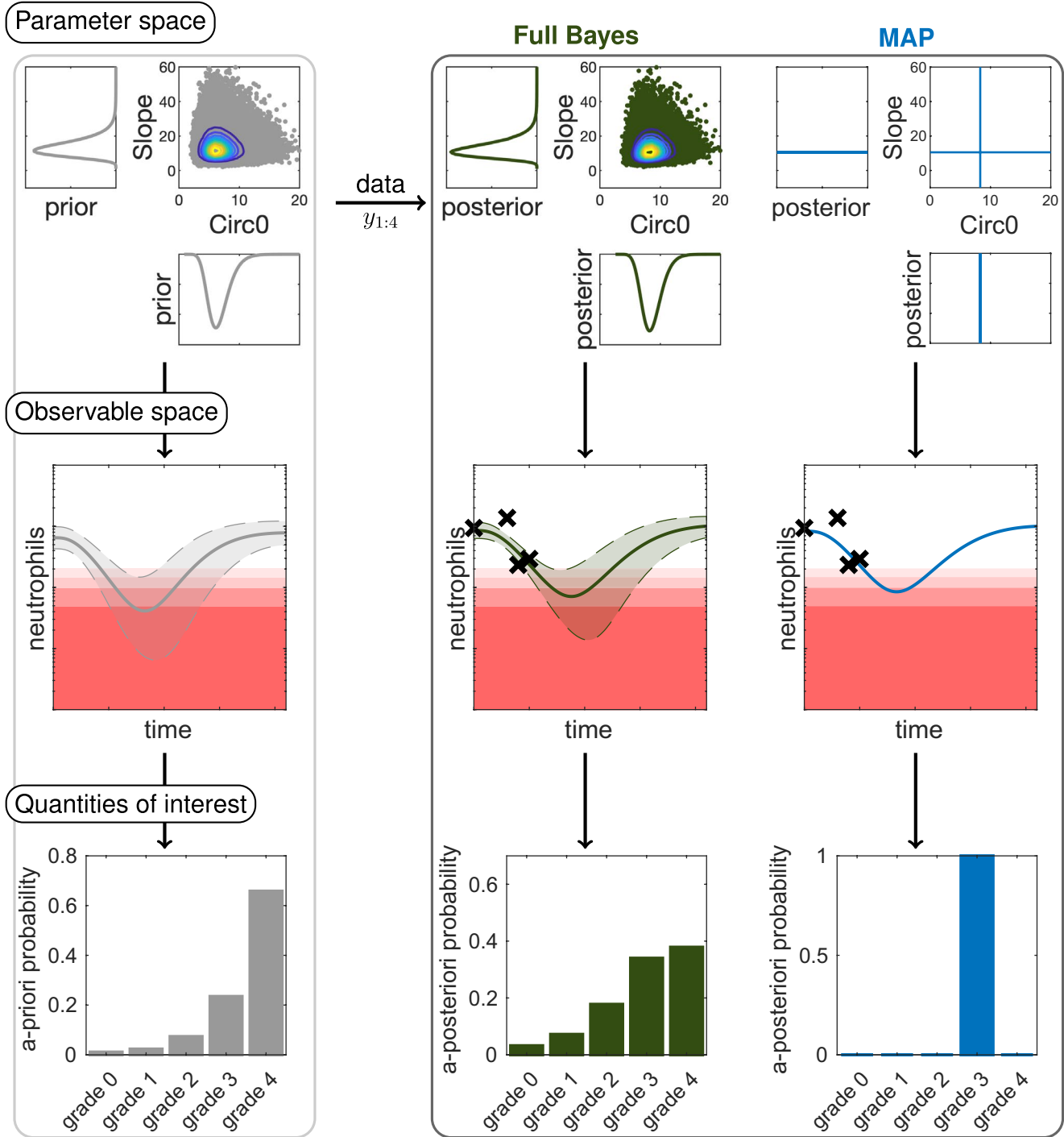


Figure 2 Overview of the workflow in model-informed precision dosing comparing full Bayesian inference to maximum *a posteriori* (MAP)-based prediction. In full Bayesian inference, uncertainties in the parameter values are propagated to uncertainties in the observable space and quantities of interest. The posterior $p(\theta|y_{1:m})$ is displayed for the parameters “Slope” (drug effect parameter) and “Circ0” (pretreatment neutrophil concentration). For the prior and full Bayes (reference) approach (sampling importance resampling with $S=10^6$) samples (dots) from the distributions are shown with contour levels. In the observable space, the point estimates (solid lines) are displayed with the central 90% confidence interval or credible intervals (dashed lines and shaded area) along with the therapeutic drug/biomarker monitoring data (crosses). The *a priori/a posteriori* probabilities are calculated for the neutropenia grades (grade 0–4). Note that y_1 corresponds to the measurement of baseline neutrophil counts (“Circ0”) and is taken into account in the posterior.

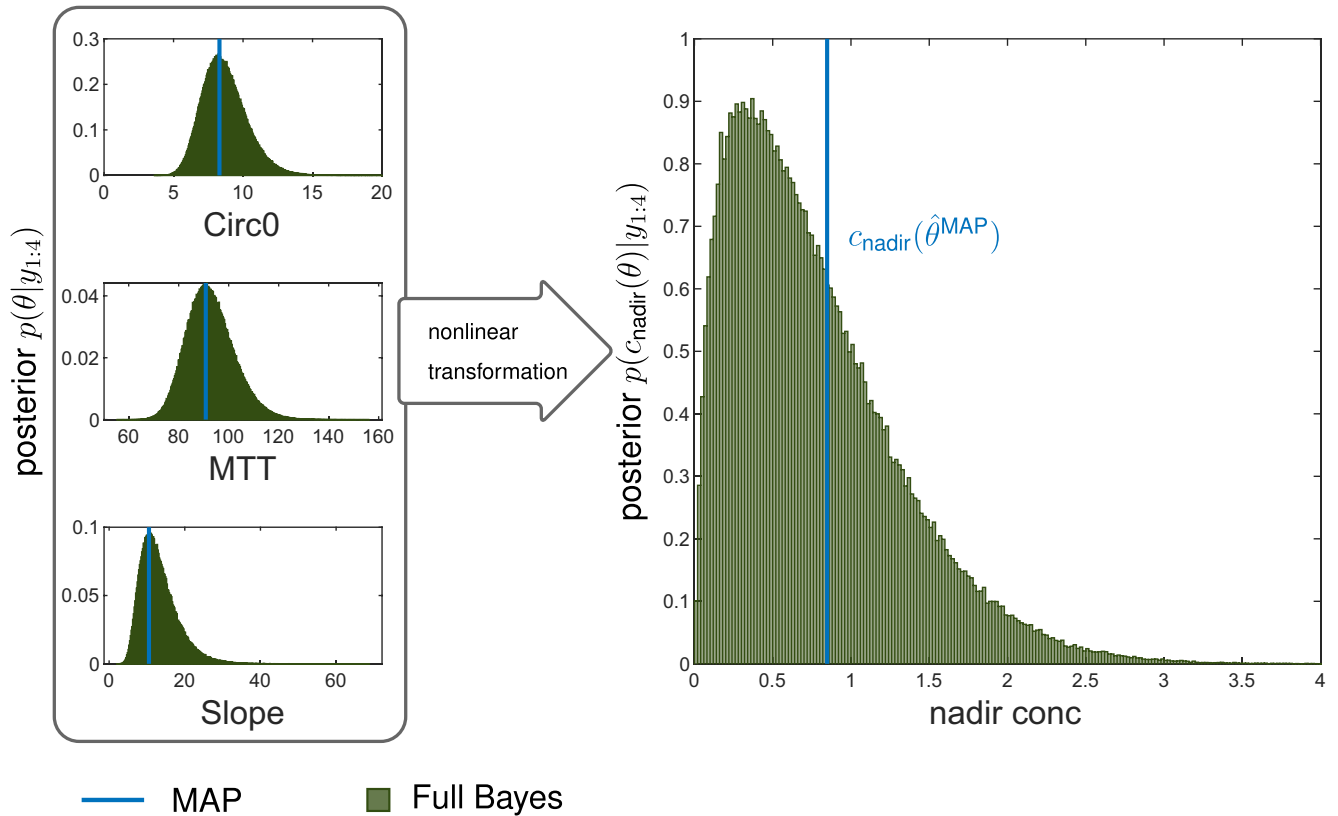


Figure 3 Maximum *a posteriori* (MAP)-based predicted nadir concentration is not the most (*a posteriori*) probable nadir concentration. We considered the single cycle study docetaxel with four observed data points $y_{1:4}$ and forecasted the nadir concentration based on the posterior $p(\theta|y_{1:4})$ using MAP estimation and the full Bayes (reference) approach (sampling importance resampling with $S=10^6$), where θ comprises the parameters baseline neutrophil counts (Circ0), mean transit time (MTT), and “Slope” (drug effect). The MAP estimate of the parameters coincides with the mode of the posterior distribution of the parameters (left panel), however, the mode is not preserved under nonlinear transformation (see text). Therefore, $c_{\text{nadir}}(\hat{\theta}^{\text{MAP}})$ with $c_{\text{nadir}}(\cdot)$ denoting some observable $T(\cdot)$ does not equal the mode of the *a posteriori* probability $p_{T(\theta)}(\cdot|y_{1:n})$ of the nadir concentration. Please also refer to Figure S1 for further illustration and analysis.

applying a nonlinear mapping is, in general, different from first applying the mapping to the full parameter posterior and then determining its MAP estimate: $T(\hat{\theta}^{\text{MAP}}) \neq \widehat{T(\theta)}^{\text{MAP}}$, see **Figure 3** for an illustration with $T(\theta) = c_{\text{nadir}}(\theta)$ and Figure S1 for more details.

Thus, MAP-based estimation lacks both, a measure of uncertainty and the feature to predict the most probable observation/quantity of interest. In addition, relevant outcomes, such as the risk of grade 4 neutropenia cannot be evaluated from the point estimate alone. MAP-based estimation, therefore, provides a biased basis for clinical decision making. In contrast, full Bayesian inference provides access to the full posterior distribution of the parameters and correctly transforms uncertainties forward to the observables and quantities of interest, allowing to compute any desired summary statistic and relevant risks (Section 5.2 in ref. 19).

Uncertainty quantifications for more comprehensive, differentiated understanding, and better informed decision making

The first scenario served to demonstrate the limitations of MAP-based estimations for the gold-standard model,⁸

however, the model does not account for the observed cumulative neutropenia over multiple cycles. Therefore, we considered for dose adaptations a model accounting for bone marrow exhaustion over multiple cycles,⁹ see paragraph about the multiple cycle study paclitaxel. We exemplarily considered the dose selection for the third treatment cycle based on prior information and patient-specific measurements during the first two cycles. The patient-specific data together with the full Bayes (reference) model fit and prediction are shown in **Figure 4a**, see also **Figures S8** and **S9**. The Crls (dashed) and prediction intervals (dotted) show the uncertainty about the “state of the patient,” without and with measurement errors, respectively.

For optimizing the dose of the third cycle, different dosing scenarios were considered: the standard dose and a -15%, -30%, and +10% adapted dose. **Figure 4b** shows the probability of the predicted grades of the third cycle for each dose. To find an effective and safe dose, the risk of being ineffective (neutropenia grade 0) should be minimized jointly with the risk of being unsafe (neutropenia grade 4). For illustration in **Figure 4b**, the dashed horizontal lines indicate a 10% and 5% level of being ineffective and unsafe,

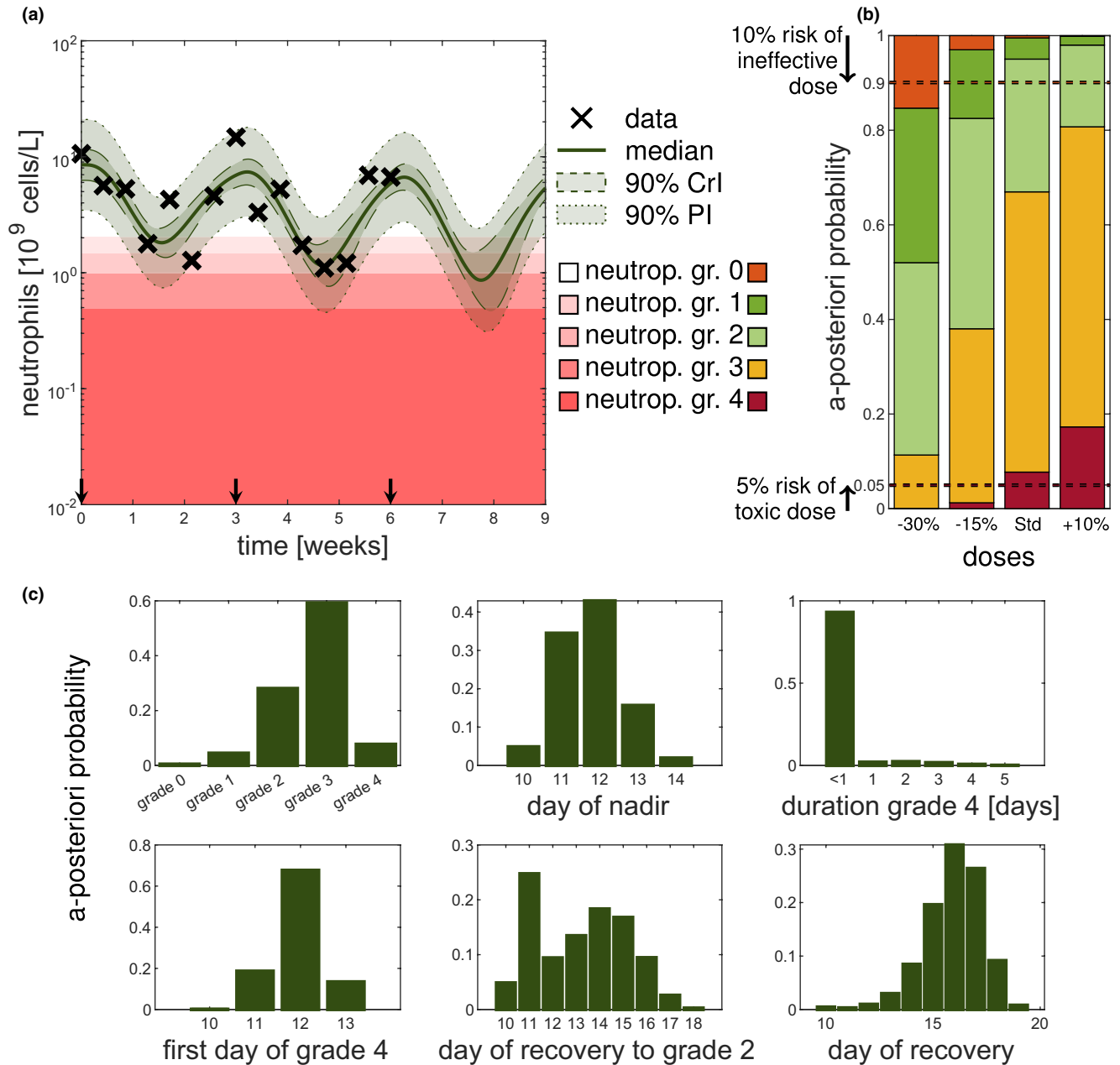


Figure 4 Uncertainty quantification by full Bayesian methods gives important information for therapy dosing selection. The scenario described in multiple cycle study paclitaxel is used and the results are shown for the full Bayes (reference) solution with sampling importance resampling (SIR) using $S = 10^6$ samples. (a) Forecasting the third cycle for different doses based on the patient's covariates and measurements of the first two cycles. (b) Full Bayesian inference allows for probabilistic statements of the different grades. Color coding of neutropenia grades shows trade-off between efficacy and toxicity. No toxicity (grade 0) is associated with poorer treatment outcome (orange) but severe neutropenia (grade 3 and 4) is also not desired (yellow and red). (c) *A posteriori* probabilities of quantities of interest for the third cycle based on the posterior at the end of second cycle (week 6) for the standard dose. Statistics, such as day of grade 4, were computed given that grade 4 is reached. Note for all displayed forecasts the full Bayes (reference) approach (SIR with $S = 10^6$) was used. CrI, credible interval; PI, prediction interval.

respectively. The standard dose and the increased dose have a risk of toxicity larger than 5% (lower horizontal line). A decrease in dose also leads to an increased risk of an ineffective dose (upper horizontal line). The 15% reduced dose is with 96% probability safe and efficacious (grades 13), with 3% probability ineffective (grade 0) and with 1% probability unacceptably toxic (grade 4). If grade 3 is also to

be avoided, the 30% reduction would be preferable, as it is with 74% probability safe and efficacious (grades 12), with 15% probability ineffective (grade 0) and with 11% probability toxic (grades 34). Thus, the choice of an optimal dose might depend on how priority is given to the risk of inefficacy and toxicity. As both risks are described by the tails of the posterior distribution, a point estimate is not able to

adequately capture them. The MAP-based predicted grades were: grade 2 (standard dose and +10% dose), grade 1 (−15% dose); and grade 0 (−30% dose), which do not only make it difficult to distinguish between some doses, but also do not reflect the true most probable grades.

Posterior-based predictions of important statistics related to the neutropenia time course can help to answer questions like “How probable is it that the patient will suffer from grade 4 neutropenia?” or “How probable is it that the patient will recover in time for the next scheduled dose so that the therapy can be continued as planned?” To answer such questions, **Figure 4c** shows important predicted quantities of interest, illustrated for the standard dose in cycle 3. We inferred that the risk of grade 4 neutropenia is 8%, and if the patient were to reach grade 4, it would be most probable (68%) on day 12. The probability that the patient’s duration in grade 4 is a day or longer is very small (< 7%). As the probability to not have been recovered until day 21 is negligible, the administration can remain scheduled on day 21 for cycle 4. Therefore, uncertainty quantification improves the decision-making process by quantifying the *a posteriori* probabilities of relevant risks and quantities of interests. Repeating the above analysis for different doses, therefore, allows for an improved distinction between dose adjustments.

Approximation accuracies comparable across different full Bayesian approaches

We next compared different established methods for uncertainty quantification with regard to their approximation accuracy. To this end, posterior inference was investigated for a patient at day 5 of the first cycle (**Figure 5**). Whereas the marginal posterior distribution for the parameter “Circ0” (pretreatment neutrophil concentration) is close to a normal distribution, the marginal posterior for the drug effect parameter (“Slope”) is closer to a log-normal distribution. Accordingly, the NAP is rather reasonable for “Circ0,” but is questionable for the “Slope” parameter. In addition, sampling from the normal distribution can lead to unrealistic (negative) parameter values (**Figure 5a**). In addition, NAP very clearly underestimated the patient’s risk to reach grade 4 neutropenia (**Figure 5b,c**), which could possibly lead to a fatal dose selection. Considering a Student’s t distribution instead of the normal approximation, as in ref. 13, did not lead to an adequate improvement (**Figure S3**). Consequently, the NAP approach can result in overoptimistic, overpessimistic, and unrealistic predictions. In contrast, the full Bayesian methods (SIR, MCMC, and PF) adequately represent the tails and respect the positivity constraint of parameter values. The resulting Crls are comparable to the reference Crls. For illustration, **Figure 6a** shows the approximation error for the predicted probability of neutropenia grades, measured in the Hellinger distance (see Eq. 13). Overall, SIR and PF showed the best approximation, whereas NAP resulted in the largest errors.

Sequential DA processes patient data most efficiently

The need for real-time inference algorithms is increasing with the possibilities to more frequently collect

patient-specific data (online collection) during treatment. Sequential DA methods (e.g., PF) provide an efficient framework for real-time data processing. At any time, all information (including associated uncertainty) is present in a collection of particles that can be interpreted as representing the current state and associated uncertainty of a patient via a virtual population. With a new datum, this information is updated. Approaches that rely on batch data analysis (i.e., MAP, SIR, and MCMC), need to redo the inference from scratch. This has impact on the computational effort as the number of data points increases. **Figure 6b** shows a comparison of the computational cost to assimilate an additional data point. All approaches show some kind of increase in effort every 21 days—due to the IOV on some parameters. Clearly, PF shows lowest and almost constant costs, whereas for batch mode approaches computational costs increase over time due to an increasing number of parameters (one additional parameter for every cycle due to the IOV, see paragraph about the multiple cycle study with paclitaxel) and an increasing integration time span to determine the likelihood. This could become computationally expensive in view of long-term treatments and higher time resolution of data points provided by new digital healthcare devices. **Figure 6c** summarizes the features of the different inference approaches. Note that all sampling-based approaches can be accelerated by parallel computing. In summary, it was found that PF processes patient monitoring most efficiently and facilitated the handling of IOV because only the IOV parameter of the current occasion needs to be considered.

DISCUSSION

In the context of chemotherapy-induced neutropenia, we illustrated the severe drawbacks of MAP-based approaches for forecasting and thereon based decision making. A prediction based on the MAP estimate does neither correspond to the most probable outcome nor does it allow to quantify relevant risks as the uncertainties are not quantified. Both are highly undesirable characteristics and make MAP-based inference difficult to interpret in a TDM setting. A normal approximation of the posterior at the MAP estimate is no alternative, as it retains the same point estimate and proved to be unsuitable in case of skewed parameter distributions. We demonstrated that full Bayesian approaches, like SIR, MCMC, or PF, provide accurate approximations to the posterior distribution, enabling comprehensive uncertainty quantification of the quantities of interest (e.g., nadir concentration). Among the three considered approaches, PF is a sequential approach that is beneficial in a more continuous monitoring context.

Uncertainty quantification in TDM is scarce. In ref. 27 the SIR algorithm was previously used in the TDM setting to construct Crls using a Student’s t distribution located at the MAP estimate as importance function. A sequential approach in the context of MAP estimation is discussed in ref. 28 with a moving estimation horizon (window of data points that are considered). A sequential DA approach has been investigated previously for glucose forecasting,^{29,30}

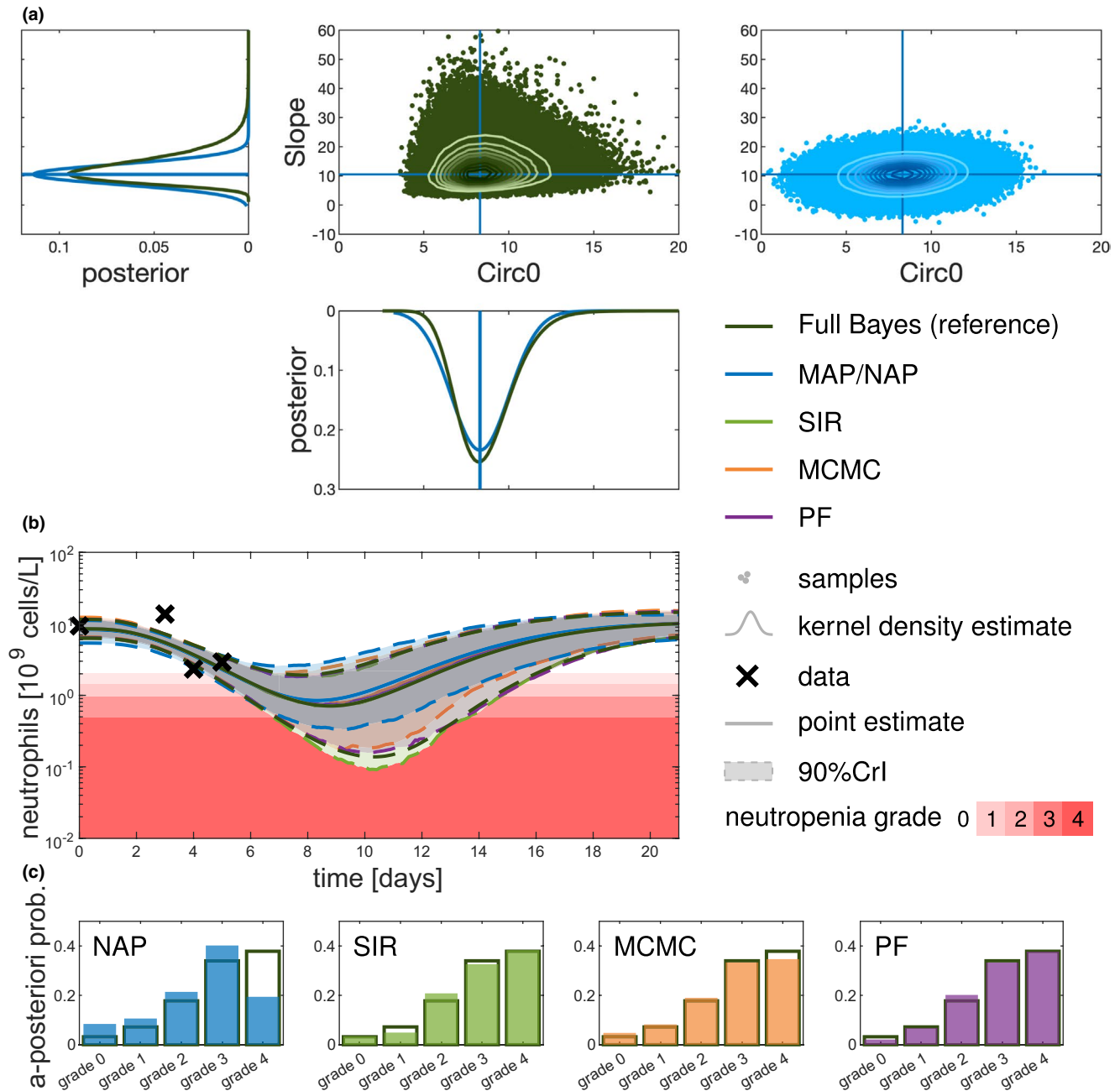


Figure 5 Comparison of uncertainty quantification at the level of parameters, observables, and quantities of interest. Exemplary comparison of the different methods for one patient after having observed four data points up to day 5. **(a)** The posterior is shown for parameters “Slope” and “Circ0” showing the kernel density estimates of the sampling distribution univariately and as scatter plots for the bivariate sampling distributions with contour plots for the full Bayesian approach (full Bayes (reference), sampling importance resampling (SIR) with $S = 10^6$) and the normal approximation located at the maximum *a posteriori* (MAP) estimate. **(b)** On the level of the observable (neutrophil concentration) the point estimates (median or MAP) are displayed along with the 90% credible intervals (CrIs). For illustration purposes, the prediction intervals are not shown here. **(c)** The forecasted *a posteriori* probability of the different neutropenia grades (0–4) is shown for the different approximations (filled bars) in comparison with the full Bayes (reference; black outlined bars). MCMC, Markov Chain Monte Carlo; NAP, normal approximation; PF, particle filter.

yet, not in combination with an NLME modeling framework and without decision-support statistics. A systematic comparison of approaches, as presented herein, is lacking.

In this study, particle filtering is applied in TDM within an NLME modeling framework to represent the current patient status via an uncertainty ensemble. A challenge in the

application of PF is the potential for weight degeneracy (i.e., a gradual separation into a few large and many very small weights). A rejuvenation approach (as applied in this study) resolves this problem, but requires specification of an additional parameter (magnitude of the rejuvenation). A too large value might result in an artificially increased uncertainty, whereas

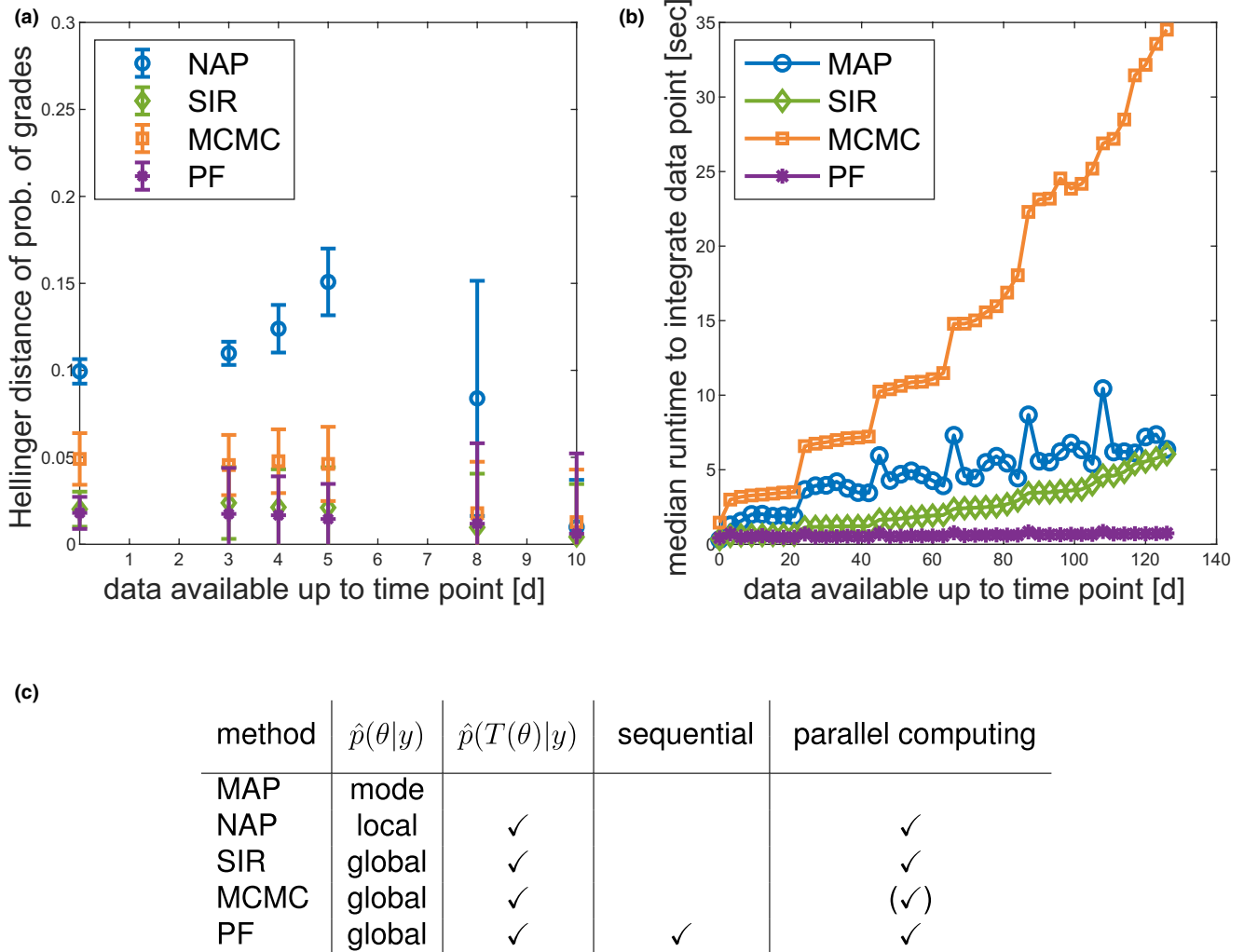


Figure 6 Comparison of methods regarding important aspects for model-informed precision dosing. (a) Approximation error (measured as Hellinger distance) of the probability of neutropenia grades (“Single cycle study docetaxel”). (b) Qualitative runtime comparison to sample from the parameter posterior. Median of $N=100$ repeated analyses (“Multiple cycle study paclitaxel”). (c) Comparison of method properties. For Markov Chain Monte Carlo (MCMC) several chains could be run in parallel, however, in this study, only one chain was considered. MAP, maximum *a posteriori*; NAP, normal approximation; PF, particle filter; SIR, sampling importance resampling.

a too small value might hinder exploration of the parameter space. In the present application context, however, IOV counteracts in addition to the rejuvenation step weight degeneracy.

Sequential data processing is not only computationally efficient and convenient for IOV handling, but has the additional advantage that already assimilated experimental data need not be stored to assimilate future data points. Sampling approaches allow a simple extension for hierarchical models to include the uncertainties in the population parameters for even more holistic uncertainty quantification. This would enable a continuous learning process between clinical trials from drug development (e.g., phase III) and continue during the acquisition of real-world data after market authorization in quantifying the diverse population of patients who have taken a given drug. For a future patient, this “historic” diversity would transform into well-quantified uncertainty in a TDM setting. The absence of need to store “historic” experimental

data can also be helpful for the exchange of information among clinics, health insurances, and pharmaceutical companies. The current knowledge, present in form of a sample of particles, can easily be exchanged without the need to exchange the experimental data. The “historic” data are implicitly present in the particles.

In view of new treatments and new mobile healthcare devices (e.g., wearables) gathering data from various sources, clinicians have to deal with new challenges and an increasing complexity of treatment decision making, which demands for comprehensive approaches that integrate data efficiently and provide informative and reliable decision support. We illustrated that comprehensive uncertainty quantification can result in a more informative, reliable, and differentiated decision support, which is not only limited to individualized chemotherapy but has the potential to improve patient care in various therapeutic areas in which TDM is indicated, such as oncology, infectious

diseases, inflammatory diseases, psychiatry, and transplantation patients.

Supporting Information. Supplementary information accompanies this paper on the *CPT: Pharmacometrics & Systems Pharmacology* website (www.psp-journal.com).

Supplementary Material - SText.pdf. Contains additional information regarding the employed methods, a detailed overview of the PK/PD models and some additional supporting analyses and figures.

Supplementary Material - SCode.zip. Contains the model, implemented algorithms and figure generating MATLAB code.

Acknowledgments. Fruitful discussions with Andrea Henrich (Idorsia Pharmaceuticals Ltd, Allschwil), Sven Mensing (AbbVie, Ludwigshafen), and Sebastian Reich (University of Potsdam, University of Reading) are kindly acknowledged.

Funding. Graduate Research Training Program PharMetrX: Pharmacometrics & Computational Disease Modelling, Berlin/Potsdam, Germany. Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1294 Data Assimilation (associated project). Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of University of Potsdam.

Conflict of Interest. C.K. and W.H. report research grants from an industry consortium (AbbVie Deutschland GmbH & Co. K.G., Boehringer Ingelheim Pharma GmbH & Co. K.G., Grünenthal GmbH, F. Hoffmann-La Roche Ltd., Merck KGaA, and Sanofi) for the PharMetrX program. In addition C.K. reports research grants from the Innovative Medicines Initiative-Joint Undertaking (“DDMoRe”) and Diurnal Ltd. All other authors declared no competing interests for this work.

Author Contributions. C.M., N.H., C.K., and W.H. wrote the manuscript. C.M., N.H., J.dW., C.K., and W.H. designed the research. C.M. performed the research. C.M., N.H., C.K., and W.H. analyzed the data.

1. Keizer, R.J., Heine, R., Frymoyer, A., Lesko, L.J., Mangat, R. & Goswami, S. Model-informed precision dosing at the bedside: scientific challenges and opportunities. *CPT Pharmacometrics Syst. Pharmacol.* **7**, 785–787 (2018).
2. Sheiner, L.B., Beal, S., Rosenberg, B. & Marathe, V.V. Forecasting individual pharmacokinetics. *Clin. Pharmacol. Ther.* **26**, 294–305 (1979).
3. Reich, S. & Cotter, C. Probabilistic Forecasting and Bayesian Data Assimilation. (Cambridge University Press, Cambridge, 2015).
4. Shuman, F.G. Numerical methods in weather prediction: II. Smoothing and filtering. *Mon. Weather Rev.* **85**, 329–332 (1958).
5. Mihaylova, L., Carmi, A.Y., Septier, F., Gning, A., Kim, S. & Godsill, S. Overview of Bayesian sequential Monte Carlo methods for group and extended object tracking. *Digit. Signal Process.* **25**, 1–16 (2014).
6. Crawford, J., Dale, D.C. & Lyman, G.H. Chemotherapy-induced neutropenia: risks, consequences, and new directions for its management. *Cancer* **100**, 228–237 (2004).
7. Netterberg, I., Nielsen, E.I., Friberg, L.E. & Karlsson, M.O. Model-based prediction of myelosuppression and recovery based on frequent neutrophil monitoring. *Cancer Chemother. Pharmacol.* **80**, 343–353 (2017).
8. Friberg, L.E., Henningson, A., Maas, H., Nguyen, L. & Karlsson, M.O. Model of chemotherapy-induced myelosuppression with parameter consistency across drugs. *J. Clin. Oncol.* **20**, 4713–4721 (2002).
9. Henrich, A. et al. Semimechanistic bone marrow exhaustion pharmacokinetic/pharmacodynamic model for chemotherapy-induced cumulative neutropenia. *J. Pharmacol. Exp. Ther.* **362**, 347–358 (2017).

10. Särkkä, S. Bayesian Filtering and Smoothing. (Cambridge University Press, Cambridge, MA, 2013).
11. Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. & Rubin, D.B. Bayesian Data Analysis 3rd edn. (Chapman and Hall/CRC, New York, NY, 2014).
12. Boos, D.D. & Stefanski, L. Essential Statistical Inference, Vol. **102**. (Springer, New York, NY, 2013).
13. Kümmel, A., Bonate, P.L., Dingemans, J. & Krause, A. Confidence and prediction intervals for pharmacometric models. *CPT Pharmacometrics Syst. Pharmacol.* **7**, 360–373 (2018).
14. Wasserman, L. All of Statistics: A Concise Course in Statistical Inference. (Springer Science & Business Media, New York, NY, 2000).
15. Ballnus, B., Hug, S., Hatz, K., Görlitz, L., Hasenauer, J. & Theis, F.J. Comprehensive benchmarking of Markov Chain Monte Carlo methods for dynamical systems. *BMC Syst. Biol.* **11**, 1–18 (2017).
16. Wakefield, J. Bayesian individualization via sampling-based methods. *J. Pharmacokin. Biopharm.* **24**, 103–131 (1996).
17. Gordon, N., Salmond, D. & Smith, A. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proc. F Radar Signal Process.* **140**, 107 (1993).
18. Arulampalam, M.S., Maskell, S., Gordon, N. & Clapp, T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans. Signal Process.* **50**, 174–188 (2002).
19. Murphy, K.P. Machine Learning: A Probabilistic Perspective, Vol. **1**. (The MIT Press, Cambridge, MA, 2012).
20. Kloft, C., Wallin, J., Henningson, A., Chatelut, E. & Karlsson, M.O. Population pharmacokinetic-pharmacodynamic model for neutropenia with patient subgroup identification: comparison across anticancer drugs. *Clin. Cancer Res.* **12**, 5481–5490 (2006).
21. Wallin, J.E., Friberg, L.E. & Karlsson, M.O. A tool for neutrophil guided dose adaptation in chemotherapy. *Comput. Methods Programs Biomed.* **93**, 283–291 (2009).
22. Joerger, M. et al. Open-label, randomized study of individualized, pharmacokinetically (PK)-guided dosing of paclitaxel combined with carboplatin or cisplatin in patients with advanced non-small-cell lung cancer (NSCLC). *Ann. Oncol.* **27**, 1895–1902 (2016).
23. Di Maio, M., Gridelli, C., Gallo, C. & Perrone, F. Chemotherapy-induced neutropenia: a useful predictor of treatment efficacy? *Nat. Clin. Pract. Oncol.* **3**, 114–115 (2006).
24. Bishop, C.M. Pattern Recognition and Machine Learning. (Springer Science+Business Media, Berlin, 2006).
25. Jermyn, I.H. Invariant Bayesian estimation on manifolds. *Ann. Stat.* **33**, 583–605 (2005).
26. Lavielle, M. Mixed Effects Models for the Population Approach. (Chapman and Hall/CRC, New York, NY, 2014).
27. Chaouch, A., Hooper, R., Csajka, C., Rousson, V., Thoma, Y. & Buclin, T. Building up a posteriori percentiles for therapeutic drug monitoring. PAGE: Abstracts of the Annual Meeting of the Population Approach Group in Europe. ISSN 1871-6032. Abstr 5979, Lisboa, Port. (2016). <www.page-meeting.org/?abstract=5979>
28. Le, T.T.T. et al. A mathematical model of white blood cell dynamics during maintenance therapy of childhood acute lymphoblastic leukemia. *Math. Med. Biol.* **36**, 471–488 (2019).
29. Albers, D.J., Levine, M., Gluckman, B., Ginsberg, H., Hripscak, G. & Mamykina, L. Personalized glucose forecasting for type 2 diabetes using data assimilation. *PLoS Comput. Biol.* **13**, e1005232 (2017).
30. Albers, D.J., Levine, M.E., Stuart, A., Mamykina, L., Gluckman, B. & Hripscak, G. Mechanistic machine learning: how data assimilation leverages physiologic knowledge using Bayesian inference to forecast the future, infer the present, and phenotype. *J. Am. Med. Informatics Assoc.* **25**, 1392–1401 (2018).
31. National Cancer Institute. Common terminology criteria for adverse events (CTCAE) version 4.03. <https://www.eortc.be/services/doc/ctc/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf>. (2010).

© 2020 The Authors. *CPT: Pharmacometrics & Systems Pharmacology* published by Wiley Periodicals, Inc. on behalf of the American Society for Clinical Pharmacology and Therapeutics. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.