HUMBOLDT-UNIVERSITÄT ZU BERLIN

Wirtschaftswissenschaftliche Fakultät

Masterarbeit

Zum Erwerb des akademischen Grades Master of Science

Parameter Uncertainty in the Optimization of Pharmacokinetic Models:
A Reproducible Bayesian Approach

Parameterunsicherheit bei der Optimierung pharmakokinetischer Modelle:
Ein reproduzierbarer bayesianischer Ansatz

vorgelegt von
Antonio Aquiles Alvarez Aguilar
Matrikelnummer: 621205
Email: antonio.alvarez@student.hu-berlin.de
Geburtsdatum und -ort: 10.08.1994 in Guayaquil, Ecuador
Arbeitsgruppe: Systems Medicine of the Liver

Erstprüfer: Dr. Matthias König
Zweitprüfer: Prof. Dr. Frank Konietschke

Table of Contents

Abstract

Bayesian methods offer a principled framework for parameter estimation and uncertainty quantification in physiologically-based pharmacokinetic (PBPK) modeling, yet the sensitivity of these approaches to key workflow hyperparameters remains insufficiently characterized. This thesis presents a reproducible methodology for evaluating Bayesian Markov Chain Monte Carlo (MCMC) parameter estimation across three pharmacokinetic models of increasing complexity: a simple chain model, a basic pharmacokinetic model, and a clinically validated PBPK model of indocyanine green (ICG). Synthetic pharmacokinetic datasets were generated from known parameter distributions and used to fit Bayesian MCMC samplers under systematically varied hyperparameters — number of timepoints, number of samples, coefficient of variation, and prior type. Posterior accuracy, point bias, computational cost, and sampling efficiency were evaluated across all conditions. Among the four hyperparameters investigated, prior specification emerged as the dominant determinant of posterior accuracy: well-specified priors yielded parameter estimates close to the true values across all models, while biased and incorrect priors introduced substantial and progressive deviations. The remaining hyperparameters had limited influence on estimation accuracy but the number of timepoints and samples directly increased computational cost. The entire workflow — built on SBML, libRoadRunner, PEtab, and pyPESTO — is fully open-source and FAIR-compliant, providing a transparent and reusable framework for validating Bayesian approaches in pharmacokinetic modeling.

Zusammenfassung

Bayesianische Methoden bieten einen prinzipiellen Rahmen für die Parameterschätzung und die Quantifizierung von Unsicherheiten in der physiologisch basierten Pharmakokinetik (PBPK) Modellierung, doch ist die Sensitivität dieser Ansätze gegenüber Hyperparametern des Workflows bislang noch unzureichend charakterisiert. Diese Arbeit stellt eine reproduzierbare Methodik zur Bewertung der bayesschen Markov-Ketten-Monte-Carlo- (MCMC-) Parameterschätzung anhand von drei pharmakokinetischen Modellen mit zunehmender Komplexität vor: einer einfachen linearen Kette, einem minimalem pharmakokinetischen Modell und einem klinisch validierten PBPK-Modell für Indocyaningrün (ICG). Synthetische pharmakokinetische Datensätze wurden aus bekannten Parameterverteilungen generiert und verwendet, um Bayes'sche MCMC-Sampler unter systematisch variierten Hyperparametern anzupassen – Anzahl der Zeitpunkte, Anzahl der Samples, Variationskoeffizient und Prior-Typ. Die Genauigkeit der Posterior-Verteilung, der Punktbias, der Rechenaufwand und die Stichproben-Effizienz wurden unter allen Bedingungen bewertet. Unter den vier untersuchten Hyperparametern erwies sich die Prior-Spezifikation als der dominierende Faktor für die posteriorische Genauigkeit: Gut spezifizierte Prioren lieferten in allen Modellen Parameterschätzungen, die nahe an den wahren Werten lagen, während verzerrte und falsche Prioren erhebliche und zunehmende Abweichungen verursachten. Die übrigen Hyperparameter hatten nur begrenzten Einfluss auf die Schätzgenauigkeit, doch die Anzahl der Zeitpunkte und Stichproben erhöhte den Rechenaufwand direkt. Der gesamte Workflow – basierend auf SBML, libRoadRunner, PEtab und pyPESTO – ist vollständig Open-Source und FAIR-konform und bietet ein transparentes und wiederverwendbares Framework zur Validierung bayesscher Ansätze in der pharmakokinetischen Modellierung.

1. Introduction

Pharmacokinetic (PK) models and physiologically-based pharmacokinetic (PBPK) models, formulated as systems of ordinary differential equations (ODEs) within the broader framework of systems biology, provide a quantitative basis for characterizing drug absorption, distribution, metabolism, and elimination in clinical settings. These models are instrumental in addressing critical pharmacological questions, including the prediction of drug–drug interactions and the optimization of dosing regimens for specific patient populations.

1.1. Pharmacokinetics

Pharmacokinetic models, including PBPK models, are structured around four fundamental processes that govern the fate of a drug within the body: absorption, distribution, metabolism, and excretion (ADME). Figure 1 provides a schematic overview of these processes.

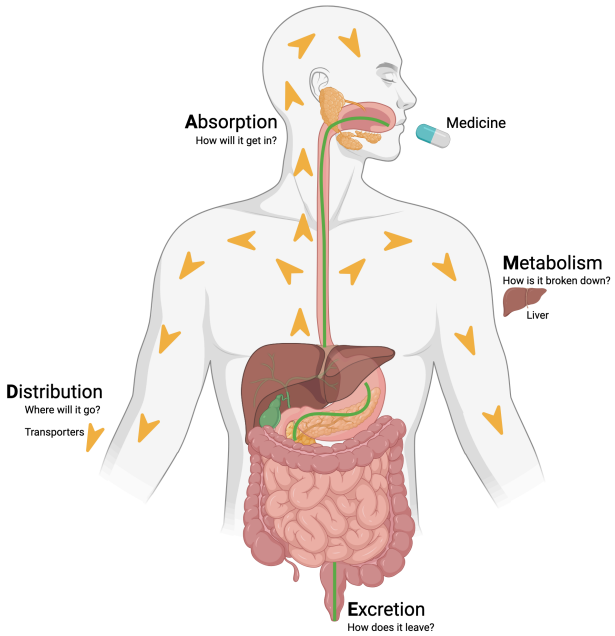

Figure 1: **Basic Diagram of ADME** Key processes are absorption, distribution, metabolism are excretion. Adapted from [1].

Absorption refers to the process by which a drug is transferred from the site of administration into the systemic circulation [2]. This process encompasses route-specific phenomena such as transporter saturation and delayed release. It begins with the release of the drug from its formulation and its subsequent dissolution in gastrointestinal fluids. Drug molecules must then traverse the epithelial barrier of the gut via passive diffusion or carrier-mediated uptake and efflux transport mechanisms. Owing to the complexity of tissue architecture and gastrointestinal physiology, absorption is not uniform along the gastrointestinal tract [3]. Accordingly, mathematical models such as PK models partition the tract into discrete segments, each characterized by distinct enzyme expression profiles, surface areas, fluid volumes, transit times, and pH values [4].

Metabolism encompasses the biotransformation of the parent drug into metabolites through the activity of metabolizing enzymes [5]. PK models frequently incorporate both hepatic and extrahepatic metabolic pathways, represented by clearance parameters and enzyme-kinetic equations. Metabolism generally proceeds through two broad phases. Phase I reactions introduce or unmask functional groups, predominantly via cytochrome P450 enzymes such as CYP3A4 and CYP2C19 [6]; common Phase I reactions include oxidation, reduction, and hydrolysis. Phase II reactions subsequently conjugate the drug or its Phase I metabolites with endogenous molecules, thereby increasing polarity and facilitating excretion [7]. In this regard,

PK models can provide mechanistic insight into drug–drug interactions, metabolite kinetics, and clinical scenario analyses.

Finally, excretion denotes the irreversible elimination of the drug and its metabolites from the body, primarily via renal and biliary routes. Urinary excretion involves the clearance of drugs and metabolites from the blood through glomerular filtration and active tubular secretion mediated by uptake and efflux transporters. Tubular reabsorption may also occur, and dedicated kidney submodels can capture this process with considerable detail [8]. With respect to biliary and fecal excretion, compounds and conjugated metabolites are transported from hepatocytes into the bile and subsequently into the intestinal lumen. PBPK models commonly include a bile compartment and a biliary route into the gut lumen, where the drug may either be eliminated or undergo enterohepatic reabsorption [9]. Additional excretion pathways, including pulmonary elimination, lactational transfer, and excretion via saliva and sweat, may also be relevant. Together with metabolism, excretion governs the residence time, overall clearance, and accumulation profile of the drug in simulations. These four ADME processes constitute the foundation of PK models, and a thorough understanding of each is essential for realizing their clinical potential.

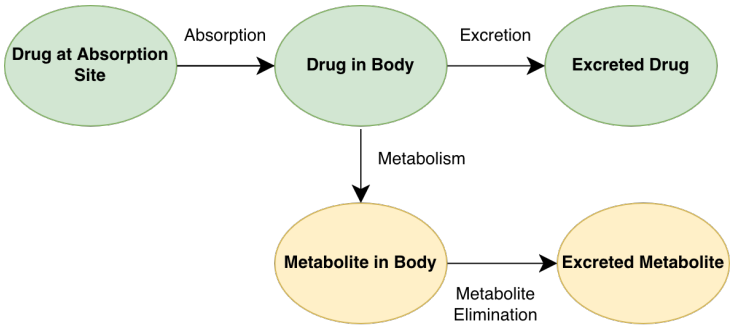

Figure 2: **A PBPK model as its most elemental.** Every circle is a compartment where the drug gets absorbed, distributed, metabolized and excreted. In PBPK, this compartment are often physioligical spaces, where organs and tissues determine parameters and equations [10].

In clinical drug development, these models are of greatest utility during the preclinical phase and Phase 0 studies [11]. They enable the assessment of drug efficacy, safety, toxicity, and bioavailability prior to the initiation of human trials. The capacity to simulate a wide range of conditions provides a basis for the rational design of subsequent clinical phases. In the context of personalized medicine, PK models facilitate the identification of optimal dosing regimens and treatment strategies by incorporating patient-specific characteristics [12]. Where the inclusion of certain patient populations in clinical trials raises ethical or methodological concerns, model-based predictions allow researchers and clinicians to evaluate patient outcomes and compare them with population-level results from clinical studies. Furthermore, PK models offer a complementary paradigm for drug research, enabling the replication of drug interactions in silico with mechanistic detail and quantifiable accuracy.

1.2. Physiologically-based Pharmacokinetic Models

Ordinary differential equations provide a mathematical framework for describing the kinetics of substances within biological systems and, as such, serve as the fundamental building blocks of PBPK models. In these models, a drug — whether exogenous or endogenous in origin — enters the biological system, undergoes distribution, and is ultimately eliminated [13]. The spatiotemporal disposition of the drug within the system constitutes the central tenet of pharmacokinetics and is explicitly represented in PBPK models. Given that the concentrations of substances within biological systems are in a constant state of flux, it is essential to formalize these processes and characterize them quantitatively.

The spatial component refers to the location of the substance within the biological system, while the temporal component describes the rate at which its quantity changes over time.

Mathematically, these two components can be combined through partial derivatives, which express rates of change with respect to both space and time. However, for practical purposes, such formulations are typically simplified to ordinary derivatives, in which spatial variation is not explicitly represented [14]. A thorough understanding of circulatory dynamics, physical chemistry, and thermodynamics is essential for the construction of such models [15].

In the present context, this reduction — or lumping — is performed on the basis of a priori knowledge of organ physiology and system biochemistry, rather than on empirical spatial and temporal kinetics, thereby rendering PBPK models tractable [16]. Compartmental models, including PBPK models, arise from this reduction, enabling the systematic formulation of modeling assumptions and the estimation of model parameters.

Several key assumptions underpin the validity of PBPK model predictions. A primary assumption is that of kinetic homogeneity, which stipulates that a substance introduced into a compartment mixes instantaneously and that each molecule has an equal probability of exiting the compartment via any available pathway [17]. A related assumption is that only unbound drug in plasma is capable of crossing biological membranes, with rapid equilibrium maintained between bound and unbound pools [18]. Additionally, anatomical and physiological structures are defined a priori with limited ranges of variability [19].

A further assumption concerns the validity of in vitro–in vivo extrapolation (IVIVE), whereby in vitro systems are scaled to in vivo conditions using standard scaling factors that are assumed to be accurate or, at minimum, unbiased on average [20]. It is also assumed that physiological parameters and drug-specific parameters are independent of one another, and that physiological parameters are transferable across compounds [21]. This separation underpins the principle that a given PBPK model framework can be applied to multiple drugs. Additional assumptions may vary from model to model depending on the specific application. Once these assumptions have been established, an experimental design can be formulated for data collection and parameter estimation.

Parameter estimation requires a site at which the drug enters the system and from which data can be collected. This site is referred to as the accessible pool. Within the accessible pool, the drug and its metabolites are assumed to be in rapid equilibrium with the sampling site [22]. Consequently, changes in this pool are directly observable through the collected data. The accessible pool thus governs much of the short-term pharmacokinetic behavior, including peak concentrations, the early distribution phase, and initial target-site exposure.

Beyond demographic descriptors of the patient population such as age and body weight, the fundamental state variable in PK modeling is the drug or substrate concentration, defined as the amount of drug present per unit volume of each compartment over time [23]. Nearly all other model quantities — including fluxes, clearances, and pharmacodynamic effects — are derived from this variable. Formally, drug concentration is expressed as the ratio of the drug amount (typically in milligrams) to the compartment volume (typically in liters) at a given time point. PBPK models track these concentration–time profiles in plasma, blood, and individual organs. Depending on the molecular size and whether the analyte is the parent drug or a metabolite, concentrations may be expressed in mass-based units (e.g., mg/L) or molar units (e.g., nM/L). This difference in units affects only the scaling convention; the underlying model structure remains unchanged [24].

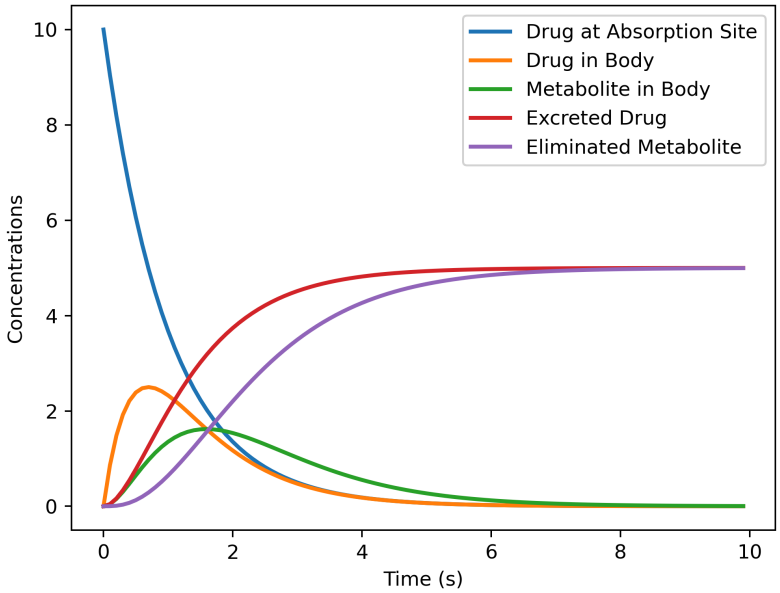

Figure 3: **Idealized drug concentrations and its metabolite in different compartments**. The time variables represents the moment the drug is first measured in each compartment. The drug starts its journey at the absorption site, where its concentrations peaks. Then it moves through the body and it gets metabolized. Finally, it gets excreted with its metabolite [10].

In addition to the accessible pool, the drug also distributes into a non-accessible pool, where it undergoes transport and transformation but cannot be directly sampled. Drug concentrations in this pool are inaccessible, as they reside in microenvironments that do not contribute directly to measurable concentrations. Examples of such analytically hidden compartments include deep tissue binding sites [25], slowly equilibrating intracellular compartments separated by membrane barriers [26] and tissue regions characterized by poor perfusion or restricted diffusion [27]. In PBPK models, the non-accessible pool can be implemented as an additional sub-compartment with slow exchange kinetics, such that the majority of the drug is sequestered and communicates with the accessible volume only gradually. Together, the accessible and non-accessible pools constitute the system.

The parameters of PBPK models serve to characterize both the accessible pool and the system as a whole. Although parameter definitions vary across models depending on the physiological attributes they represent, certain parameters are common to all PBPK models [28]. These can be classified as either accessible pool parameters or system parameters.

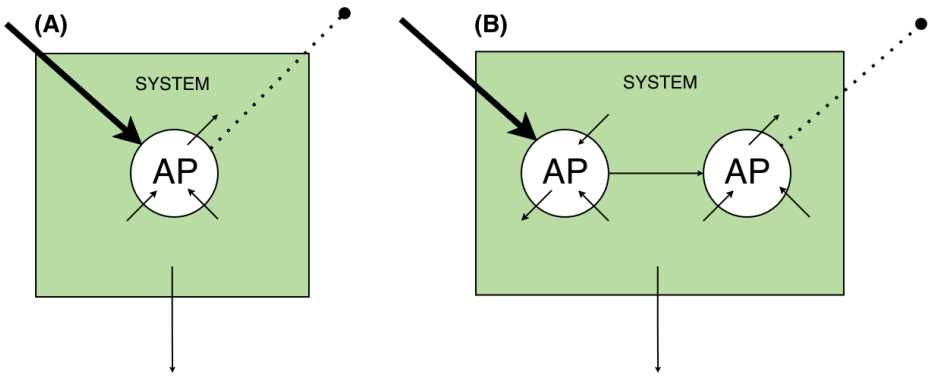

Figure 4: **Two systems with accessible pools**. System (A) has an accessible pool where a drug input, represented by a bold arrow, can enter and data, represented by a dashed line, can be collected. On the other hand, system (B) has one accessible pool where a drug input can enter and another accessible pool where data can be collected. Both have loss of input and exchanges between the accessible pool and the overall system [29].

Three widely used accessible pool parameters are described here. First, the volume of distribution of the accessible pool, V_a, (measured in liters), represents the volume in which the drug distributes uniformly and instantaneously upon entering the system. Second, the elimination rate constant, k_ε, (measured in inverse time units), denotes the fraction of drug cleared from the accessible pool per unit time. Third, the clearance rate, CL_a quantifies the volume of the accessible pool from which the drug is completely removed per unit time.

Three system-level parameters are of particular importance. The total equivalent volume of distribution, V_t, represents the total volume into which the drug distributes throughout the system. The extent of absorption, F is the fraction of the administered dose that reaches a second accessible pool following administration into the first accessible pool. The absorption rate constant, k_a, describes the fraction of the drug that appears per unit time in the second accessible pool following administration into the first. Both F and k_a are components of bioavailability [22].

Given that PK models can attain considerable complexity, the approach outlined in this thesis begins with a simple chain model comprising only two compartments that exchange a substrate between them. Subsequently, more clinically relevant models that approximate real physiological systems are employed. Detailed specifications of all models, including their parameterization and estimation procedures, are presented in Section 2.

Despite the care taken in model design to ensure clinical relevance, inter- and intra-individual variability remains a persistent concern in the development and application of PBPK models [30]. Pharmacokinetic parameters exhibit substantial variability across patients [31], and this variability arises from two broad categories of sources: exogenous and endogenous factors.

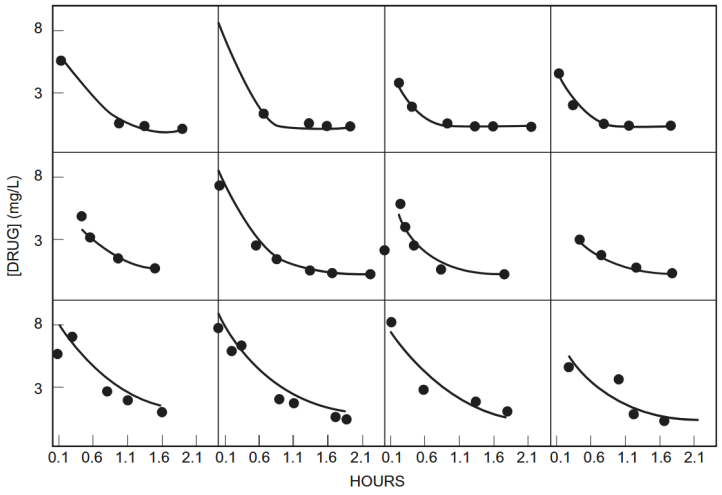

Figure 5: **Drug concentration in the same compartment but on different subjects.** Concentration data was collected on different times. The variability is evident between individuals [29].

Exogenous factors comprise external influences that may alter drug exposure and metabolism, frequently through effects on transporters or metabolizing enzymes [32]. Key examples include diet, lifestyle, environmental chemicals, and concomitant medications. Endogenous factors, by contrast, are intrinsic characteristics of the individual that influence ADME processes and pharmacodynamics [33]. Relevant endogenous factors include age, sex, body size, organ function, and genetic variation.

The interplay between endogenous and exogenous factors acting on ADME processes gives rise to interindividual variability in drug disposition [34]. For instance, organ function determines the baseline capacity for drug handling, whereas lifestyle factors may introduce time-dependent modifications in drug concentrations.

The variability of drug concentrations across individuals within a population is a central concern of population pharmacokinetics [[35]. This discipline provides a framework for modeling inter- and intra-individual variation and for estimating population-level parameter means. Several methodological approaches exist for fitting population data to individual-level PBPK models. Among the most widely used is the nonlinear mixed-effects modeling framework, popularized by the NONMEM software, in which population mean parameters are estimated alongside random effects characterized by probability distributions [36].

1.3. Statistical Modelling

For ODE-based models such as PBPK models to be applicable in practice, observational data from the process under study must be collected. The parameters of these models are then estimated by fitting the model to the observed data using a defined statistical methodology, subject to a quantifiable margin of error. Because ODEs describe time courses as particular solutions of a system of equations, the statistical framework must incorporate assumptions that accommodate these data structures

To formalize these statistical approaches, an ODE-based model can be expressed in the following general form::

$$\dot{x}_i(t_j) = \frac{d}{dt}x = f(x_i(t_j), \theta) + d_{ij} \quad x_i(0) = x_{i0} \tag{1}$$

where f denotes the system dynamics, x_i is the state of the model for subject i at time t_j, θ represents the model parameters, d_{ij} are system disturbances and x_{i0} are the initial conditions. Observations for each subject can then be expressed as:

$$y_{ij} = g\big(x_i(t_j), \hat{\theta}\big) + \varepsilon_{ij} \tag{2}$$

where g is an approximation of the true dynamics f, $\hat{\theta}$ denotes the estimated parameters and ε_{ij} represents measurement noise. In general, f is unknown as the true dynamics of the underlying natural process have not been fully elucidated. Consequently, g serves as an approximation of f represents the idealized dynamics of drug absorption, distribution, metabolism, and excretion, while g corresponds to the pharmacokinetic model employed for a specific drug.

Classical Approach

Numerous parameter estimation strategies have been developed to obtain $\hat{\theta}$ and novel methods continue to be proposed in the literature. The predominant approach in pharmacokinetics is nonlinear least squares estimation. Under this framework, the residual errors ε_{ij} are assumed to be independent and identically distributed as $N(0, \sigma^2)$ and the objective function to be minimized is given by:

$$J\big(\hat{\theta}\big) = \sum_{i=1}^{n}\sum_{j=1}^{T} \| y_{ij} - g\big(x_i(t_j), \hat{\theta}\big) \| \tag{3}$$

For a given candidate $\hat{\theta}$, the ODE system is numerically integrated to produce a complete realization of $g\left(x(t_i), \hat{\theta}\right)$. The objective function is then evaluated, and a gradient-based optimization algorithm iteratively proposes new candidates until a minimum is identified. Commonly employed optimization algorithms include Gauss–Newton, Levenberg–Marquardt, and the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm.

This estimation strategy is grounded in the frequentist approach to statistical inference [37]. The frequentist paradigm evaluates parameter estimates through the concept of repeated sampling: under the assumption that data are drawn repeatedly from the same population, population parameters can be quantified with a specified degree of precision. In this framework, the data are treated as random while the parameters are regarded as fixed quantities [38]. Uncertainty is expressed through confidence intervals, which are interpreted as containing the true parameter value in a specified proportion (typically 95%) of hypothetical repeated experiments [39].

This framework, however, presents certain limitations in the pharmacokinetic setting. Because data generation in pharmacokinetic studies is resource-intensive, the assumption of repeated sampling is difficult to justify in practice [40]. Moreover, within the frequentist paradigm, prior scientific knowledge may inform the clinical trial design but is not formally incorporated into the data analysis itself [41]. In pharmacokinetics, where data are continually generated through adaptive experimentation, models should possess the flexibility to assimilate newly available evidence on an ongoing basis [42].

Bayesian Approach

Bayesian statistical analysis is founded upon Bayes' rule of probability:

$$p(\theta|x) \propto p(x|\theta) \cdot p(\theta) \tag{4}$$

where $p(\theta|x)$ is the posterior probability of θ, $p(x|\theta)$ is the likelihood of the observed data x given θ and $p(\theta)$ is the prior probability of θ. This rule formalizes the integration of prior information with observed evidence to obtain the posterior probability distribution. Bayesian inference is thus primarily concerned with characterizing the posterior distribution of the parameters of interest. By embedding the likelihood from Equation 4 within the dynamical system framework of Equation 1, a Bayesian approach for dynamic systems can be formulated as:

$$p(\theta|x,t) \propto p(x|\theta,t) \cdot p(\theta) \tag{5}$$

Various forms of the likelihood function are available for dynamical systems, each requiring the time variable t to be explicitly specified [43]. However, the introduction of a complex likelihood renders the posterior distribution analytically intractable – a common situation in applied Bayesian statistics. This challenge has given rise to the field of Bayesian computation, in which numerical methods are employed to draw samples from an approximate posterior distribution. These methods are predominantly based on variants of Markov Chain Monte Carlo (MCMC) sampling [44].

The Bayesian approach offers several advantages over classical frequentist methods. It accommodates small sample sizes effectively, permits the explicit incorporation of prior information, and supports adaptive decision-making [45]. Because Bayesian estimators do not rely on asymptotic theory, they can yield valid probabilistic statements about model parameters even with limited data, provided that the model and priors are appropriately specified [41]. In this regard, mechanistic PBPK models stand to benefit from a Bayesian framework in which informative priors are assigned to model components and subsequently updated with observed data [46].

The Bayesian approach is particularly attractive in the present context owing to its capacity to quantify interindividual variability [47]. It enables the estimation of full parameter distributions rather than point estimates alone. The resulting model outputs, termed posterior samples, can

be analyzed to identify clinically relevant subgroups with distinct pharmacokinetic profiles, which is of direct relevance to precision dosing and patient stratification [48].

However, the increased computational complexity of Bayesian methods necessitates adequate computational resources to obtain the approximate posterior distribution, a sufficiently large sample size, and execution times that permit timely analysis of results [49]. Furthermore, priors must be specified judiciously, as uninformative or biased priors can exert a substantial influence on the posterior estimates [50].

To mitigate these limitations, two key strategies are adopted in the present work. First, a large number of samples are drawn from the MCMC sampler by allowing it to run for an extended number of iterations and retaining a large set of posterior samples upon completion. This strategy also serves to attenuate autocorrelation among successive samples, a common source of bias that is routinely addressed in Bayesian workflows [51]. Second, priors are specified in a scientifically informed manner for all parameters of interest. For example, parameters such as body weight are assigned prior distributions with mean values and standard deviations that reflect clinically established ranges.

By combining the clinical applicability of PBPK models with the probabilistic uncertainty quantification afforded by Bayesian inference, an approach for characterizing interindividual parameter variability can be formulated. However, given the considerable complexity inherent in both PBPK models and Bayesian computation, it is necessary to validate the chosen approach under a defined set of model hyperparameters. The central question addressed in this thesis is: given a PBPK model, does a Bayesian parameter estimation strategy yield reliable, clinically relevant results?

To address this question, this thesis proposes a reproducible methodology. PBPK parameters are sampled from predefined probability distributions and used to generate synthetic pharmacokinetic datasets, which are subsequently employed to fit a Bayesian MCMC sampler. The resulting posterior distributions are then compared with the original parameter distributions. By repeating this procedure across different sampler configurations, the sensitivity of the optimization results to the choice of hyperparameters can be systematically assessed.

1.4. Reproducibility and Standardization

Standardization and reproducibility remain significant challenges in the field of systems biology [52]. In the case of PBPK models specifically, many published models are not openly available: the underlying software is often proprietary, clinical datasets are inaccessible, source code is not provided, and workflows do not adhere to the principles of findability, accessibility, interoperability, and reusability (FAIR) [53].

With respect to reproducibility, many published models lack accessible equations, executable workflows, and publicly available data [54]. Even when the overall structure of a study is transparent, independent research groups frequently encounter difficulties in reproducing existing workflows. Model specifications may be ambiguous or contain errors when simulation protocols are insufficiently detailed or initial conditions are inadequately reported.

In the PBPK context, the high dimensionality of these models and the multiscale coupling of parameters across organ compartments and physiological processes compound these challenges. Where such complexity leaves room for ambiguity, published results may not be reproducible by independent investigators [55]. Moreover, when PBPK models employ virtual populations, the algorithms used to generate these populations are often insufficiently specified, potentially introducing undesirable variability.

To address these challenges, the approach presented in this thesis adheres to the recommendations of the FAIR guidelines [56]. The complete code implementation is released as open-source software, accompanied by automated tests that verify code execution. Furthermore, all models

are written in a standardized and transparent format and are made available through community repositories.

1.5. Objective and Scope

The objectives of this thesis are as follows:

1. To develop a reproducible parameter optimization approach applicable to a range of pharmacokinetic (PK) computational models.
2. To evaluate the proposed workflow using PK models of increasing complexity, from simple compartmental models to physiologically-based pharmacokinetic (PBPK) models.
3. To investigate the influence of key hyperparameters on both the PBPK model specification and the Bayesian-based optimization procedure.

2. Materials and Methods

The approach aligned on this thesis requires several software suites to work in tandem. First, models need to be defined in a format that is well known and accessible to researchers. Second, the simulations from the models must be very fast as the approach iterates through several model specifications. Third, everything needs to be stored in a format that the sampler algorithm can read. Finally, a bayesian sampler should read the simulations data and collect results. Below, the solution for each part is described throughly. Also, an introduction to the models used is outlined below.

It is imperative for computational models to be repeatable, allowing researchers to copy the same source code and data to repeat the stud, and reproducible, where models results are replicated after rebuilding in a researchers computational environment. The methods outlined below meet those requirements. The code repository that accompanies this methodology can be found at https://github.com/matthiaskoenig/parameter-variability [57].

2.1. Pharmacokinetic Models

To review the results of this approach, three different models with different levels of complexity are evaluated. They are ordered from the most simple to the most complex and clinically relevant. Their parameters and output results are described in detail.

Simple Chain

The first model used is called simple chain. Here a chemical reaction decreases the concentration amount of S_1 and increases the amount of substance S_2, all in one same homogenous compartment. This is an ODE system in the most basic. How fast S1 converts into S2 depends on the parameter k_1.

$$\begin{aligned} R_1 &= k_1 \cdot S_1 \\ \frac{\mathrm{d}}{\mathrm{d}t} S_1 &= -R_1 \\ \frac{\mathrm{d}}{\mathrm{d}t} S_2 &= R_1 \end{aligned} \tag{6}$$

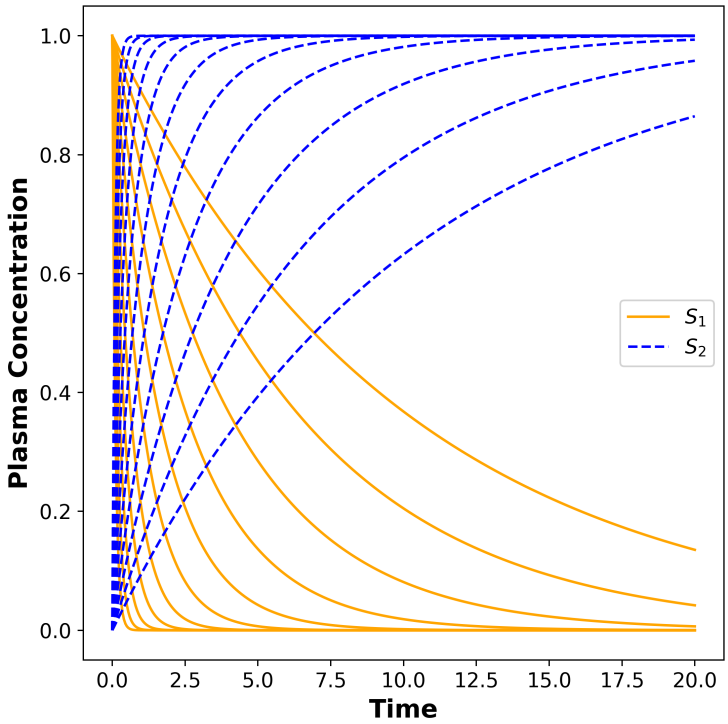

Figure 6: **Parameter Scan of k_1 for the Simple Chain model.** The parameter k_1 is varied between -1 and 1. The curves closer to the edge result from its negative values and the ones diverging result from its positives.

Simple PK

Simple PK is designed as a simple pharmacokinetic model. There are three different compartments: the gut, the central and the peripheral compartment. For each, there is a specific volume V, while they share a clearance CL, absorption k_{abs} and distribution Q parameters. With it, a researcher can model a substance being absorbed into the gut and distributed throughout peripheral and central compartment.

$$
\begin{aligned}
\text{Absorption} &= k_{\text{abs}} \cdot y_{\text{gut}} \\
\text{Clearance} &= CL \cdot y_{\text{cent}} \\
R_1 &= Q \cdot y_{\text{cent}} \\
R_2 &= Q \cdot y_{\text{peri}} \\
\\
\frac{\mathrm{d}}{\mathrm{d}t} y_{\text{gut}} &= -\frac{\text{Absorption}}{V_{\text{gut}}} \\
\frac{\mathrm{d}}{\mathrm{d}t} y_{\text{cent}} &= \frac{\text{Absorption} - \text{Clearance} - R_1 + R_2}{V_{\text{cent}}} \\
\frac{\mathrm{d}}{\mathrm{d}t} y_{\text{peri}} &= \frac{R_1 - R_2}{V_{\text{peri}}}
\end{aligned}
\tag{7}
$$

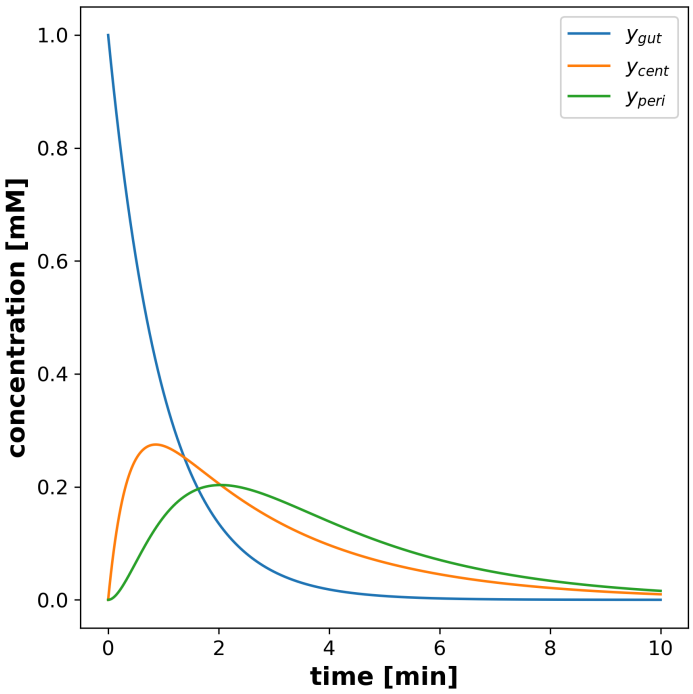

Figure 7: **Resulting concentration curves for the Simple Chain model.** The parameters k_{abs} and CL are both set at 1. Like Figure 3, the drug travels through different compartments. However, the time is now represented in minutes.

ICG PBPK

The Indocyanine green (ICG) PBPK model is the most complex and the most clinically relevant model that this approach encounters. This is an established and validated model of ICG distribution and hepatic elimination. It was developed to measure the ICG plasma concentration in different organ compartments [58]. Their main purpose is to aid on the accurate evaluation of liver function by dynamic live function tests (DLFTs).

These examinations are based on time-dependent elimination of test substances, which in turn provides a key tool for the assessments of the liver's functions [59]. The ICG achieves this objective by bounding to plasma proteins, being eliminated only by the liver and excreted into the bile [60]. As it does not get reabsorbed by the intestinal tissue, it does not undergo enterohepatic circulation. Then, its concentration can be measured in invasive and non-invasive methods such as *ex-vivo* photometric measurement, which is the gold standard for determining the ICG plasma curve [61].

A key obstacle in DLFTs based on the elimination of test substances like ICG is the significant amount of inter-individual variability in their elimination [62]. For instance, subjects with liver disease have a reduced ICG elimination that can be altered by surgical interventions like partial hepatectomy. Additionally, factors like the amount of transport proteins (OTAP1B3), binding to plasma proteins and hepatic blood flow can influence ICG's elimination.

Most of these factors are highly correlated with anthropometric factors such as age, sex and body weight. Age has a major effect in live function, flow and volume. Also, the susceptibility of live r disease increases with age [63]. Sex plays a role in liver function when other factors are controlled for [64]. In particular, body weight is of importance as there is a link between internal organ volumes and a subjects' weight, meaning that the ICG elimination rate can be affected [65].

Hence, this PBPK model gives an opportunity to accurately model the ICG concentration curves in physiologically relevant compartments and the inter-individual variability of several

subjects. Specifically, the body weight and the liver volume would be varied on this model for this thesis' objective.

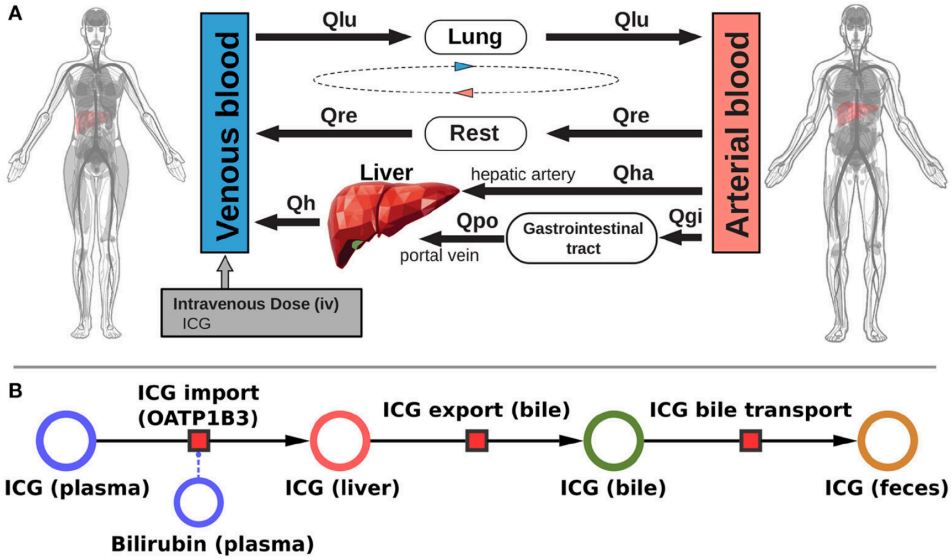

Figure 8: **Indocyanine Green PBPK model schema. (A)** Whole body PBPK model of ICG. **(B)** Liver submodel of ICG. Unlike Simple Chain and Simple PK, ICG PBPK compartments correspond to defined clinical spaces, where the drug goes through its ADME phases in parts [58].

Given the complexity of parameters and compartments, only parameters of body weight Bw and liver maximum volume $V_{\max}$ are used for the Bayesian approach. As well, because there are several compartments that are available for data analysis, only 2 compartments are going to be concentrated

2.2. Bayesian Approach

The approach outlined on this thesis starts with the sampling of the "true" parameters, which are based on the parameter of the pharmacokinetics model. Second, the parameters are fed to a simulation process that outputs the concentration curves, also called time courses. Third, parameter priors are defined and recorded, in conjunction with the time courses, in a readable format. Fourth, a bayesian computational model uses the time courses and priors to sample posterior estimates of the parameters. Finally, the true sampling parameters and the posterior parameters are compared to answer if the Bayesian model was successful in retrieving the true sampling parameters.

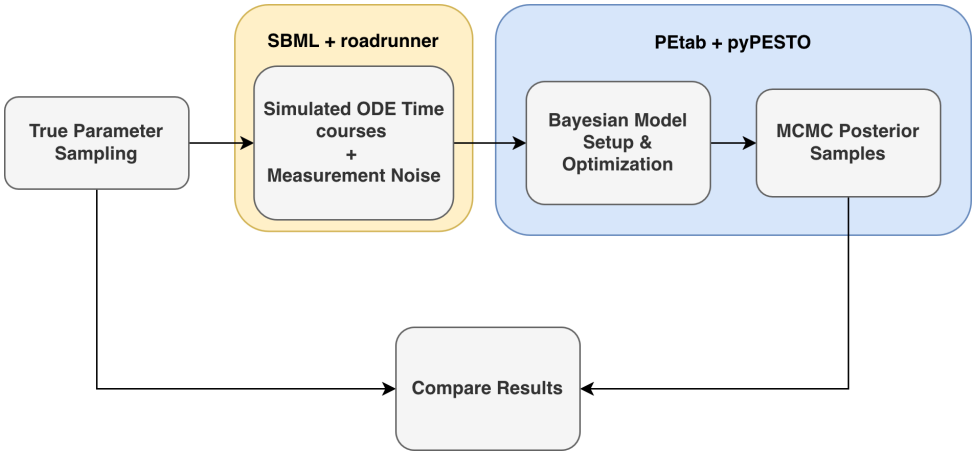

Figure 9: **Outline of the Bayesian approach.** It is a sequential approach that moves from sampling true parameters to estimate them using the simulated concentration time courses.

For each of those steps, open-sourced software for reproducible research is used, which is already visible in Figure 9 above. The reasons for their inclusion are also described.

True Parameter Sampling

Before the pharmacokinetics models are defined and the optimizer is introduced, the approach must sample the parameters of interest for the models based on the groups of interest. In particular, this approach defines the common sex-based group: MALE and FEMALE, capitalized for highlight.

To safeguard that the model parameters should always have a positive value, their distribution is decided to always follow a log normal distribution.

$$f\big(x|\mu_{\log},\sigma_{\log}\big) = \frac{1}{x\sigma_{\log}\sqrt{2\pi}}\exp\left(-\frac{\big(\ln x-\mu_{\log}\big)^2}{2\sigma_{\log}^2}\right) \tag{8}$$

where $\mu_{\log}$ is the location parameter and $\sigma_{\log}$ is the scale parameter.

However, most of our parameters are on the linear scale. Hence, it is important to transform them into a log scale. This is really important specially to codify the location $\mu_{\log}$ and scale $\sigma_{\log}$ parameter of Equation 8. The transformation is as follows:

$$\sigma_{\log} = \sqrt{\ln\left(1+\left(\frac{\sigma}{\mu}\right)^2\right)}, \quad \mu_{\log} = \ln(\mu) - \frac{\sigma_{\log}^2}{2} \tag{9}$$

where σ is the standard deviation of the parameter of interest and μ is the mean of the parameter of interest. Both are on the linear scale.

This distribution is a popular choice to describe the variability of pharmacokinetic parameters [66]. The main reason is that pharmacokinetic parameters are affected by multiplicative random influences, making their distribution often positively skewed in nature, which supports log transformation to restore normality [67].

SBML as Standardized Model Description

The System Biology Markup Language (SBML) is an XML-based representation format for storing and communicating computational models of biological processes [68]. With a widespread software support and a community of developers and users, SBML is an open standard that can represent different classes of biological phenomena. This includes cell signaling pathways, regulatory networks, metabolic networks and infectious diseases.

The SBML framework is very popular for researchers working, designing and implementing PBPK models. In it, a modeler can add physiological compartments, chemical reactions, specific model parameters and mathematical formulation that summarizes the system processes. For risk assessment, there are guidelines for documentation and evaluations of the models [69]. Finding an open source model written on the SBML standard is very accessible and requires to search it in one of the several publicly available databases like Biomodels [70].

It also possesses shared compatibility with several programming languages and framework. Statistical programming languages like R [71], numeric computing platforms like Matlab [53] and general-purpose programming languages like Python [72] all have support to read and create SBML models. As it allows to formulate ODE models, there is also several solvers that can read SBML models and find solutions to the model parameters given some initial condition [73].

libRoadRunner as ODE Model Solver

As an SBML solver, libRoadRunner is a high performance simulation engine for systems biology [74]. It is based on C++ and has interfaces for Python and Julia programming languages. This

software architecture makes is computationally fast and flexible, where extremely large models or repeated runs finish in reasonable timeframes.

It is specifically optimized for SBML-based models, where they are compiled directly into machine code using a just-in-time (JIT) compiler. Compared to other simulation solutions, the speed of its optimizer stands out and makes it suitable for large-scale simulation in emerging applications [75]. It is also has a community of developers from several research backgrounds and they also share interest in maintaining SBML models, giving confidence on the project's continue longevity.

One common aspect when measuring concentrations in dynamic systems is the involvement of error sources, often from measurement devices. To account for these event, a noise function is added to the synthetic dataset from the solver, making the data closer to what it would look in a real-world scenario. The error is applied as follows:

$$\dot{z} = z + z \cdot \nu \cdot c, \quad \nu \sim N(0, 1) \tag{10}$$

where z is the data without noise, ν is a sample from a random normal distribution and c is a coefficient of variation.

So far, libRoadRunner solves an SBML ODE-based model fast once given a parameter set, resulting in concentration curves for each compartment of the PBPK model. Then, a noise function adds a degree of measurement noise to the generated data. Hence, a researcher can create several synthetic datasets while varying the model initial parameters. Now, the need for a recording strategy arises, where the initial parameters of the model and the simulated synthetic data are saved in an organized structure.

PEtab as Data Recorder

PEtab is a parameter estimation problem definition format for model and data specification [76]. It allows model development, simulation, optimization and uncertainty analysis in one format. Specifically, a researcher can define a set of table files where model paths, data observed and parameters of interest are defined. Then, an optimizer can read through this tabular format and map its parameters in an unambiguous way. This format serves as a connection between the model parameter, the synthetic dataset and the optimizer.

pyPESTO as Optimizer

In order to estimate the parameters from the synthetic datasets, pyPESTO offers a modular framework with scalable algorithms for estimation and uncertainty quantification [77]. Orginally developed for Matlab, it is now a Python-based tool that brings a broad spectrum of interfaced tools and modern methods.

It has a long list of global and local optimizers including, Fides [78], NLopt [79], Ipopt [80] and Scipy [81]. Moreover, it provides a multi-start globalization strategy, which works well for biological systems. Noise parameters and hierarchical optimization are also other features imbedded in the software.

The main reason that the approach outline here has gravitated towards pyPESTO is its uncertainty analysis capabilities. It provides Bayesian sampling algorithms specialized for ODE models with modular parallel tempering [82], acceptance rate based scale [83] and adaptive estimation of the correlation structure [84]. These features are really important for ODE models as the ODE needs to be solved inside the MCMC sampler for each sample [85] and there can be numerical instability when computing the likelihood gradients [86].

2.3. Bayesian Model

Priors

One of the main advantages of a Bayesian approach is ability to incorporate prior scientific information directly into the model definition. Hence, for each of the parameters of interest of pharmacokinetic models, a prior distribution is defined.

To assess the influence of prior specification on this approach, three prior configurations were defined. In the Exact prior setting, the prior distribution parameters match those used to generate the true parameter values. In the Biased prior setting, the location and scale parameters deviate from the true values, though not substantially. In the Incorrect prior setting, the prior distribution parameters differ markedly from the true values, representing a scenario of severe prior misspecification.

For all cases, a zero truncated normal distribution is used to safeguard from drawing negative values during the sampling process.

Likelihood

In theory, the likelihood function could be any probability distribution. As the priors are linear and follow a truncated normal distribution, the suitable likelihood distribution for our purpose is a normal distribution, which is also a common choice in many Bayesian dynamical systems models [87].

This likelihood decision allows for adding process noise, which is a common concern in ODE based systems [43]. Many of those error sources combine and their sum is well approximated as Gaussian. Moreover, for nonlinear ODEs, normal priors and likelihood lead to posteriors that are amenable to Laplace approximations and gradient-based MCMC, which are the standard in dynamical systems inference. Also, ODE solvers uses nonlinear least squares, which is equivalent to the maximum likelihood estimation for the Gaussian distribution. Hence, it is fitting that a normal likelihood is used for this approach.

Bayesian Estimates and Sampler Diagnostics

Bayesian statistic offer several ways to analyze the posterior samples effectively. At its most basic, samples are a sequence where any sufficient statistic can be applied. However, due to our very specicfic application and data origins, the choice of statistic is one that has to be made carefully.

Moreover, for any approach that involves Bayesian computation, it is imperative to assess the quality of the sampling procedure. As MCMC can have several tuning parameters, there is a possibility of chains diverging or samples becoming correlated, which makes chains biased toward the true posterior samples.

As the parameter scale for the PK models presented is positive, the posterior median $\tilde{\theta}$ is a robust alternative to the posterior mean.

Let $\theta_1, ..., \theta_m$ be posterior samples of data y. Then, the posterior median $\tilde{\theta}$ is the 50th percentile of those samples.

$$P\Big(\theta \leq \tilde{\theta} \mid y\Big) = 0.5 \quad \text{and equivalently} \quad P\Big(\theta \geq \tilde{\theta} \mid y\Big) = 0.5 \tag{11}$$

To compute it in practice, the samples are ordered ascending sequentially as $\theta^{(1)}, ..., \theta^{(m)}$, where $\theta^{(1)}$ is the first sample and $\theta^{(m)}$ is the last sample. Then, the posterior sample median $\tilde{\theta}$ is defined as:

$$\tilde{\theta} = \begin{cases} \theta^{(\frac{m+1}{2})} & \text{if } m \text{ is odd} \\ \dfrac{\theta^{(\frac{m}{2})} + \theta^{(\frac{m}{2}+1)}}{2} & \text{if } m \text{ is even} \end{cases} \tag{12}$$

Unlike the posterior mean, this statistic is robust to outliers and data contamination, which both can be quite prevalent in the context of early clinical trials. Also, the true parameters distribution are in log normal, a skewed distribution where the median excels at finding its mode [88].

To measure the distance between the posterior sample median and true mean for each of the parameters, a simple statistic was devised for it. It measure the relative difference between the median and the true mean. The point bias b_p is defined as:

$$b_p = \frac{\tilde{\theta} - \mu}{\mu} \quad , \quad b_p \in \mathbb{R} \tag{13}$$

where $\tilde{\theta}$ is the sample median and μ is the true parameter mean.

Finally, one way to measure MCMC chain diagnostics is Effective Sample Size (ESS). It measures the amount by which autocorrelation within MCMC chains increases uncertainty in estimates [89]. It is a function of the actual draws N of a chain and a measure of autocorrelation:

$$N_{\text{eff}} = \frac{N}{1 + 2 \cdot \sum_{t=1}^{\infty} \rho_t} \tag{14}$$

where ρ_t is the autocorrelation of a sequence at different lags that increases to infinity.

Unfortunately, ρ_t is an intractable integral, making ESS incalculable. Instead, a correlation estimate $\hat{\tau}$ can be constructed for a finite sample size. For this implementation, pyPESTO uses the Sokal's adaptive truncated periodogram estimator [90]. Then, the estimated effective sample size becomes:

$$\hat{n}_{\text{eff}} = \frac{m \cdot n}{1 + 2 \cdot \sum_{t=1}^{T} \hat{\tau}_t} \tag{15}$$

where m are the number of chains, n are the number of draws in a chain and T are the amount of lags possible for n.

3. Results

Within this thesis a reproducible parameter optimization approach applicable to a range of pharmacokinetic (PK) computational models was developed. The workflow was applied to three PK models of increasing complexity, from simple compartmental models to physiologically-based pharmacokinetic (PBPK) models. The influence of key hyperparameters consisting of number of samples, timepoints, coefficient of variation and priors was studied (Table 1) on the three models.

Prior Types	**Samples**	**Time points**	**Coefficients of Variation**
Exact	1	2	0
Biased	2	3	0.001
Incorrect	3	4	0.01
	4	5	0.05
	5	10	0.1
	10	20	0.2
	20	40	0.5
	40	80	1.0
	80	160	
	160		

Table 1: **Hyperparameters of the Bayesian Approach with their respective values**. Each experiment is a realization of the hyperparameter values. The definition for prior types are found in Section 3.4.

To evaluate the effects that these hyperparameters have on the approach, 2160 simulation experiments are performed for each of the PK model. In total, 6480 experiments were carried out for the three models. Each experiment corresponds to a combination of the hyperparameters in Table 1. For example, a single experiment evaluates the simple chain model experiment with *Exact* prior type, 4 samples, 3 time points and 0.01 coefficient of variation.

3.1. Number of Timepoints

Simple Chain

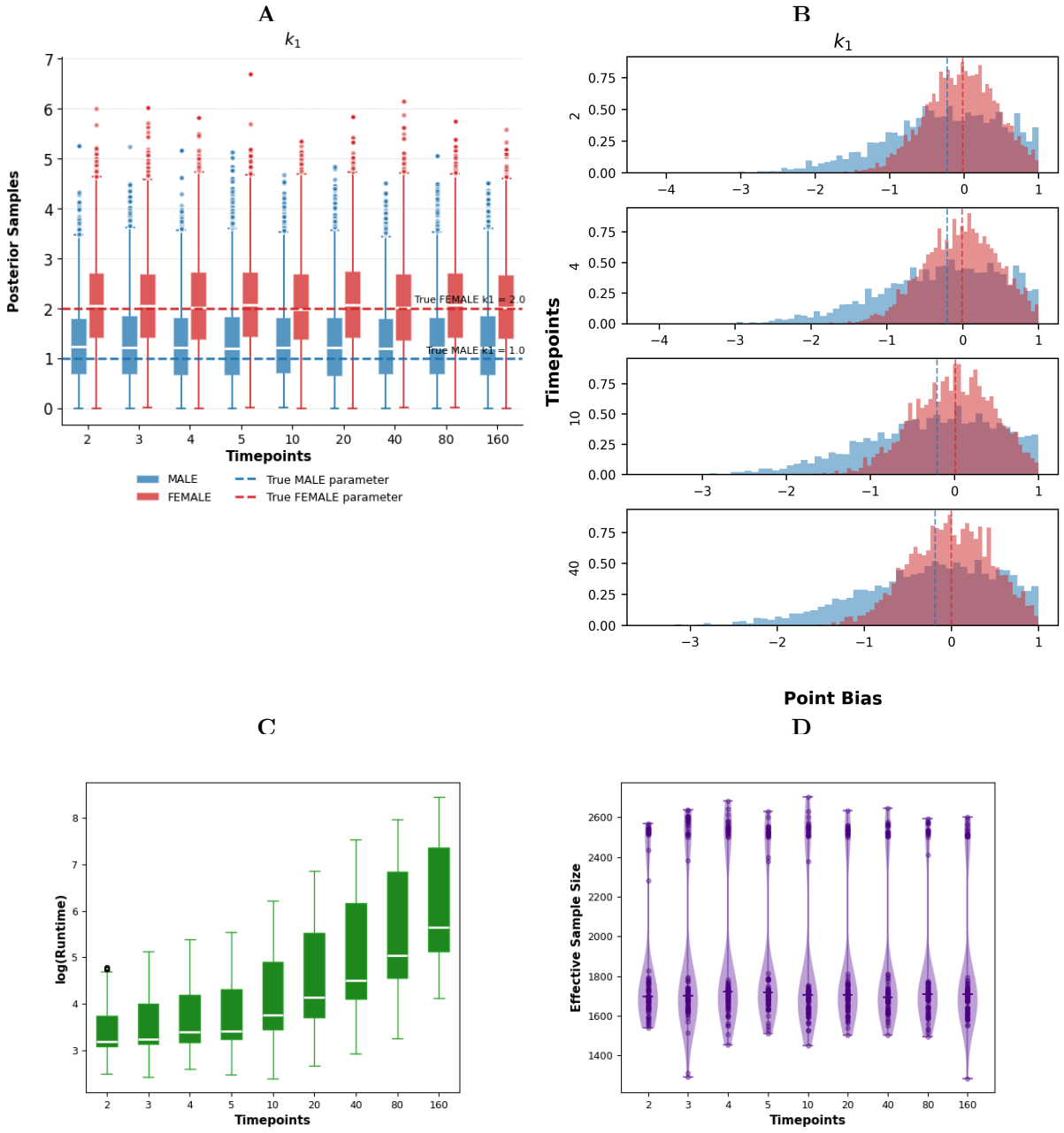

Figure 10: **Effect of varying number of time points for the Simple Chain model**. The time points for which data was generated were varied in $\{2, 3, 4, 5, 10, 20, 40, 80, 160\}$. For A and B, a subset of the results are analysed with prior type Exact (no bias), 5 samples, and a CV of 0.1 (n=8). For the analytic panels C and D, a subset of the results are analysed with prior type Exact (no bias) and all samples and CV combinations (n=720). **(A)** Box plot of the posterior samples of k_1. In dashed lines, the true parameter value of k_1 for MALE and FEMALE. **(B)** Histogram of point biases between posterior samples and the true parameter of k_1. **(C)** Box plot of optimization runtime. **(D)** Effective Sample Size violin plot of the Bayesian sampler chains (chain length 5000).

Fig 10.A presents boxplots of the posterior samples for parameter k_1 in the Simple Chain model across varying numbers of timepoints (2 to 160) for both male and female groups, with true parameter values of 1.0 and 2.0, respectively. Across all timepoint levels, the median posterior estimates for the male group remained close to the true value of 1.0, while the female group medians were consistently near the true value of 2.0. The interquartile range of the posterior samples appeared to narrow slightly as the number of timepoints increased for both groups. Both groups exhibited numerous outliers extending well beyond the interquartile range, with

values reaching up to approximately 6.0 across all timepoint conditions. Overall, the posterior distributions for the male group appeared more concentrated around the true value compared to those of the female group, which displayed greater spread at all timepoint levels.

Fig 10.B presents histograms of the point bias distribution for parameter k_1 in the Simple Chain model at four timepoint levels (2, 4, 10, and 40) for both male (blue) and female (red) groups. Both distributions peaked near zero across all timepoint levels, with the female distribution showing a slightly sharper peak closer to zero compared to the male distribution. The distributions were asymmetric, with long left tails extending to approximately -4 or -5, while being bounded near 1.0 on the right side. The overall shape and spread of both distributions remained similar across the different timepoint values. Substantial overlap between the male and female distributions was observed at all four timepoint levels.

Fig 10.C presents boxplots of the log-transformed runtime of the Bayesian MCMC sampler for the Simple Chain model across varying numbers of timepoints (2 to 160). The median log(Runtime), indicated by the red line, increased steadily from approximately 3.2 at 2 timepoints to approximately 5.6 at 160 timepoints. The interquartile ranges also widened progressively as the number of timepoints increased, with the largest spread observed at 80 and 160 timepoints. A small number of outliers were present at the lowest timepoint level of 2, while no outliers were visible at higher timepoint values. Overall, both the central tendency and variability of the log(Runtime) grew with increasing numbers of timepoints.

Fig 10.D presents violin plots of the Effective Sample Size (ESS) for the Simple Chain model across timepoint levels ranging from 2 to 160. The median ESS, indicated by the horizontal line within each violin, remained approximately constant at around 1700 across all timepoint levels. The majority of data points were concentrated in two regions: a dense cluster near 1700 and a secondary grouping near 2500, visible at every timepoint condition. The distributions spanned from approximately 1300 to 2700, with the lower tails reaching their most extreme values at 3 and 160 timepoints. The overall shape and spread of the violins were consistent across all timepoint levels examined.

Simple PK

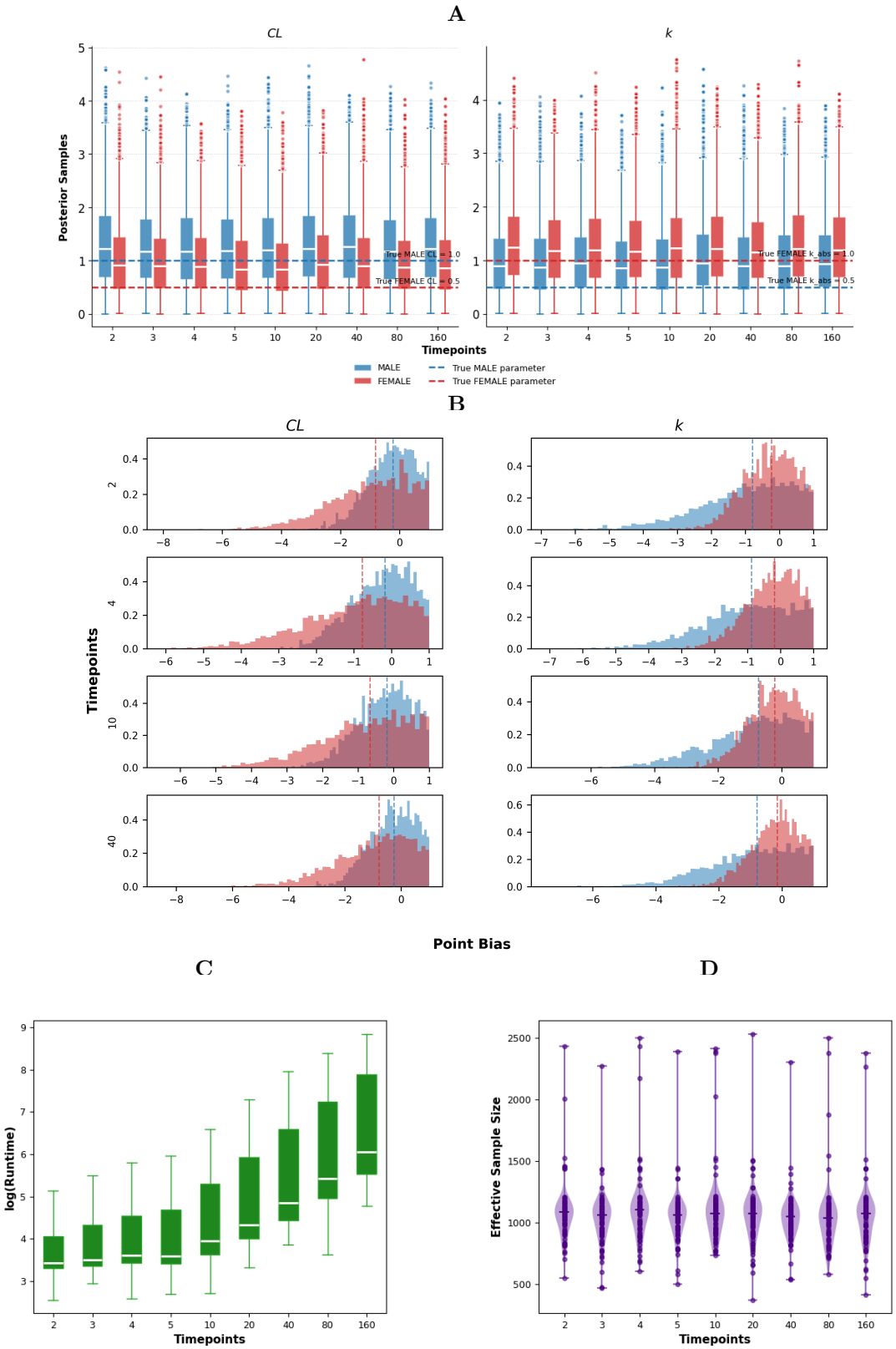

Figure 11: **Effect of varying number of timepoints for the Simple PK model**. The time points for which data was generated were varied in $\{2, 3, 4, 5, 10, 20, 40, 80, 160\}$. For A and B subset of results are analysed with prior type Exact (no bias), 5 samples, and a CV of 0.1 (n=8). For the analytic panels C and D subset of results are analysed with prior type Exact (no bias) and all samples and CV combinations (n=720). **(A)** Box plot of the posterior samples of CL and k. In dashed lines, the true parameter value for MALE and FEMALE. **(B)** Histogram of point biases between posterior samples and the true parameter of CL and k. **(C)** Box plot of optimization runtime. **(D)** Effective Sample Size violin plot of the Bayesian sampler chains (chain length 5000).

At a glance, Fig 11.A presents boxplots of the posterior samples for parameters CL (left panel) and k (right panel) in the Simple PK model across varying numbers of timepoints for male and female groups. For CL, the male median posterior estimates remained close to the true value across all timepoint levels, while the female medians were consistently above their true value of 0.5. For k, the female medians remained near their true value, while the male medians were consistently above their true value of 0.5. Both parameters exhibited outliers reaching up to approximately 5.0 across all timepoint levels for both groups. The interquartile ranges for both parameters and both groups remained relatively stable across the range of timepoints examined.

Fig 11.B displays histograms of the point bias distribution for parameters CL (left panels) and k (right panels) in the Simple PK model at four timepoint levels for male (blue) and female (red) groups. For CL, the male distribution peaked near zero across all timepoint levels, while the female distribution was shifted further to the left, with its peak located between approximately -1 and 0. For k, the pattern was reversed: the female distribution peaked closer to zero, whereas the male distribution was displaced toward more negative values. Both parameters exhibited long left tails extending to approximately -6 to -9, with the distributions bounded near 1.0 on the right side. The overall shape and spread of the distributions for both parameters remained largely consistent across the four timepoint levels, with substantial overlap between the male and female groups at each level.

Fig 11.C displays boxplots of the log-transformed runtime of the Bayesian MCMC sampler for the Simple PK model at timepoint levels ranging from 2 to 160. The median log(Runtime) rose gradually from approximately 3.5 at 2 timepoints to approximately 6.0 at 160 timepoints. A progressive widening of the interquartile ranges accompanied this increase, with the narrowest boxes observed at 2 and 3 timepoints and the widest at 80 and 160 timepoints. No outliers were visible at any of the timepoint levels. The overall pattern of increasing central tendency and spread closely resembled that observed in the Simple Chain model, though the median values were slightly higher at each corresponding timepoint level.

Fig 11.D displays violin plots of the Effective Sample Size for the Simple PK model across timepoint levels from 2 to 160. The median ESS remained stable across all timepoint conditions, with the bulk of the data points densely concentrated around the median region. Scattered individual points extended toward higher and lower values at every timepoint level, with the upper extremes reaching above 2000 and the lower extremes falling below 500 in several cases. The violins were widest around the median, tapering at both ends, and maintained a consistent shape throughout all conditions.

ICG PBPK

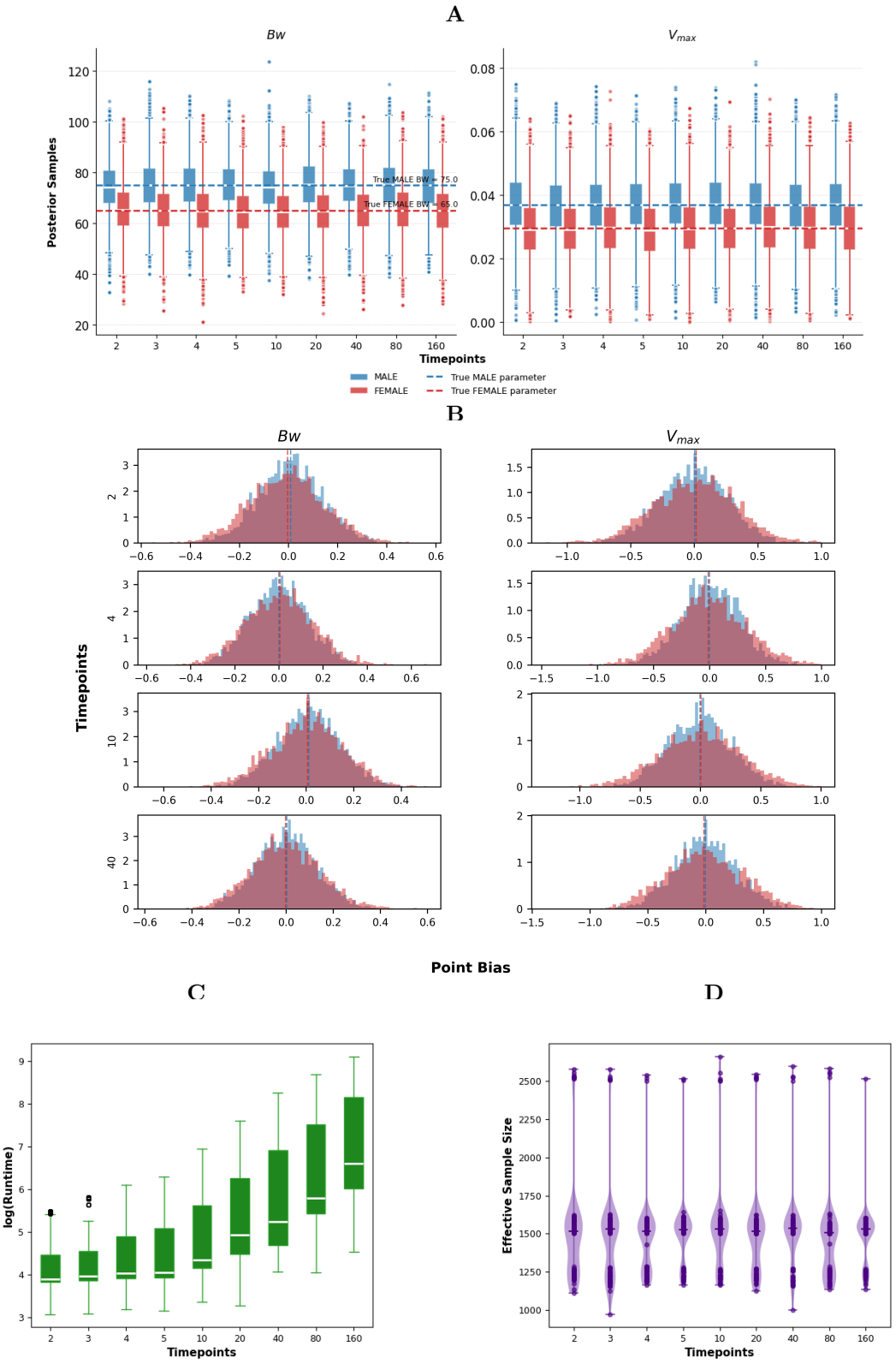

Figure 12: **Effect of varying number of timepoints for the ICG PBPK model**. The time points for which data was generated were varied in $\{2, 3, 4, 5, 10, 20, 40, 80, 160\}$. For A and B subset of results are analysed with prior type Exact (no bias), 5 samples, and a CV of 0.1 (n=8). For the analytic panels C and D subset of results are analysed with prior type Exact (no bias) and all samples and CV combinations (n=720). **(A)** Box plot of the posterior samples of Bw and $V_{\max}$. In dashed lines, the true parameter value for MALE and FEMALE. **(B)** Histogram of point biases between posterior samples and the true parameter of Bw and $V_{\max}$. **(C)** Box plot of optimization runtime. **(D)** Effective Sample Size violin plot of the Bayesian sampler chains (chain length 5000).

Fig 12.A presents boxplots of the posterior samples for parameters Bw (left panel) and $V_{\max}$ (right panel) in the ICG PBPK model across varying numbers of timepoints for male and female groups. For Bw, the median posterior estimates for both groups remained close to their respective true values across all timepoints levels, with both groups displaying substantial overlap in their interquartile ranges throughout. For $V_{\max}$, the male and female median estimates were close to one another across all timepoints levels, with considerable overlap between the two distributions. Both parameters exhibited outliers across all timepoints conditions, with Bw outliers ranging from approximately 20 to 120 and $V_{\max}$ outliers extending from near 0.0 to approximately 0.08. The interquartile ranges for both parameters and both groups remained relatively consistent across all timepoint levels.

Fig 12.B illustrates the point bias distributions for parameters Bw (left panels) and $V_{\max}$ (right panels) in the ICG PBPK model across four timepoint levels (2, 4, 10, and 40), with male (blue) and female (red) groups shown in each panel. Both parameters produced approximately symmetric distributions centered near zero at all timepoint levels, in contrast to the skewed distributions observed in the previous models. The Bw distributions spanned a range of approximately -0.6 to 0.6, while $V_{\max}$ exhibited a wider range extending from roughly -1.5 to 1.0. The male and female distributions for both parameters overlapped almost entirely, with their peaks closely aligned near zero across all four timepoint conditions. No notable changes in the shape or spread of the distributions were observed as the number of timepoints increased from 2 to 40.

Fig 12.C shows boxplots of the log-transformed runtime of the Bayesian MCMC sampler for the ICG PBPK model across timepoint values from 2 to 160. Median log(Runtime) increased from approximately 3.9 at 2 timepoints to approximately 6.5 at 160 timepoints. The interquartile ranges expanded as the number of timepoints grew, with the most compact boxes at 2 and 3 timepoints and the broadest at 80 and 160 timepoints. Outliers were present at the 2 and 3 timepoint levels, appearing above the upper whiskers, while no outliers were observed at the remaining levels. Compared to both the Simple Chain and Simple PK models, the ICG PBPK model exhibited higher median log(Runtime) values at each corresponding timepoint level.

Fig 12.D presents violin plots of the Effective Sample Size for the ICG PBPK model across timepoint levels from 2 to 160. The median ESS was stable across all conditions, situated around 1550. The majority of data points were concentrated near the median, with a secondary cluster of points visible near 2500 at every timepoint level. The lower tails of the distributions extended to approximately 1000 at their most extreme, while the upper values reached above 2600. The shape of the violins remained consistent across all timepoint levels, with the widest portion centered around the median region.

3.2. Number of Samples

Simple Chain

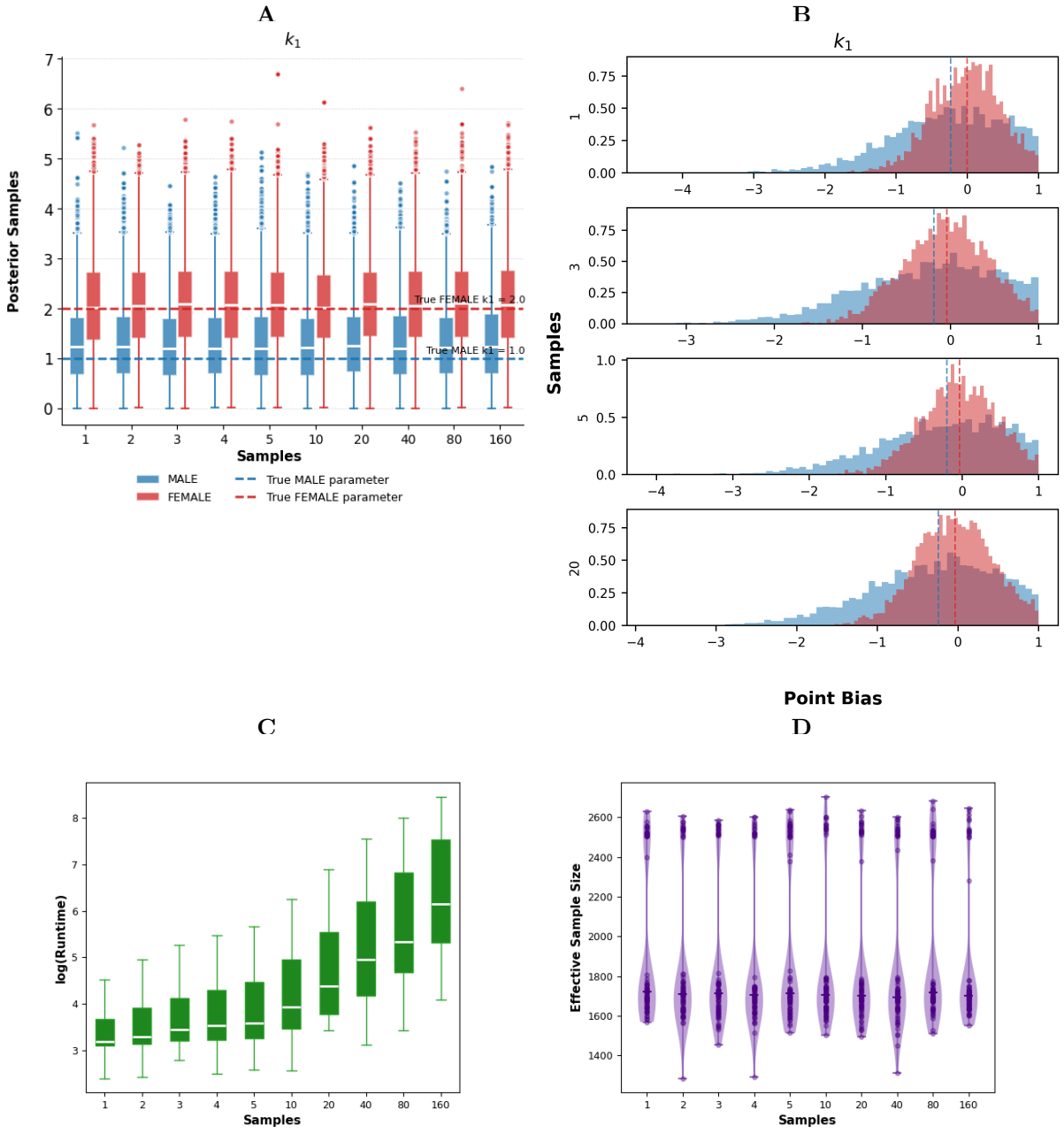

Figure 13: **Effect of varying number of samples for the Simple Chain model**. The samples for which data was generated were varied in $\{1, 2, 3, 4, 5, 10, 20, 40, 80, 160\}$. For A and B subset of results are analysed with prior type Exact (no bias), 5 time points, and a CV of 0.1 (n=8). For the analytic panels C and D subset of results are analysed with prior type Exact (no bias) and all time points and CV combinations (n=720). **(A)** Box plot of the posterior samples of k_1. In dashed lines, the true parameter value for MALE and FEMALE. **(B)** Histogram of point biases between posterior samples and the true parameter of k_1 **(C)** Box plot of optimization runtime. **(D)** Effective Sample Size violin plot of the Bayesian sampler chains (chain length 5000).

Fig 13.A presents boxplots of the posterior samples for parameter k_1 in the Simple Chain model across varying numbers of samples for male and female groups. The median posterior estimates for the male group remained close to the true value across all sample sizes, while the female group medians were consistently near their true value. Both groups displayed outliers reaching up to approximately 6.0 across all sample size conditions. The interquartile ranges for both groups remained relatively stable across the range of samples examined. The male group distributions were consistently more concentrated around lower values compared to the female group distributions across all sample sizes.

Fig 13.B shows histograms of the point bias distribution for parameter k_1 in the Simple Chain model across four sample sizes for male (blue) and female (red) groups. At all sample sizes, both distributions peaked near zero, with the female distribution consistently exhibiting a sharper peak closer to zero than the male distribution. The distributions were right-bounded near 1.0 and displayed long left tails extending to approximately -3 to -5. Considerable overlap between the male and female groups was present at each sample size level. The shape, spread, and overall range of both distributions remained largely unchanged across the four sample sizes examined.

Fig 13.C presents boxplots of the log-transformed runtime of the Bayesian MCMC sampler for the Simple Chain model across sample sizes ranging from 1 to 160. The median log(Runtime) increased from approximately 3.2 at 1 sample to approximately 6.1 at 160 samples. The interquartile ranges were narrowest at 1 and 2 samples and widest at 80 and 160 samples. No outliers were observed at any of the sample size levels. The upper whiskers extended from approximately 3.7 at 1 sample to approximately 7.5 at 160 samples.

Fig 13.D displays violin plots of the Effective Sample Size for the Simple Chain model across sample sizes ranging from 1 to 160. The median ESS remained constant across all sample sizes, positioned around 1700. Data points were concentrated in two distinct regions at every sample size: a dense cluster around the median and a secondary grouping near 2550. The lower tails extended to their most extreme values at sample sizes of 2, 4, and 40, reaching below 1300, while the upper extremes consistently fell near 2650 across all conditions. The overall shape and spread of the violins did not vary notably across the range of sample sizes examined.

Simple PK

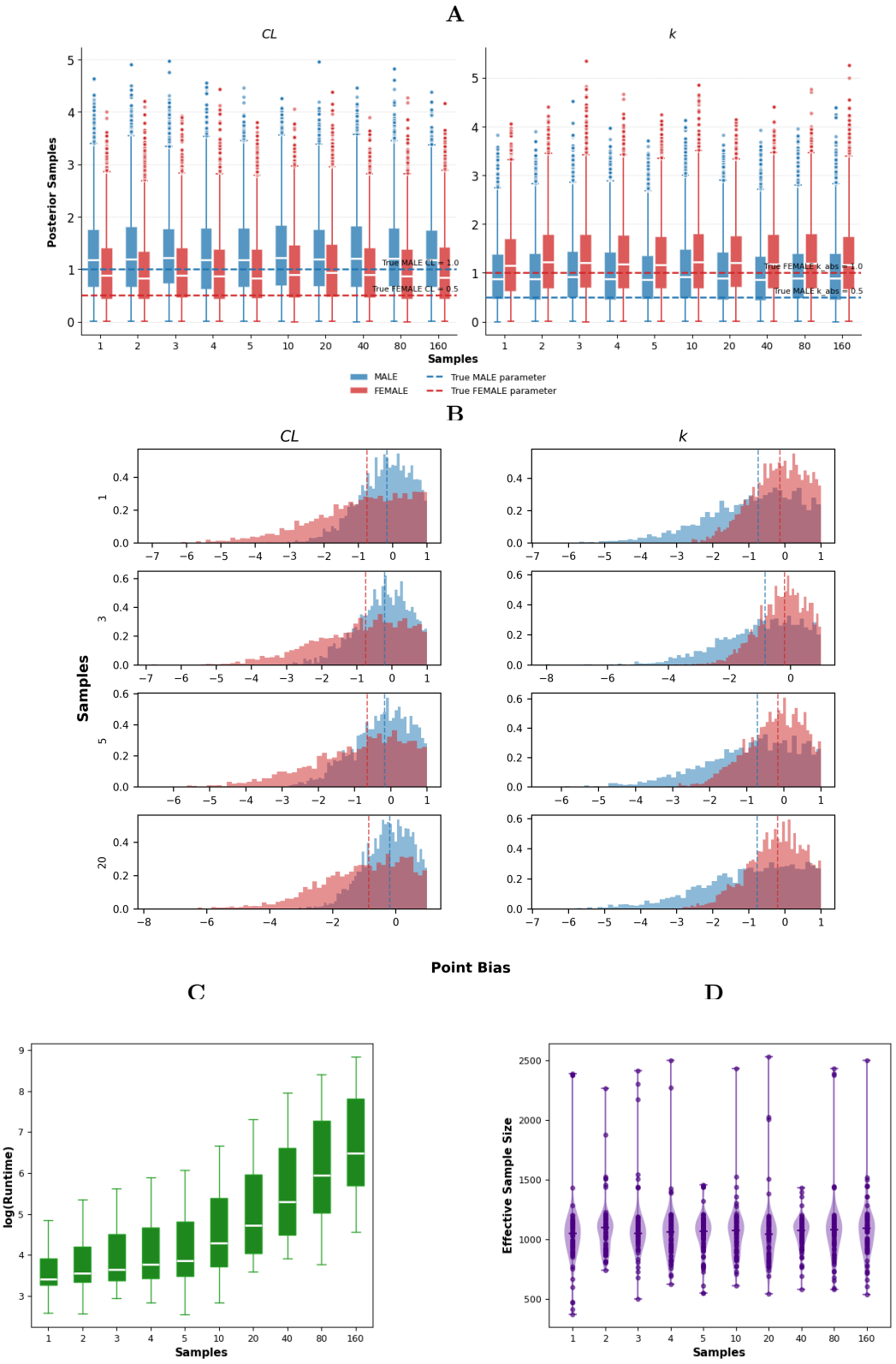

Figure 14: **Effect of varying number of samples for the Simple PK model**. The samples for which data was generated were varied in $\{1, 2, 3, 4, 5, 10, 20, 40, 80, 160\}$. For A and B subset of results are analysed with prior type Exact (no bias), 5 time points, and a CV of 0.1 (n=8). For the analytic panels C and D subset of results are analysed with prior type Exact (no bias) and all time points and CV combinations (n=720). **(A)** Box plot of the posterior samples of CL and k. In dashed lines, the true parameter value for MALE and FEMALE. **(B)** Histogram of point biases between posterior samples and the true parameter of CL and k. **(C)** Box plot of optimization runtime. **(D)** Effective Sample Size violin plot of the Bayesian sampler chains (chain length 5000).

Fig 14.A displays the distribution of posterior samples for parameters CL (left panel) and k (right panel) in the Simple PK model as the number of samples varies from. In the left panel, male medians were situated near the true value at all sample sizes, whereas female medians consistently fell above their reference value. A similar pattern was observed in the right panel for k, where female medians aligned with their true value, while male medians were positioned above 0.5 throughout all conditions. Outliers were present for both parameters across all sample sizes, extending up to approximately 5.0 for both groups. No notable changes in the spread of the interquartile ranges were observed for either parameter as the number of samples increased from 1 to 160.

Fig 14.B depicts the point bias distributions for parameters CL (left panels) and k (right panels) in the Simple PK model across four sample sizes for male (blue) and female (red) groups. For CL, the male distribution peaked near zero at all sample sizes, while the female distribution was shifted toward more negative values, with its peak located between approximately -1 and 0. The reverse was observed for k, where the female distribution peaked closer to zero and the male distribution was displaced toward negative values. All distributions exhibited long left tails, extending to approximately -7 to -8, and were bounded near 1.0 on the right. The overall shape and range of the distributions for both parameters did not change notably across the four sample sizes.

Fig 14.C displays boxplots of the log-transformed runtime of the Bayesian MCMC sampler for the Simple PK model across sample sizes from 1 to 160. The median log(Runtime) ranged from approximately 3.5 at 1 sample to approximately 6.5 at 160 samples. The interquartile ranges were narrowest at the lower sample sizes of 1 and 2, and widest at 80 and 160 samples. No outliers were present at any sample size level. The lower whiskers spanned from approximately 3.0 at 1 sample to approximately 5.5 at 160 samples, while the upper whiskers reached from approximately 4.8 at 1 sample to approximately 8.8 at 160 samples.

Fig 14.D shows violin plots of the Effective Sample Size for the Simple PK model across sample sizes from 1 to 160. The median ESS was stable across all conditions, located around 1100. The bulk of the data points were densely packed around the median region, with scattered individual points extending toward higher and lower values at every sample size. The upper extremes reached above 2300 at most sample sizes, while the lower tails descended to values between approximately 400 and 600. The violins maintained a consistent shape across all sample sizes, with the widest portion situated around the median.

ICG PBPK

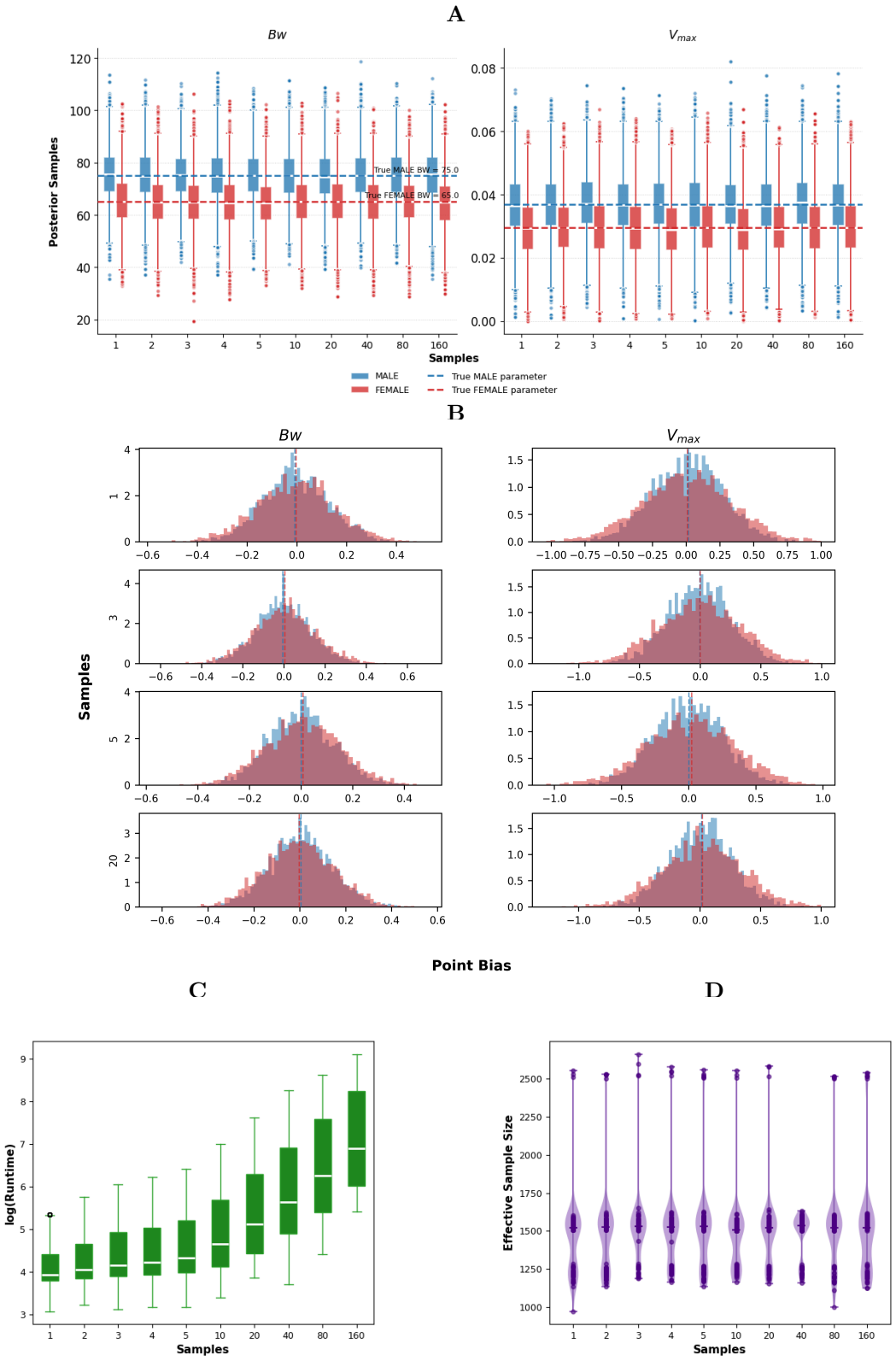

Figure 15: **Effect of varying number of samples for the ICG PBPK model**. The samples for which data was generated were varied in $\{1, 2, 3, 4, 5, 10, 20, 40, 80, 160\}$. For A and B subset of results are analysed with prior type Exact (no bias), 5 time points, and a CV of 0.1 (n=8). For the analytic panels C and D subset of results are analysed with prior type Exact (no bias) and all time points and CV combinations (n=720). **(A)** Box plot of the posterior samples of Bw and $V_{\max}$. In dashed lines, the true parameter value for MALE and FEMALE. **(B)** Histogram of point biases between posterior samples and the true parameter of Bw and $V_{\max}$. **(C)** Box plot of optimization runtime. **(D)** Effective Sample Size violin plot of the Bayesian sampler chains (chain length 5000).

Fig 15.A shows the posterior sample distributions for parameters Bw (left panel) and $V_{\max}$ (right panel) in the ICG PBPK model across sample sizes ranging from 1 to 160. For Bw, the median estimates for both groups were situated near their respective reference values across all conditions, though considerable overlap between the two distributions was present throughout. In the right panel, $V_{\max}$ medians of both groups were close together at all sample sizes, with the male and female interquartile ranges largely overlapping. Outliers spanned from approximately 20 to 120 for Bw and from near 0.0 to 0.08 for $V_{\max}$, and were observed at every sample size level for both groups. The width of the interquartile ranges for both parameters showed no substantial variation as the number of samples increased.

Fig 15.B characterizes the point bias distributions for parameters Bw (left panels) and $V_{\max}$ (right panels) in the ICG PBPK model at sample sizes of 1, 3, 5, and 20, with male (blue) and female (red) groups overlaid in each panel. The Bw histograms formed roughly bell-shaped distributions concentrated around zero, spanning from approximately -0.6 to 0.6, while the $V_{\max}$ distributions covered a broader interval from roughly -1.0 to 1.0. Across all sample sizes, the peaks of both group distributions were closely aligned near zero for both parameters. The male and female histograms were nearly indistinguishable in their coverage and height at every sample size condition. Neither the distributional form nor the range of values exhibited any marked differences as the number of samples varied from 1 to 20.

Fig 15.C shows boxplots of the log-transformed runtime of the Bayesian MCMC sampler for the ICG PBPK model across sample sizes from 1 to 160. The median log(Runtime) increased from approximately 3.9 at 1 sample to approximately 6.9 at 160 samples. The interquartile ranges were most compact at sample sizes of 1 through 5 and expanded at higher sample sizes, with the broadest boxes at 80 and 160. A single outlier was observed above the upper whisker at a sample size of 1, while no outliers were present at the remaining levels. The upper whiskers extended from approximately 4.5 at 1 sample to approximately 8.5 at 160 samples, and the lower whiskers ranged from approximately 3.5 at 1 sample to approximately 5.5 at 160 samples.

Fig 15.D presents violin plots of the Effective Sample Size for the ICG PBPK model across sample sizes ranging from 1 to 160. The median ESS remained steady at approximately 1550 throughout all sample size conditions. A secondary cluster of data points was visible near 2500 at every sample size, separated from the main concentration around the median. The lower extremes of the distributions reached down to approximately 1000, most notably at sample sizes of 1 and 80. The violin shapes were uniform across all sample sizes, with the broadest portion centered near the median and narrow tails extending toward both ends.

3.3. Coefficient of Variation

Simple Chain

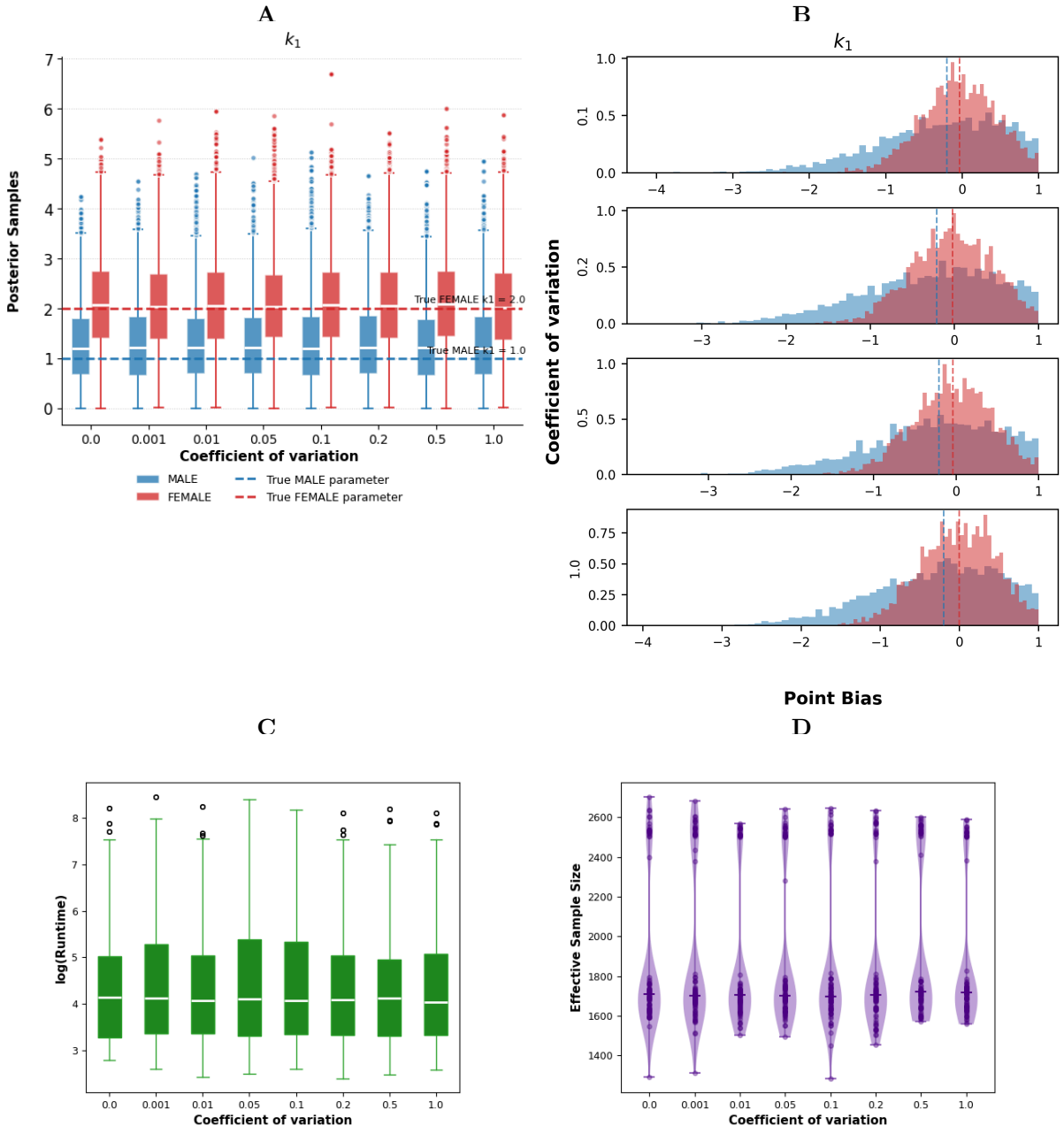

Figure 16: **Effect of varying the coefficient of variation for the Simple Chain model**. The CV for which data was generated were varied in $\{0, 0.001, 0.01, 0.05, 0.1, 0.2, 0.5, 1\}$. For A and B subset of results are analysed with prior type Exact (no bias), 5 time points, and 5 samples. For the analytic panels C and D subset of results are analysed with prior type Exact (no bias) and all time points and samples (n=720). **(A)** Box plot of the posterior samples of k_1. In dashed lines, the true parameter value for MALE and FEMALE. **(B)** Histogram of point biases between posterior samples and the true parameter of k_1 **(C)** Box plot of optimization runtime. **(D)** Effective Sample Size violin plot of the Bayesian sampler chains (chain length 5000).

Fig 16.A presents boxplots of the posterior samples for parameter k_1 in the Simple Chain model across varying levels of the coefficient of variation for male and female groups. The male median estimates were positioned near the true value at all coefficient of variation levels, and similarly, the female medians remained close to their true value throughout. Outliers were observed for both groups at every level, reaching up to approximately 7.0 at the highest coefficient of variation. The interquartile ranges for both groups remained relatively consistent across the lower coefficient of variation values, with a slight widening visible at 0.5 and 1.0. The female

group consistently exhibited a wider spread in its posterior samples compared to the male group across all conditions.

Fig 16.B presents the point bias distributions for parameter k_1 in the Simple Chain model at four coefficient of variation levels (0.1, 0.2, 0.5, and 1.0) for male (blue) and female (red) groups. Both groups exhibited their highest density near zero across all levels, with the female distribution maintaining a more pronounced peak than the male distribution throughout. The left tails of both distributions extended to approximately -3 to -4, while the right side remained bounded near 1.0 at every coefficient of variation level. The male distribution seems flat in all the coefficient of variation conditions studied, while the female distribution retained a relatively sharp peak. Overlap between the two groups was evident at all levels, though the degree of separation between their peaks appeared to increase at the highest coefficient of variation value.

Fig 16.C presents boxplots of the log-transformed runtime of the Bayesian MCMC sampler for the Simple Chain model across coefficient of variation levels ranging from 0.0 to 1.0. The median log(Runtime) remained approximately constant at around 4.0 across all coefficient of variation levels. The interquartile ranges were similarly consistent, spanning roughly from 3.5 to 5.0 at each level. Outliers were present above the upper whiskers at most coefficient of variation levels, reaching values of approximately 7.5 to 8.3. The lower whiskers extended to approximately 2.8 to 3.0 across all conditions.

Fig 16.D displays violin plots of the Effective Sample Size for the Simple Chain model across coefficient of variation levels from 0.0 to 1.0. The median ESS was consistent across all levels, located around 1700. Data points were grouped into two distinct clusters at every condition: a primary concentration near the median and a secondary grouping near 2550. The lower tails reached their most extreme values at coefficient of variation levels of 0.0, 0.001, and 0.1, descending to approximately 1300. The violins at 0.5 and 1.0 appeared slightly narrower than those at the lower coefficient of variation levels, though the overall distributional shape remained similar across all conditions.

Simple PK

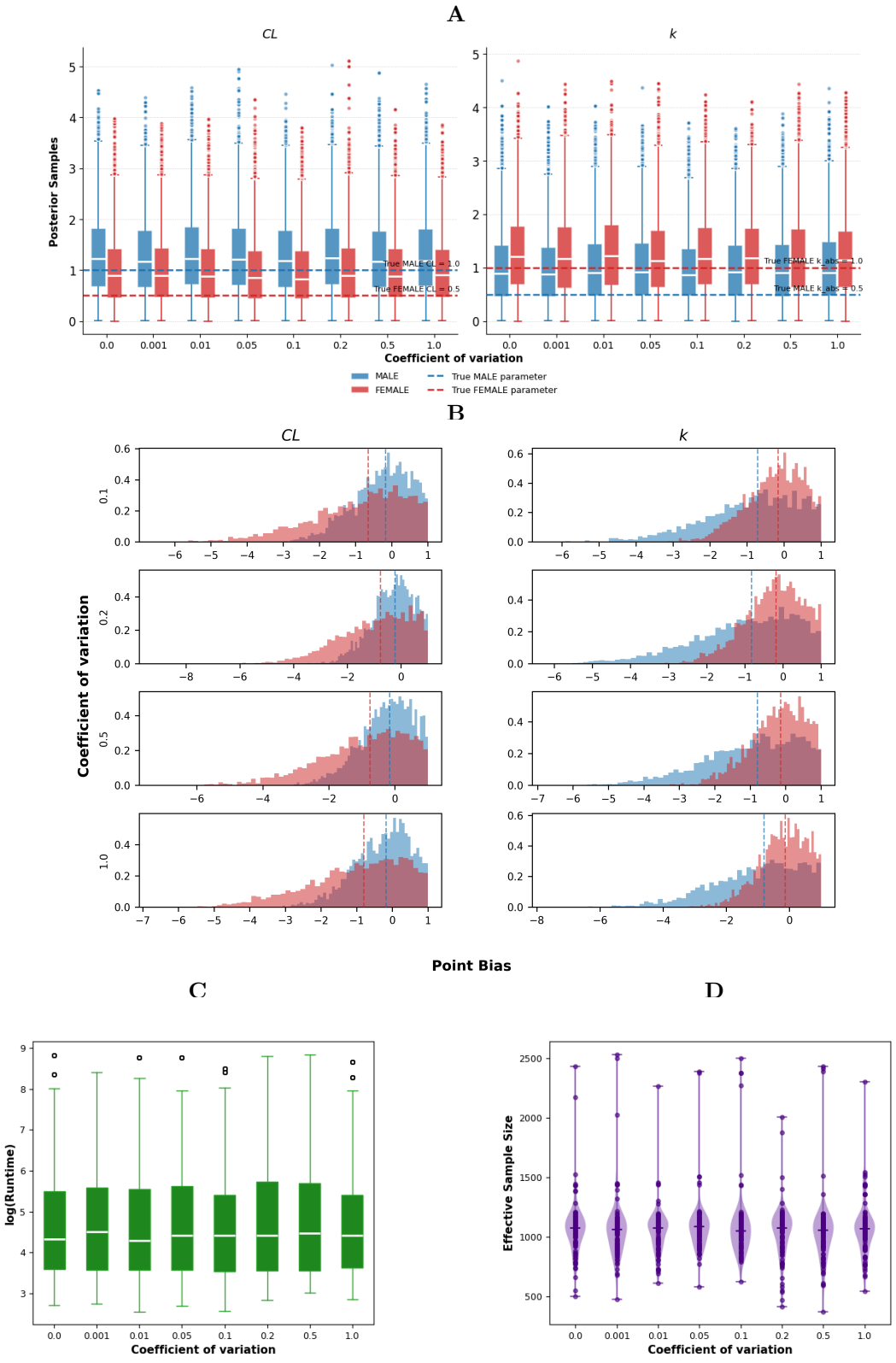

Figure 17: **Effect of varying the coefficient of variation for the Simple PK model**. The CV for which data was generated were varied in $\{0, 0.001, 0.01, 0.05, 0.1, 0.2, 0.5, 1\}$. For A and B subset of results are analysed with prior type Exact (no bias), 5 time points, and 5 samples. For the analytic panels C and D subset of results are analysed with prior type Exact (no bias) and all time points and samples (n=720). **(A)** Box plot of the posterior samples of CL and k. In dashed lines, the true parameter value for MALE and FEMALE. **(B)** Histogram of point biases between posterior samples and the true parameter of CL and k. **(C)** Box plot of optimization runtime. **(D)** Effective Sample Size violin plot of the Bayesian sampler chains (chain length 5000).

Fig 17.A depicts the posterior sample distributions for parameters CL (left panel) and k (right panel) in the Simple PK model as the coefficient of variation varies. For CL, male medians were located near the true value across all levels, while female medians remained consistently above their reference value. In the right panel, female medians were close to 1.0 at all levels, whereas male medians were situated above their true value throughout. Outliers extended up to approximately 5.0 for both parameters and were present at every coefficient of variation level. The interquartile ranges for both parameters and groups appeared to widen at higher coefficient of variation values.

Fig 17.B shows the point bias distributions for parameters CL (left panels) and k (right panels) in the Simple PK model across four coefficient of variation levels for male (blue) and female (red) groups. For CL, the male distribution peaked near zero while the female distribution was displaced toward negative values at all levels; similarly for k, the female distribution peaked closer to zero while the male distribution was shifted to the left. Both parameters produced long left tails at every coefficient of variation level, with the tails extending to approximately -6 to -8, and all distributions were bounded near 1.0 on the right side. At a coefficient of variation of 1.0, both parameters exhibited a more uniform spread of values across the negative range compared to the lower coefficient of variation conditions, where the distributions were more concentrated near their peaks. The separation between the male and female distributions remained consistent for both parameters across all four levels.

Fig 17.C depicts boxplots of the log-transformed runtime of the Bayesian MCMC sampler for the Simple PK model at coefficient of variation levels from 0.0 to 1.0. The medians were stable across all levels, situated around the same value with no discernible trend. Similarly, the interquartile ranges remained relatively uniform throughout the different conditions. Outliers appeared above the upper whiskers at several coefficient of variation levels, with the highest reaching close to 9.0. The overall spread of the distributions, as captured by the whiskers, did not differ markedly across the range of coefficient of variation values examined.

Fig 17.D presents violin plots of the Effective Sample Size for the Simple PK model across coefficient of variation levels from 0.0 to 1.0. The median ESS was stable at approximately 1100 across all conditions. The data points were primarily concentrated around the median, with individual points scattered toward higher and lower values at every level. The upper extremes reached above 2000 at most coefficient of variation levels, while the lower tails descended to values between approximately 400 and 600. The violin at 0.5 appeared wider than those at other levels, though the overall shape and central tendency of the distributions remained consistent throughout.

ICG PBPK

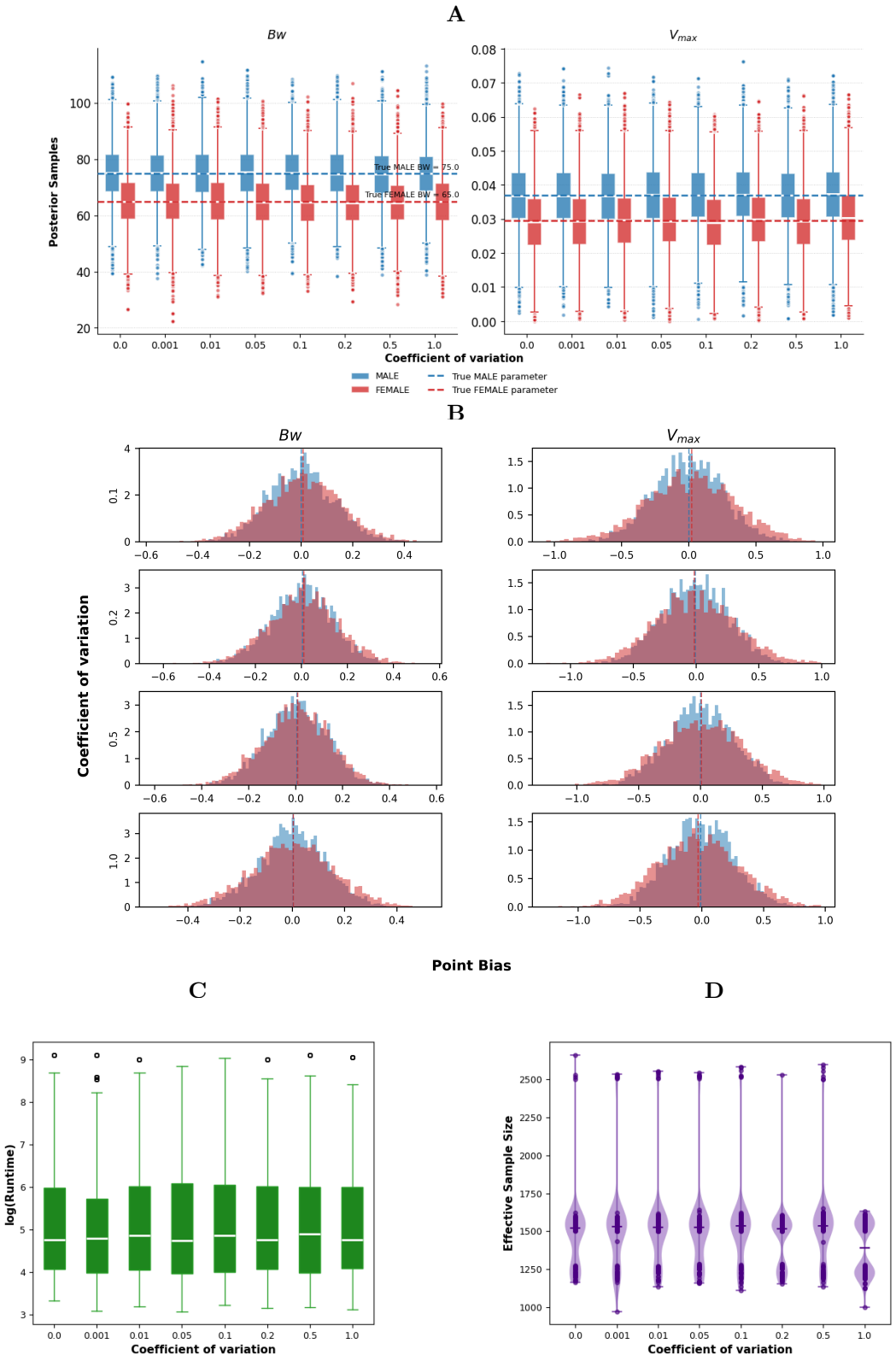

Figure 18: **Effect of varying the coefficient of variation for the ICG PBPK model**. The CV for which data was generated were varied in $\{0, 0.001, 0.01, 0.05, 0.1, 0.2, 0.5, 1\}$. For A and B subset of results are analysed with prior type Exact (no bias), 5 time points, and 5 samples. For the analytic panels C and D subset of results are analysed with prior type Exact (no bias) and all time points and samples (n=720). **(A)** Box plot of the posterior samples of Bw and $V_{\max}$. In dashed lines, the true parameter value for MALE and FEMALE. **(B)** Histogram of point biases between posterior samples and the true parameter of Bw and $V_{\max}$. **(C)** Box plot of optimization runtime. **(D)** Effective Sample Size violin plot of the Bayesian sampler chains (chain length 5000).

Fig 18.A illustrates the posterior sample distributions for parameters Bw (left panel) and $V_{\max}$ (right panel) in the ICG PBPK model across coefficient of variation levels ranging from 0.0 to 1.0. For Bw, both groups had median estimates positioned near their respective true values at all levels, with substantial overlap between the male and female distributions throughout. In the right panel, the male and female medians for $V_{\max}$ were situated close to each other across all coefficient of variation levels, with the interquartile ranges of both groups largely overlapping. Outliers ranged from approximately 20 to 120 for Bw and from near 0.0 to 0.08 for $V_{\max}$, appearing at every level for both groups. The spread of the interquartile ranges for both parameters did not vary substantially across the different coefficient of variation levels.

Fig 18.B portrays the point bias distributions for parameters Bw (left panels) and $V_{\max}$ (right panels) in the ICG PBPK model at different coefficient of variation values with male (blue) and female (red) groups overlaid. For Bw, the distributions were approximately symmetric and centered near zero at all levels, ranging from roughly -0.6 to 0.6. The $V_{\max}$ distributions spanned a wider interval, from approximately -1.0 to 1.0, and were also concentrated around zero throughout. The male and female histograms nearly completely overlapped for both parameters at every coefficient of variation level. No discernible changes in the shape, spread, or central tendency of the distributions were observed across the four coefficient of variation conditions.

Fig 18.C illustrates boxplots of the log-transformed runtime of the Bayesian MCMC sampler for the ICG PBPK model across coefficient of variation levels from 0.0 to 1.0. The median values were consistent across all conditions, remaining at approximately the same level throughout. The interquartile ranges were comparable at every coefficient of variation level, with the boxes spanning a similar extent in each case. Outliers were observed above the upper whiskers at most levels, reaching values near 9.0. The coefficient of variation at 0.001 displayed a slightly narrower box and lower upper whisker compared to the other conditions.

Fig 18.D illustrates violin plots of the Effective Sample Size for the ICG PBPK model across coefficient of variation levels from 0.0 to 1.0. The median ESS was situated around 1550 for levels 0.0 through 0.5, while at 1.0 the median was noticeably lower at approximately 1400. The secondary cluster of data points near 2500 was present at all levels from 0.0 to 0.5 but was absent at 1.0. At a coefficient of variation of 1.0, the violin was notably smaller and more compact, with data points concentrated between approximately 1000 and 1300, along with the lower median. The lower extremes of the distributions reached approximately 1000 at coefficient of variation levels 0.001 and 1.0.

3.4. Prior Type

Simple Chain

Parameter	Group	Prior Type	μ	σ
k_1	Male	Exact	1	1
		Biased	2	2
		Incorrect	5	1
	Female	Exact	2	1
		Biased	3	2
		Incorrect	6	1

Table 2: **Description of Prior types for the Simple Chain model.** *Exact* is the same distribution as the true parameter distribution, *Biased* is a partially informed prior to be close to *Exact* and *Incorrect* is a misleading prior to be quite distant from *Exact*.

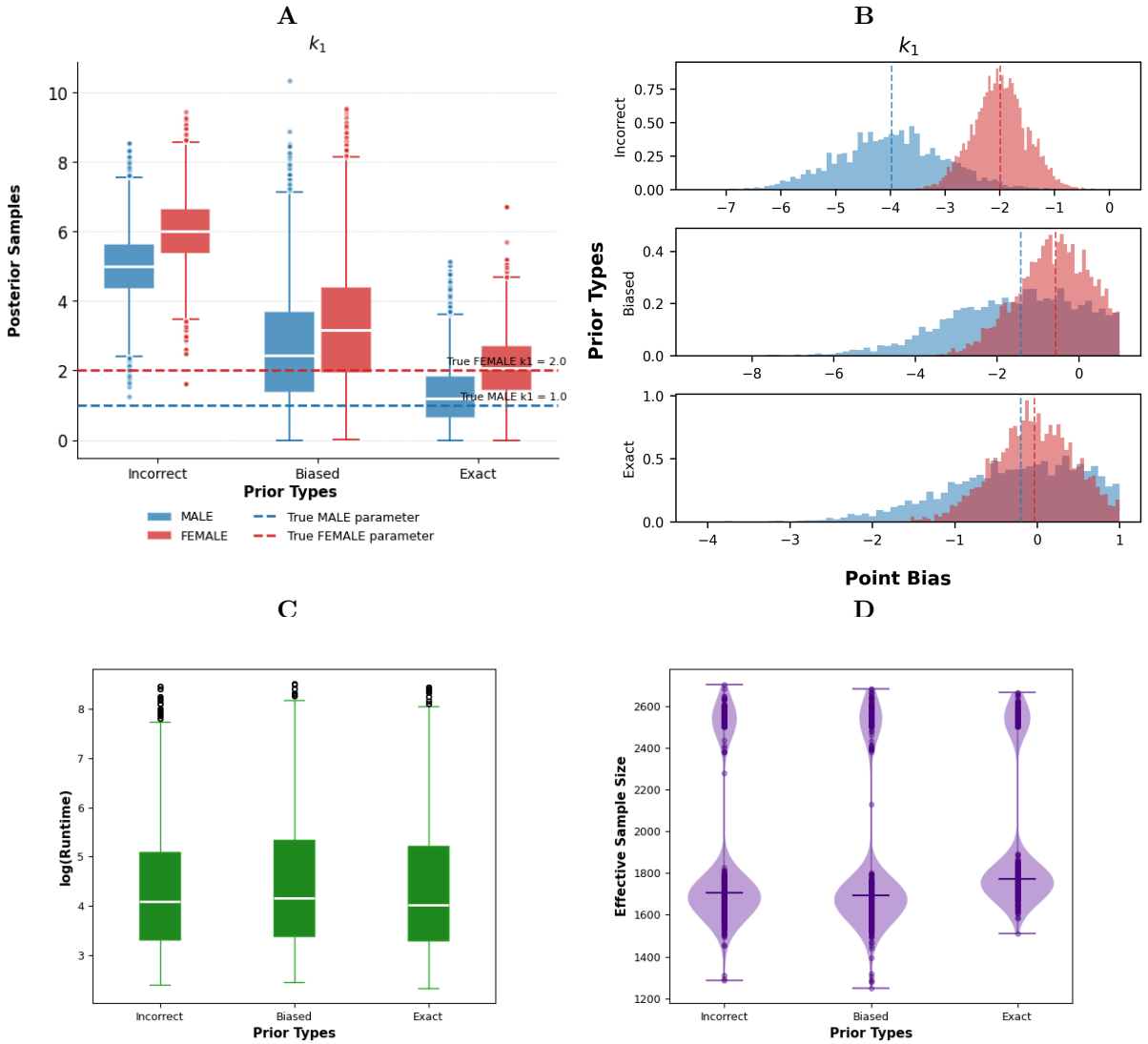

Figure 19: **Effect of varying number of prior types for the Simple Chain model**. The prior types for which data was generated were varied with prior type Exact (no bias), Biased and Incorrect. The parameter k_1 was optimized for the two groups MALE (blue) and FEMALE (red). For A and B subset of results are analysed with for 5 samples, 5 time points and a CV of 0.1 (n=3). For the analytic panels C and D subset of results are analysed with all prior types, timepoints, samples and CV combinations (n=2160). **(A)** Box plot of the posterior samples of k_1. In dashed lines, the true parameter value of k_1 for MALE and FEMALE. **(B)** Histogram of point biases between posterior sample median and the true parameter of k_1. **(C)** Box plot of optimization runtime. **(D)** Effective Sample Size violin plot of the Bayesian sampler chains (chain length 5000).

Fig 19.A presents boxplots of the posterior samples for parameter k_1 in the Simple Chain model across three prior types — Exact, Biased, and Incorrect — for male and female groups. Under the Exact prior, the medians for both groups were positioned close to their true values, with relatively narrow interquartile ranges. The Biased prior yielded higher median estimates for both groups compared to the Exact prior, along with wider interquartile ranges and outliers reaching up to approximately 10.5 for the male group. Under the Incorrect prior, both male and female medians were shifted well above their respective true values, with the male median near 5.0 and the female median near 5.5. Overall, both the central tendency and the spread of the posterior distributions increased progressively from the Exact to the Biased to the Incorrect prior type.

Under the Exact prior (bottom panel), Fig 19.B reveals that both the male (blue) and female (red) point bias distributions for parameter k_1 in the Simple Chain model peaked near zero, with left tails extending to approximately -4. When the Biased prior was applied, both distributions shifted substantially toward more negative values, with the male distribution centered around -3 and the female around -1, while the left tails reached approximately -8. The most pronounced displacement was observed under the Incorrect prior, where the male distribution peaked near -4 and the female near -2, with minimal overlap between the two groups compared to the other prior types. Across all three conditions, the distributions remained right-bounded near 1.0. A progressive separation between the male and female distributions, as well as a shift of both groups away from zero, was evident when moving from the Exact to the Biased to the Incorrect prior.

Fig 19.C presents boxplots of the log-transformed runtime of the Bayesian MCMC sampler for the Simple Chain model across three prior types: Incorrect, Biased, and Exact. The medians for all three conditions were closely aligned, falling near 4.0. The interquartile ranges were similar across the three prior types, with the Biased prior displaying a marginally taller box than the other two. Outliers were present above the upper whiskers for all three conditions, clustering between approximately 8.0 and 8.5. The lower whiskers extended to comparable values across the three prior types, reaching down to approximately 2.7 to 3.0.

Fig 19.D displays violin plots of the Effective Sample Size for the Simple Chain model across three prior types: Incorrect, Biased, and Exact. The median ESS for the Incorrect and Biased priors was approximately 1700, while the Exact prior had a slightly higher median near 1775. All three distributions featured two clusters of data points: a primary concentration around the median region and a secondary grouping near 2550. The lower tails of the Incorrect and Biased priors extended to approximately 1300 and 1250, respectively, whereas the Exact prior had a higher lower extreme near 1500. The upper extremes reached approximately 2700 for all three prior types.

Simple PK

Parameter	Group	Prior Type	μ	σ
k	Male	Exact	0.5	1
		Biased	1.5	1.5
		Incorrect	5	1
	Female	Exact	1	1
		Biased	2.0	1.5
		Incorrect	6	1
CL	Male	Exact	1	1
		Biased	1.5	1.5
		Incorrect	4	1
	Female	Exact	0.5	1
		Biased	1.0	1.5
		Incorrect	3	1

Table 3: **Description of Prior types for the Simple PK Model.** *Exact* is the same distribution as the true parameter distribution, *Biased* is a partially informed prior to be close to *Exact* and *Incorrect* is a misleading prior to be quite distant from *Exact*.

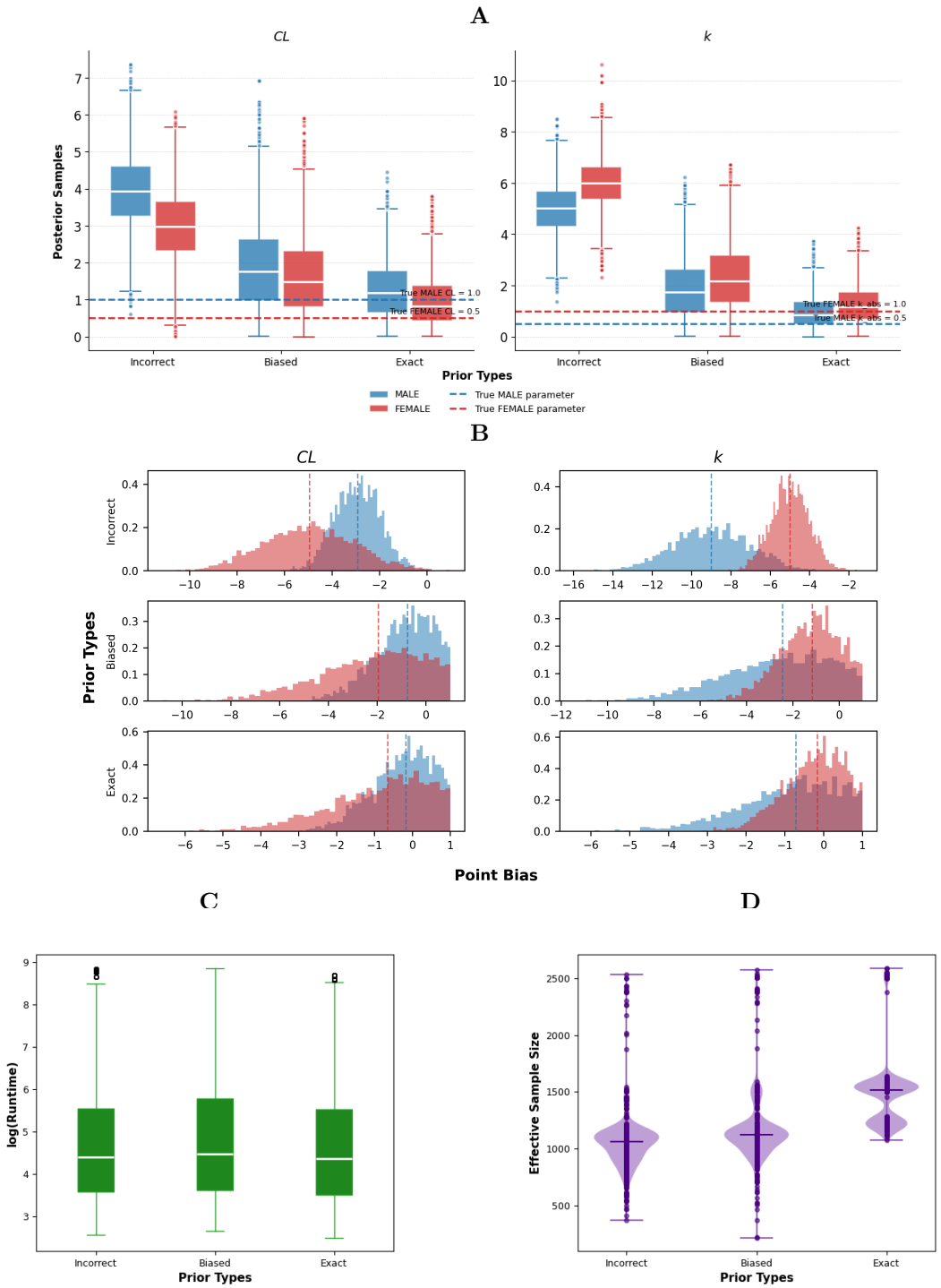

Figure 20: **Effect of varying prior types for the Simple PK model**. The prior types for which data was generated were varied with prior type Exact (no bias), Biased and Incorrect. For A and B subset of results are analysed with 5 time points, 5 samples and a CV of 0.1 (n=3). For the analytic panels C and D subset of results are analysed with all prior types, time points, samples and CVs (n=2160). **(A)** Box plot of the posterior samples of CL and k. In dashed lines, the true parameter value for MALE and FEMALE. **(B)** Histogram of point biases between posterior samples and the true parameter of CL and k. **(C)** Box plot of optimization runtime. **(D)** Effective Sample Size violin plot of the Bayesian sampler chains (chain length 5000).

Fig 20.A displays the posterior sample distributions for parameters CL (left panel) and k (right panel) in the Simple PK model across three prior types: Exact, Biased, and Incorrect. For CL, the Exact prior produced medians close to the true values for both groups, while the Biased and Incorrect priors resulted in progressively higher medians and wider interquartile ranges. A comparable pattern was observed for k, where the true values are 0.5 for males and 1.0 for females; under the Exact prior, medians were near the true values, whereas the Biased prior shifted both group medians upward, and the Incorrect prior further elevated them to

approximately 5.5 for males and 5.8 for females. The range of outliers also increased across prior types for k, extending up to approximately 10.5 under the Incorrect prior. For both parameters, the degree of separation between the male and female distributions diminished under the Biased and Incorrect priors relative to the Exact prior.

Fig 20.B displays the point bias distributions for parameters CL (left panels) and k (right panels) in the Simple PK model across three prior types: Exact, Biased, and Incorrect. Under the Exact prior, both parameters showed distributions with their peaks situated near zero, with left tails reaching approximately -6 for both CL and k. The Biased prior shifted both group distributions toward more negative values for each parameter, with left tails extending to roughly -10 for CL and -8 for k. The largest displacement from zero occurred under the Incorrect prior, where the CL distributions extended to approximately -10 and the k distributions reached as far as -16, with both male and female peaks located well below zero. Across both parameters, the gap between the male and female distributions widened progressively from the Exact to the Biased to the Incorrect prior, while the overall range of values also expanded.

Fig 20.C displays boxplots of the log-transformed runtime of the Bayesian MCMC sampler for the Simple PK model under three prior types: Incorrect, Biased, and Exact. The median log(Runtime) was approximately 4.4 for the Incorrect and Biased priors and slightly lower at around 4.3 for the Exact prior. The Biased prior exhibited the widest interquartile range of the three, while the Incorrect and Exact priors had comparable box widths. Outliers appeared above the upper whiskers for the Incorrect and Exact priors, reaching approximately 8.8, whereas no outliers were observed for the Biased prior. The lower whiskers extended to similar values across all three conditions, reaching approximately 2.8 to 3.0.

Fig 20.D presents violin plots of the Effective Sample Size for the Simple PK model across three prior types: Incorrect, Biased, and Exact. The Exact prior exhibited the highest median ESS at approximately 1500, while the Incorrect and Biased priors had lower medians near 1075 and 1125, respectively. The Exact prior also displayed a notably different distributional shape, with data points concentrated in two compact clusters around 1250 and 1550, and a narrower overall spread. The Incorrect and Biased priors showed wider violins with data points more dispersed across the range, and their lower tails extended to approximately 400 and 250, respectively, compared to the Exact prior's lower extreme near 1100. All three prior types had upper extremes reaching approximately 2550.

ICG PBPK

Parameter	Group	Prior Type	μ	σ
Bw	Male	Exact	75	10
		Biased	60	20
		Incorrect	10	10
	Female	Exact	65	10
		Biased	50	20
		Incorrect	30	10
$V_{\max}$	Male	Exact	0.037	0.01
		Biased	0.1	0.1
		Incorrect	1	0.01
	Female	Exact	0.029	0.01
		Biased	0.09	0.1
		Incorrect	0.8	0.01

Table 4: **Description of Prior types for the ICG PBPK model.** *Exact* is the same distribution as the true parameter distribution, *Biased* is a partially informed prior to be close to *Exact* and *Incorrect* is a misleading prior to be quite distant from *Exact*.

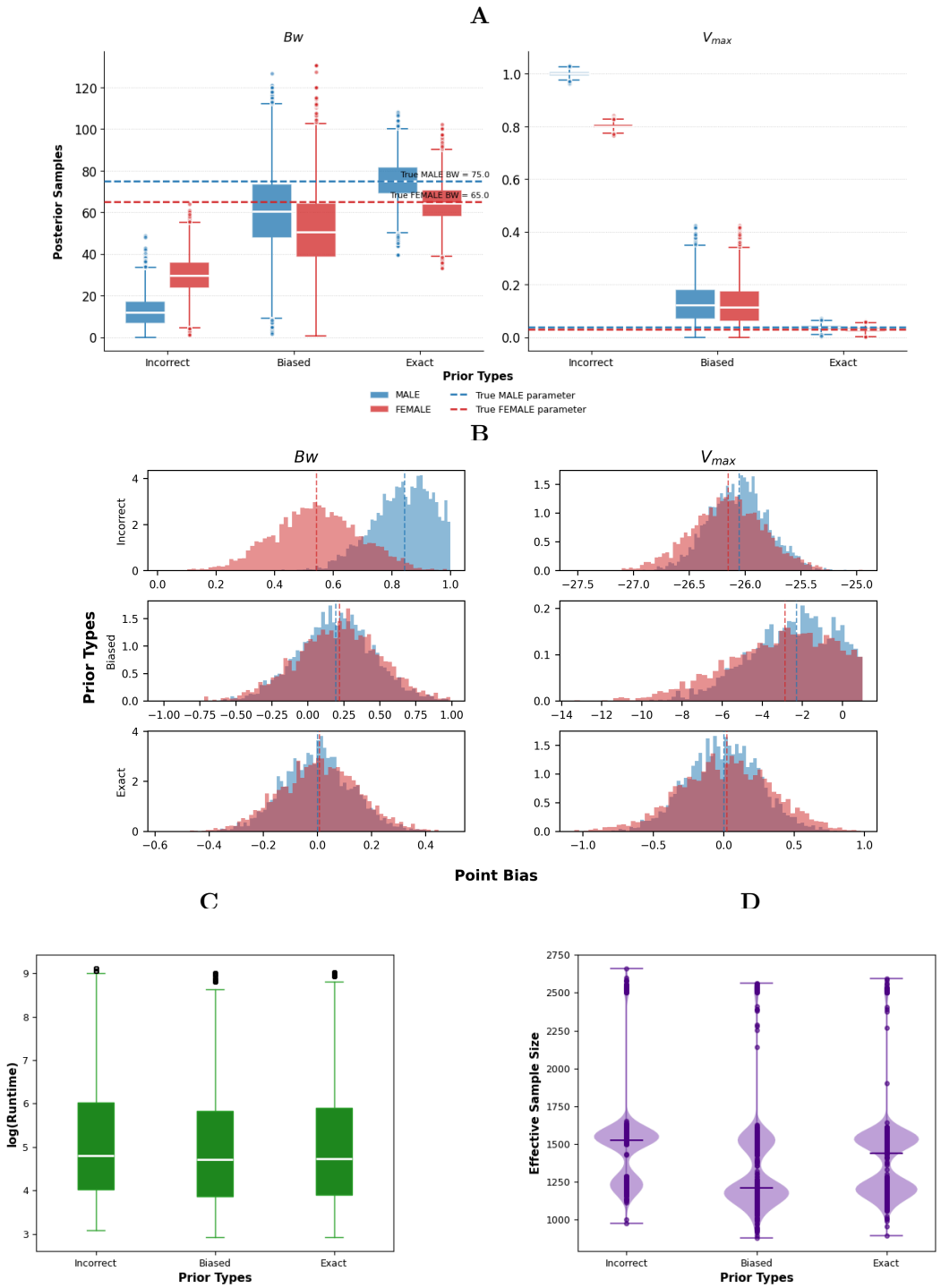

Figure 21: **Effect of varying the prior types for the ICG PBPK model**. The prior types for which data was generated were varied with prior type Exact (no bias), Biased and Incorrect. For A and B subset of results are analysed with 5 time points, 5 samples and a CV of 0.1 (n=3). For the analytic panels C and D subset of results are analysed with all prior types, time points and samples (n=2160). **(A)** Box plot of the posterior samples of Bw and $V_{\max}$. In dashed lines, the true parameter value for MALE and FEMALE. **(B)** Histogram of point biases between posterior samples and the true parameter of Bw and $V_{\max}$. **(C)** Box plot of optimization runtime. **(D)** Effective Sample Size violin plot of the Bayesian sampler chains (chain length 5000).

Fig 21.A shows the posterior sample distributions for parameters Bw (left panel) and $V_{\max}$ (right panel) in the ICG PBPK model across three prior types: Exact, Biased, and Incorrect. For Bw, the Exact prior produced medians near the true values for both groups, while the Biased prior shifted both medians downward to approximately 55, and the Incorrect prior further reduced them to approximately 10 for males and 30 for females. In contrast, $V_{\max}$ showed medians close to the true values under the Exact prior, elevated medians under the Biased prior with wider interquartile ranges, and substantially higher medians under the Incorrect prior — reaching

approximately 0.95 for males and 0.80 for females. For both parameters, the Incorrect prior produced the largest deviation of the medians from the true values, with the male and female distributions becoming more separated compared to the Exact and Biased conditions. The interquartile ranges varied notably across prior types, with Bw displaying the widest spread under the Exact prior and $V_{\max}$ under the Biased prior.

Fig 21.B presents the point bias distributions for parameters Bw (left panels) and $V_{\max}$ (right panels) in the ICG PBPK model under three prior types: Exact, Biased, and Incorrect. With the Exact prior, both parameters yielded distributions centered near zero for both male (blue) and female (red) groups, with Bw spanning approximately -0.6 to 0.4 and $V_{\max}$ ranging from roughly -1.0 to 1.0. Under the Biased prior, the Bw distributions broadened to approximately -1.0 to 1.0 while remaining near zero, whereas $V_{\max}$ shifted considerably toward negative values with peaks around -4 to -5. The Incorrect prior produced the most pronounced deviations from zero for both parameters: Bw distributions were displaced entirely into positive territory, with the male group peaking near 0.85 and the female near 0.5, while $V_{\max}$ distributions were shifted far into negative values, centered around -26. Notably, the Incorrect prior also introduced a clear separation between the male and female distributions for Bw, while for $V_{\max}$ the two groups remained largely overlapping despite the extreme shift from zero.

Fig 21.C shows boxplots of the log-transformed runtime of the Bayesian MCMC sampler for the ICG PBPK model across three prior types: Incorrect, Biased, and Exact. The medians were nearly identical for all three conditions, situated around 4.8. The interquartile ranges were comparable across the three prior types, spanning from roughly 4.0 to 6.0. Outliers were observed above the upper whiskers for all three conditions, reaching approximately 9.0. The lower whiskers extended to approximately 3.0 for each prior type.

Fig 21.D illustrates violin plots of the Effective Sample Size for the ICG PBPK model across three prior types: Incorrect, Biased, and Exact. The Incorrect prior had the highest median ESS at approximately 1550, followed by the Exact prior near 1450, and the Biased prior with the lowest median around 1200. The Incorrect prior displayed a relatively compact main distribution with a secondary cluster near 2500, while the Biased prior exhibited a wider spread of data points between approximately 900 and 1600, with scattered points extending upward to 2550. The Exact prior showed the broadest violin shape, with data points distributed across a range from approximately 900 to 2600. The lower extremes were comparable for all three conditions, reaching approximately 900 to 1000.

4. Discussion

The results presented above reveal that, among the four hyperparameters examined, the type of prior had the most pronounced effect on the accuracy of the posterior estimates across all three models. While the number of timepoints, the number of samples, and the coefficient of variation had limited impact on the posterior distributions and point bias, the choice of prior – particularly the Biased and Incorrect priors – led to substantial deviations from the true parameter values. This was observed consistently in the Simple Chain, Simple PK, and ICG PBPK models, with the magnitude of the displacement increasing progressively from the Exact to the Biased to the Incorrect prior.

In contrast, the number of timepoints and samples primarily influenced the computational cost, as reflected by the steady increase in log(Runtime), without yielding notable changes in either the posterior accuracy or the Effective Sample Size. The coefficient of variation similarly had minimal impact across all metrics examined. The stability of the ESS across most hyperparameter conditions indicates that the MCMC sampler maintained consistent performance regardless of the data configuration, with the exception of specific conditions such as the highest coefficient of variation level in the ICG PBPK model and certain prior types in the Simple PK model.

4.1. Prior Type

A notable finding is the difference in how the three models responded to the prior specification. In the Simple Chain and Simple PK models, the Biased and Incorrect priors shifted the posterior distributions and point bias away from the true values in a progressive manner, with both male and female groups affected. The posterior boxplots showed that under the Exact prior, medians for both groups were positioned close to their true values, while the Biased prior elevated these estimates, and the Incorrect prior displaced them further still.

This pattern was consistent across both parameters in the Simple PK model, where CL and k medians moved increasingly away from their reference values under each successive level of prior misspecification. In the ICG PBPK model, the Incorrect prior produced particularly striking effects: the Bw parameter was displaced entirely into positive territory relative to the true values, while $V_{\{\max\}}$ estimates were shifted to values far above the reference, reaching approximately 0.95 for males and 0.80 for females. This suggests that in more complex, physiologically informed models, misspecified priors may propagate through the model structure in ways that differ from simpler formulations, potentially affecting parameters in opposing directions.

In contrast to the marked influence of the prior type on posterior accuracy, the remaining three hyperparameters — number of timepoints, number of samples, and coefficient of variation — had limited effects on the posterior distributions across all three models, when the prior structure is *Exact*. The posterior boxplots for these hyperparameters showed that the medians and interquartile ranges remained relatively stable regardless of the values tested. In the Simple Chain model, the median estimates for k_1 were consistently near the true values for both groups across all timepoint, sample size, and coefficient of variation levels. A similar pattern was observed in the Simple PK and ICG PBPK models, where neither increasing the number of observations nor varying the noise level produced notable shifts in the posterior estimates.

The only exception was a slight widening of the interquartile ranges at higher coefficient of variation values in the Simple Chain and Simple PK models, though the central tendency of the posteriors remained largely unaffected. These observations indicate that, within the ranges examined, prior specification was the dominant factor determining the accuracy of the parameter recovery, while the data configuration hyperparameters played a comparatively minor role.

4.2. Point Bias

The point bias distributions further corroborated the findings from the posterior boxplots, providing a more detailed view of the relative distance between the posterior samples and the true parameter values. In the Simple Chain and Simple PK models, the distributions were consistently asymmetric, exhibiting long left tails extending to values as extreme as -8 or -16 under certain conditions, while being bounded near 1.0 on the right side. This indicates that overestimation of the true parameter values was constrained while underestimation was not. In the Simple PK model, a systematic pattern emerged in which the parameter with the smaller true value (0.5) — CL for females and k for males — consistently exhibited greater displacement from zero in the point bias distributions, while the parameter with the larger true value (1.0) was more accurately recovered. This asymmetry was present regardless of which hyperparameter was being varied. This constraint is based on the prior definitions of PyPesto. As parameters can only be positive, this limit allows the parameters to always be interpretable.

The ICG PBPK model yielded markedly different point bias distributions compared to the other two models. For both Bw and $V_{\max}$, the distributions were approximately symmetric and centered near zero under the Exact prior, with the male and female groups nearly indistinguishable across all timepoint, sample size, and coefficient of variation conditions. This symmetry may reflect the additional physiological constraints embedded within the model structure. However, when the prior type was varied, the point bias distributions for the ICG PBPK model revealed pronounced deviations: under the Biased prior, $V_{\max}$ shifted substantially toward negative values with peaks around -4 to -5, while under the Incorrect prior the displacement was extreme, with the distribution centered around -26. For Bw, the Incorrect prior shifted the distributions entirely into positive territory. These results demonstrate that while the ICG PBPK model produced well-centered point bias distributions under favorable prior conditions, it was equally susceptible to prior misspecification as the simpler models, with the magnitude of the resulting bias being considerable.

4.3. Runtime Considerations

The log-transformed runtime increased steadily with both the number of timepoints and the number of samples across all three models. In the Simple Chain model, the median log(Runtime) rose from approximately 3.2 at the lowest levels to approximately 5.6–6.1 at the highest levels of timepoints and samples, respectively. The Simple PK model exhibited a similar trajectory, with median values ranging from approximately 3.5 to 6.0–6.5, and the ICG PBPK model showed the highest runtimes overall, with medians increasing from approximately 3.9 to 6.5–6.9. The interquartile ranges also expanded at higher timepoint and sample levels across all models, indicating greater variability in execution time. These results highlight a practical consideration when scaling Bayesian PBPK approaches: while adding more timepoints or samples does not appear to improve posterior accuracy or MCMC diagnostics within the ranges examined, it does carry a substantial computational cost that grows with model complexity.

In contrast, neither the coefficient of variation nor the prior type had a discernible effect on the runtime across any of the three models. When the coefficient of variation was varied from 0.0 to 1.0, the median log(Runtime) remained approximately constant for the Simple Chain, Simple PK, and ICG PBPK models, with stable interquartile ranges at each level. Similarly, the three prior types — Exact, Biased, and Incorrect — produced nearly identical runtime distributions within each model, with closely aligned medians and comparable spreads. Outliers were present at several coefficient of variation and prior type conditions across all models, reaching values near 8.0–9.0, but these occurred without any systematic pattern tied to a specific hyperparameter value. These findings indicate that the computational burden of the Bayesian MCMC sampler was driven primarily by the volume of data — as determined by the number of timepoints and samples — rather than by the noise characteristics or the specification of the prior distributions.

4.4. Sampling Diagnostics

The Effective Sample Size remained largely stable across the number of timepoints, number of samples, and coefficient of variation for all three models. In the Simple Chain model, the median ESS was approximately 1700 and did not vary across any of these hyperparameter conditions, with data points consistently distributed in two clusters: a primary concentration near the median and a secondary grouping near 2550. The Simple PK model exhibited a lower median ESS around 1100 with more dispersed data points, while the ICG PBPK model had a median near 1550 with a similar two-cluster pattern to the Simple Chain. The shape and spread of the violin plots remained consistent across the range of values tested for each of these three hyperparameters, with the exception of the ICG PBPK model at a coefficient of variation of 1.0, where the ESS was noticeably lower, the median dropped to approximately 1400, and the secondary cluster of high ESS values near 2500 disappeared.

The prior type had a more varied effect on the ESS across the three models. In the Simple Chain model, the ESS distributions were similar for all three prior types, with only a marginally higher median for the Exact prior. In the Simple PK model, the differences were more pronounced: the Exact prior produced a higher and more compact ESS distribution with a median near 1500, while the Incorrect and Biased priors had lower medians around 1075 and 1125, respectively, with data points extending to lower extremes near 400 and 250. The ICG PBPK model showed a different pattern, with the Incorrect prior yielding the highest median ESS near 1550, the Exact prior near 1450, and the Biased prior the lowest at approximately 1200, accompanied by the widest spread of data points. These observations indicate that while the MCMC sampler generally maintained consistent performance regardless of the data configuration, specific combinations of model complexity and prior misspecification can affect sampling efficiency in model-dependent ways.

Taken together, these findings underscore the importance of prior selection in Bayesian PBPK modeling, as it was the only hyperparameter that substantially influenced the accuracy of the posterior estimates. The relative insensitivity of the results to the number of timepoints, samples, and coefficient of variation — within the ranges tested — suggests that the approach is robust to variations in data configuration, but remains vulnerable to inadequate prior knowledge. Future work could explore strategies for prior elicitation and sensitivity analysis to mitigate the impact of prior misspecification in more complex PBPK frameworks.

5. Conclusion

This work investigated the effects of four hyperparameters — number of timepoints, number of samples, coefficient of variation, and prior type — on the performance of a Bayesian approach applied to three Pharmacokinetic models of increasing complexity: the Simple Chain, the Simple PK, and the ICG PBPK models. The performance was evaluated in terms of posterior accuracy, point bias, computational cost, and MCMC sampling efficiency as measured by the Effective Sample Size.

The results demonstrated that, among the four hyperparameters examined, the type of prior had the most substantial impact on the accuracy of the posterior estimates. Under the Exact prior, all three models recovered parameter values close to the true values for both male and female groups. The Biased and Incorrect priors progressively displaced the posterior distributions away from the true values, with the Incorrect prior producing the most severe deviations across all models. The point bias distributions confirmed these findings, revealing that the displacement from the true values increased in magnitude from the Exact to the Biased to the Incorrect prior. Notably, the ICG PBPK model — despite producing well-centered and symmetric point bias distributions under favorable prior conditions — was equally susceptible to prior misspecification, with extreme shifts observed for both Bw and $V_{\max}$ under the Incorrect prior.

In contrast, the number of timepoints, the number of samples, and the coefficient of variation had limited influence on the posterior accuracy and point bias across all three models. The posterior distributions and point bias histograms remained largely unchanged as these hyperparameters were varied within the tested ranges. However, both the number of timepoints and the number of samples had a direct effect on the computational cost, with the log-transformed runtime increasing steadily as either hyperparameter grew. The ICG PBPK model consistently required the highest runtimes, followed by the Simple PK and Simple Chain models. Neither the coefficient of variation nor the prior type affected the runtime in any of the three models. The Effective Sample Size was stable across most conditions, with isolated exceptions observed in the ICG PBPK model at the highest coefficient of variation level and in the Simple PK model under misspecified priors.

Several structural differences between the models were observed. The Simple Chain and Simple PK models produced asymmetric point bias distributions with long left tails, whereas the ICG PBPK model yielded approximately symmetric distributions under the Exact prior. Additionally, in the Simple PK model, parameters with smaller true values were consistently less accurately recovered than those with larger true values, a pattern that persisted across all hyperparameter conditions. These model-specific behaviors highlight the role that model structure and parameterization play in shaping the outcomes of Bayesian inference.

In summary, this work identified prior specification as the primary determinant of posterior accuracy in Bayesian PBPK modeling, while the data configuration hyperparameters — within the ranges tested — had a comparatively minor role. The computational cost scaled with the volume of data but was unaffected by the noise level or prior choice. These findings carry practical implications for the design of Bayesian PBPK workflows: careful attention to prior elicitation is essential, as misspecified priors can lead to substantial bias in parameter estimates regardless of the amount or quality of the available data. Future work could explore adaptive prior selection strategies, prior sensitivity analyses, and the extension of this framework to additional PBPK models with greater physiological detail, to further assess the generalizability and robustness of these findings.

Bibliography

[1] E. P. A. O. T. Innovation, "Pharmacokinetics." Accessed: Mar. 01, 2026. [Online]. Available: https://toolbox.eupati.eu/glossary/pharmacokinetics/

[2] H. Jones and K. Rowland-Yeo, "Basic Concepts in Physiologically Based Pharmacokinetic Modeling in Drug Discovery and Development," *CPT: Pharmacometrics & Systems Pharmacology*, vol. 2, no. 8, pp. 1–12, Aug. 2013, doi: 10.1038/psp.2013.41.

[3] C. X. Zhang and S. L. Arnold, "Potential and Challenges in Application of Physiologically Based Pharmacokinetic Modeling in Predicting Diarrheal Disease Impact on Oral Drug Pharmacokinetics," *Drug Metabolism and Disposition*, vol. 53, no. 1, p. 100014, Jan. 2025, doi: 10.1124/dmd.122.000964.

[4] F. Chen *et al.*, "Physiologically Based Pharmacokinetic Modeling to Understand the Absorption of Risperidone Orodispersible Film," *Frontiers in Pharmacology*, vol. 10, p. 1692, Feb. 2020, doi: 10.3389/fphar.2019.01692.

[5] O. Ozbek, D. E. Genc, and K. O. Ulgen, "Advances in Physiologically Based Pharmacokinetic (PBPK) Modeling of Nanomaterials," *ACS Pharmacology & Translational Science*, vol. 7, no. 8, pp. 2251–2279, Aug. 2024, doi: 10.1021/acsptsci.4c00250.

[6] Y.-C. Zhao *et al.*, "A Preliminary Exploration of Liver Microsomes and PBPK to Uncover the Impact of CYP3A4/5 and CYP2C19 on Tacrolimus and Voriconazole Drug-Drug Interactions," *Scientific Reports*, vol. 15, no. 1, p. 6389, Feb. 2025, doi: 10.1038/s41598-025-91356-7.

[7] B. Hossam Abdelmonem *et al.*, "Decoding the Role of CYP450 Enzymes in Metabolism and Disease: A Comprehensive Review," *Biomedicines*, vol. 12, no. 7, p. 1467, July 2024, doi: 10.3390/biomedicines12071467.

[8] S. Li, X. Ye, Q. Wang, Z. Cheng, F. Liu, and F. Xie, "Is the GFR-based Scaling Approach Adequate for Predicting Pediatric Renal Clearance of Drugs with Passive Tubular Reabsorption? Insights from PBPK Modeling," *CPT: Pharmacometrics & Systems Pharmacology*, vol. 14, no. 1, pp. 152–163, Jan. 2025, doi: 10.1002/psp4.13254.

[9] D. Deepika and V. Kumar, "The Role of "Physiologically Based Pharmacokinetic Model (PBPK)" New Approach Methodology (NAM) in Pharmaceuticals and Environmental Chemical Risk Assessment," *International Journal of Environmental Research and Public Health*, vol. 20, no. 4, p. 3473, Feb. 2023, doi: 10.3390/ijerph20043473.

[10] T. N. Tozer and M. Rowland, *Essentials of Pharmacokinetics and Pharmacodynamics*, Second edition. Philadelphia: Wolters Kluwer, 2016.

[11] T. Saeheng, K. Na-Bangchang, and J. Karbwang, "Utility of Physiologically Based Pharmacokinetic (PBPK) Modeling in Oncology Drug Development and Its Accuracy: A Systematic Review," *European Journal of Clinical Pharmacology*, vol. 74, no. 11, pp. 1365–1376, Nov. 2018, doi: 10.1007/s00228-018-2513-6.

[12] N. Marsousi, J. A. Desmeules, S. Rudaz, and Y. Daali, "Usefulness of PBPK Modeling in Incorporation of Clinical Conditions in Personalized Medicine," *Journal of Pharmaceutical Sciences*, vol. 106, no. 9, pp. 2380–2391, Sept. 2017, doi: 10.1016/j.xphs.2017.04.035.

[13] R. Ota and F. Yamashita, "Application of Machine Learning Techniques to the Analysis and Prediction of Drug Pharmacokinetics," *Journal of Controlled Release*, vol. 352, pp. 961–969, Dec. 2022, doi: 10.1016/j.jconrel.2022.11.014.

[14] Y. Arav, "Advances in Modeling Approaches for Oral Drug Delivery: Artificial Intelligence, Physiologically-Based Pharmacokinetics, and First-Principles Models," *Pharmaceutics*, vol. 16, no. 8, p. 978, July 2024, doi: 10.3390/pharmaceutics16080978.

[15] K. Rowland Yeo, E. Gil Berglund, and Y. Chen, "Dose Optimization Informed by PBPK Modeling: State-of-the Art and Future," *Clinical Pharmacology & Therapeutics*, vol. 116, no. 3, pp. 563–576, Sept. 2024, doi: 10.1002/cpt.3289.

[16] H. Huang, W. Zhao, N. Qin, and X. Duan, "Recent Progress on Physiologically Based Pharmacokinetic (PBPK) Model: A Review Based on Bibliometrics," *Toxics*, vol. 12, no. 6, p. 433, June 2024, doi: 10.3390/toxics12060433.

[17] L. Aarons, "Physiologically Based Pharmacokinetic Modelling: A Sound Mechanistic Basis Is Needed," *British Journal of Clinical Pharmacology*, vol. 60, no. 6, pp. 581–583, Dec. 2005, doi: 10.1111/j.1365-2125.2005.02560.x.

[18] H. Jones and K. Rowland-Yeo, "Basic Concepts in Physiologically Based Pharmacokinetic Modeling in Drug Discovery and Development," *CPT: Pharmacometrics & Systems Pharmacology*, vol. 2, no. 8, pp. 1–12, Aug. 2013, doi: 10.1038/psp.2013.41.

[19] W. Zhang, Q. Zhang, Z. Cao, L. Zheng, and W. Hu, "Physiologically Based Pharmacokinetic Modeling in Neonates: Current Status and Future Perspectives," *Pharmaceutics*, vol. 15, no. 12, p. 2765, Dec. 2023, doi: 10.3390/pharmaceutics15122765.

[20] M. Le Merdy, K. X. Szeto, J. Perrier, M. B. Bolger, and V. Lukacova, "PBPK Modeling Approach to Predict the Behavior of Drugs Cleared by Metabolism in Pregnant Subjects and Fetuses," *Pharmaceutics*, vol. 16, no. 1, p. 96, Jan. 2024, doi: 10.3390/pharmaceutics16010096.

[21] Y. Cao and W. J. Jusko, "Applications of Minimal Physiologically-Based Pharmacokinetic Models," *Journal of Pharmacokinetics and Pharmacodynamics*, vol. 39, no. 6, pp. 711–723, Dec. 2012, doi: 10.1007/s10928-012-9280-2.

[22] D. M. Foster and P. Vicini, "Noncompartmental and Compartmental Approaches to Pharmacokinetic Data Analysis," *Atkinson's Principles of Clinical Pharmacology*. Elsevier, pp. 113–135, 2022. doi: 10.1016/B978-0-12-819869-8.00035-5.

[23] P. Chou, A. Shannar, Y. Pan, P. D. Dave, J. Xu, and A.-N. T. Kong, "Application of Physiologically-Based Pharmacokinetic (PBPK) Model in Drug Development and in Dietary Phytochemicals," *Current Pharmacology Reports*, vol. 11, no. 1, p. 45, Aug. 2025, doi: 10.1007/s40495-025-00427-w.

[24] M. Ibarra, M. Vázquez, and P. Fagiolino, "Current PBPK Models: Are They Predicting Tissue Drug Concentration Correctly?," *Drugs in R&D*, vol. 20, no. 4, pp. 295–299, Dec. 2020, doi: 10.1007/s40268-020-00325-0.

[25] P.-J. De Sutter, E. Gasthuys, and A. Vermeulen, "Comparison of Monoclonal Antibody Disposition Predictions Using Different Physiologically Based Pharmacokinetic Modelling Platforms," *Journal of Pharmacokinetics and Pharmacodynamics*, vol. 51, no. 6, pp. 639–651, Dec. 2024, doi: 10.1007/s10928-023-09894-4.

[26] N. Nowak, S. E. Escher, and K. Schwarz, "A Unified Whole Lung \textsc{PBK} Model for Inhalational Uptake of Gases and Aerosols in Men," *CPT: Pharmacometrics & Systems Pharmacology*, vol. 14, no. 12, pp. 2173–2185, Dec. 2025, doi: 10.1002/psp4.70117.

[27] R. Zang *et al.*, "Design and Measurement of Drug Tissue Concentration Asymmetry and Tissue Exposure-Effect (Tissue PK-PD) Evaluation," *Journal of Medicinal Chemistry*, vol. 65, no. 13, pp. 8713–8734, July 2022, doi: 10.1021/acs.jmedchem.2c00502.

[28] B. Mukherjee, "Fundamentals of Pharmacokinetics," in *Pharmacokinetics: Basics to Applications*, Singapore: Springer Singapore, 2022, pp. 1–20. doi: 10.1007/978-981-16-8950-5_1.

[29] A. Atkinson, *Principles of Clinical Pharmacology*. S.l.: Elsevier Academic Press, 2021.

[30] A. Kreutz, X. Chang, H. T. Hogberg, and B. A. Wetmore, "Advancing Understanding of Human Variability through Toxicokinetic Modeling, in Vitro-in Vivo Extrapolation, and New Approach Methodologies," *Human Genomics*, vol. 18, no. 1, p. 129, Nov. 2024, doi: 10.1186/s40246-024-00691-9.

[31] C. M. Schacht *et al.*, "Evaluating the Impact of Anatomical and Physiological Variability on Human Equivalent Doses Using PBPK Models," *Toxicological Sciences*, vol. 200, no. 2, pp. 241–264, Aug. 2024, doi: 10.1093/toxsci/kfae067.

[32] T. Wolff and M. Strecker, "Endogenous and Exogenous Factors Modifying the Activity of Human Liver Cytochrome P-450 Enzymes," *Experimental and Toxicologic Pathology*, vol. 44, no. 5, pp. 263–271, Jan. 1992, doi: 10.1016/S0940-2993(11)80241-1.

[33] J. Tibbitts, D. Canter, R. Graff, A. Smith, and L. A. Khawli, "Key Factors Influencing ADME Properties of Therapeutic Proteins: A Need for ADME Characterization in Drug Discovery and Development," *mAbs*, vol. 8, no. 2, pp. 229–245, Feb. 2016, doi: 10.1080/19420862.2015.1115937.

[34] Y. S. Lin, K. E. Thummel, B. D. Thompson, R. A. Totah, and C. W. Cho, "Sources of Interindividual Variability," *Enzyme Kinetics in Drug Metabolism*, vol. 2342. Springer US, New York, NY, pp. 481–550, 2021. doi: 10.1007/978-1-0716-1554-6_17.

[35] M. A. Hedaya, *Basic Pharmacokinetics*, 3rd ed. New York: Routledge, 2023. doi: 10.4324/9781003161523.

[36] H. Chandasana, M. Bush, M. Ait-Khaled, B. Wynne, S. Min, and R. Mehta, "Population Pharmacokinetic Analysis of Dolutegravir in Treatment-Experienced Adults Living with HIV-1," *The Journal of Clinical Pharmacology*, vol. 64, no. 11, pp. 1407–1418, Nov. 2024, doi: 10.1002/jcph.2494.

[37] R. Li, E. Kimoto, Y.-A. Bi, D. Tess, and M. V. S. Varma, "Physiologically Based Pharmacokinetic Model of OATP1B Substrates with a Nonlinear Mixed Effect Approach: Estimating Empirical In Vitro-to-In Vivo Scaling Factors," *Clinical Pharmacokinetics*, vol. 63, no. 8, pp. 1177–1189, Aug. 2024, doi: 10.1007/s40262-024-01408-w.

[38] I. Fornacon-Wood, H. Mistry, C. Johnson-Hart, C. Faivre-Finn, J. P. O'Connor, and G. J. Price, "Understanding the Differences Between Bayesian and Frequentist Statistics," *International Journal of Radiation Oncology*Biology*Physics*, vol. 112, no. 5, pp. 1076–1082, Apr. 2022, doi: 10.1016/j.ijrobp.2021.12.011.

[39] A. Hazra, "Using the Confidence Interval Confidently," *Journal of Thoracic Disease*, vol. 9, no. 10, pp. 4124–4129, Oct. 2017, doi: 10.21037/jtd.2017.09.14.

[40] Y. Xiao, L. Xu, Y. Qian, and Y. Xu, "Identification and Characterization of Critical Values in Therapeutic Drug Monitoring: A Retrospective Analysis," *Scientific Reports*, vol. 14, no. 1, p. 11520, May 2024, doi: 10.1038/s41598-024-62402-7.

[41] S. J. Ruberg *et al.*, "Application of Bayesian Approaches in Drug Development: Starting a Virtuous Cycle," *Nature Reviews Drug Discovery*, vol. 22, no. 3, pp. 235–250, Mar. 2023, doi: 10.1038/s41573-023-00638-0.

[42] S. Greenhill, S. Rana, S. Gupta, P. Vellanki, and S. Venkatesh, "Bayesian Optimization for Adaptive Experimental Design: A Review," *IEEE Access*, vol. 8, pp. 13937–13948, 2020, doi: 10.1109/ACCESS.2020.2966228.

[43] W. C. Roda, "Bayesian Inference for Dynamical Systems," *Infectious Disease Modelling*, vol. 5, pp. 221–232, Jan. 2020, doi: 10.1016/j.idm.2019.12.007.

[44] Z. Li, "A Review of Bayesian Posterior Distribution Based on MCMC Methods," *Computing and Data Science*, vol. 1513. Springer Nature Singapore, Singapore, pp. 204–213, 2021. doi: 10.1007/978-981-16-8885-0_17.

[45] E. W. Rognli, R. Zahl-Olsen, S. S. Rekdal, A. Hoffart, and T. B. Bertelsen, "Editorial Perspective: Bayesian Statistical Methods Are Useful for Researchers in Child and Adolescent Mental Health," *Journal of Child Psychology and Psychiatry*, vol. 64, no. 2, pp. 339–342, Feb. 2023, doi: 10.1111/jcpp.13662.

[46] J. R. Wedagedera *et al.*, "Population \textsc{PBPK} Modeling Using Parametric and Nonparametric Methods of the Simcyp Simulator, and Bayesian Samplers," *CPT: Pharmacometrics & Systems Pharmacology*, vol. 11, no. 6, pp. 755–765, June 2022, doi: 10.1002/psp4.12787.

[47] D. Hardiansyah, B. B. Patrianesha, K. Shi, B. Saboury, A. Rahmim, and G. Glatting, "An Overview of Physiologically Based Pharmacokinetic (PBPK) and Population Pharmacokinetic (PopPK) Models." Accessed: Feb. 22, 2026. [Online]. Available: https://arxiv.org/abs/2509.09755

[48] W. A. Chiu *et al.*, "Bayesian Estimation of Human Population Toxicokinetics of PFOA, PFOS, PFHxS, and PFNA from Studies of Contaminated Drinking Water," *Environmental Health Perspectives*, vol. 130, no. 12, p. 127001, Dec. 2022, doi: 10.1289/EHP10103.

[49] J. Barido-Sottani, O. Schwery, R. C. M. Warnock, C. Zhang, and A. M. Wright, "Practical Guidelines for Bayesian Phylogenetic Inference Using Markov Chain Monte Carlo (MCMC)," *Open Research Europe*, vol. 3, p. 204, Aug. 2024, doi: 10.12688/openreseurope.16679.3.

[50] A. Elmokadem, Y. Zhang, T. Knab, E. Jordie, and W. Gillespie, "Bayesian \textsc{PBPK} Modeling Using R/Stan/Torsten and Julia/ \textsc{SciML} /Turing.Jl," *CPT: Pharmacometrics & Systems Pharmacology*, vol. 12, no. 3, pp. 300–310, Mar. 2023, doi: 10.1002/psp4.12926.

[51] L. F. South, M. Riabiz, O. Teymur, and C. J. Oates, "Postprocessing of MCMC," *Annual Review of Statistics and Its Application*, vol. 9, no. 1, pp. 529–555, Mar. 2022, doi: 10.1146/annurev-statistics-040220-091727.

[52] K. Tiwari *et al.*, "Reproducibility in Systems Biology Modelling," *Molecular Systems Biology*, vol. 17, no. 2, p. e9982, Feb. 2021, doi: 10.15252/msb.20209982.

[53] E. Domínguez-Romero *et al.*, "Making PBPK Models More Reproducible in Practice," *Briefings in Bioinformatics*, vol. 25, no. 6, p. bbae569, Sept. 2024, doi: 10.1093/bib/bbae569.

[54] W. Lin, Y. Chen, J. D. Unadkat, X. Zhang, D. Wu, and T. Heimbach, "Applications, Challenges, and Outlook for PBPK Modeling and Simulation: A Regulatory, Industrial and Academic Perspective," *Pharmaceutical Research*, vol. 39, no. 8, pp. 1701–1731, Aug. 2022, doi: 10.1007/s11095-022-03274-2.

[55] S. Frechen and A. Rostami-Hodjegan, "Quality Assurance of PBPK Modeling Platforms and Guidance on Building, Evaluating, Verifying and Applying PBPK Models Prudently under the Umbrella of Qualification: Why, When, What, How and By Whom?," *Pharmaceutical Research*, vol. 39, no. 8, pp. 1733–1748, Aug. 2022, doi: 10.1007/s11095-022-03250-w.

[56] L. Vogt *et al.*, "FAIR 2.0: Extending the FAIR Guiding Principles to Address Semantic Interoperability," *Scientific Data*, vol. 12, no. 1, p. 688, Apr. 2025, doi: 10.1038/s41597-025-05011-x.

[57] M. König, "Matthiaskoenig/Parameter-Variability." Accessed: Mar. 15, 2026. [Online]. Available: https://github.com/matthiaskoenig/parameter-variability

[58] A. Köller, J. Grzegorzewski, and M. König, "Physiologically Based Modeling of the Effect of Physiological and Anthropometric Variability on Indocyanine Green Based Liver Function Tests," *Frontiers in Physiology*, vol. 12, p. 757293, Nov. 2021, doi: 10.3389/fphys.2021.757293.

[59] P. Sharma, "Value of Liver Function Tests in Cirrhosis," *Journal of Clinical and Experimental Hepatology*, vol. 12, no. 3, pp. 948–964, May 2022, doi: 10.1016/j.jceh.2021.11.004.

[60] E. Levesque, E. Martin, D. Dudau, C. Lim, G. Dhonneur, and D. Azoulay, "Current Use and Perspective of Indocyanine Green Clearance in Liver Diseases," *Anaesthesia Critical Care & Pain Medicine*, vol. 35, no. 1, pp. 49–57, Feb. 2016, doi: 10.1016/j.accpm.2015.06.006.

[61] H. O. Wheeler, W. I. Cranston, and J. I. Meltzer, "Hepatic Uptake and Biliary Excretion of Indocyanine Green in the Dog.," *Experimental Biology and Medicine*, vol. 99, no. 1, pp. 11–14, Oct. 1958, doi: 10.3181/00379727-99-24229.

[62] A. Köller, J. Grzegorzewski, H.-M. Tautenhahn, and M. König, "Prediction of Survival After Partial Hepatectomy Using a Physiologically Based Pharmacokinetic Model of Indocyanine Green Liver Function Tests," *Frontiers in Physiology*, vol. 12, p. 730418, Nov. 2021, doi: 10.3389/fphys.2021.730418.

[63] T. Radonjić *et al.*, "Aging of Liver in Its Different Diseases," *International Journal of Molecular Sciences*, vol. 23, no. 21, p. 13085, Oct. 2022, doi: 10.3390/ijms232113085.

[64] A. Kasarinaite, M. Sinton, P. T. K. Saunders, and D. C. Hay, "The Influence of Sex Hormones in Liver Function and Disease," *Cells*, vol. 12, no. 12, p. 1604, June 2023, doi: 10.3390/cells12121604.

[65] C. Roberts, L. Jackson, M. Halliwell, and R. Branch, "The Relationship between Liver Volume, Antipyrine Clearance and Indocyanine Green Clearance before and after Phenobarbitone Administration in Man.," *British Journal of Clinical Pharmacology*, vol. 3, no. 5, pp. 907–913, Oct. 1976, doi: 10.1111/j.1365-2125.1976.tb00646.x.

[66] M. Guidi, C. Csajka, and T. Buclin, "Parametric Approaches in Population Pharmacokinetics," *The Journal of Clinical Pharmacology*, vol. 62, no. 2, pp. 125–141, 2022, doi: 10.1002/jcph.1633.

[67] L. F. Lacey, O. N. Keene, J. F. Pritchard, and A. Bye, "Common Noncompartmental Pharmacokinetic Variables: Are They Normally or Log-Normally Distributed?," *Journal of Biopharmaceutical Statistics*, vol. 7, no. 1, pp. 171–178, Jan. 1997, doi: 10.1080/10543409708835177.

[68] S. M. Keating *et al.*, "SBML Level 3: An Extensible Format for the Exchange and Reuse of Biological Models," *Molecular Systems Biology*, vol. 16, no. 8, p. e9110, Aug. 2020, doi: 10.15252/msb.20199110.

[69] 万紀子清水 and Y. Hiroshi, *OECD Environment, Health and Safety Publications Series on Testing and Assessment No. 331*. OECD, 2021. Accessed: Mar. 07, 2026. [Online]. Available: https://cir.nii.ac.jp/crid/1020013168676971264

[70] R. S. Malik-Sheriff *et al.*, "BioModels—15 Years of Sharing Computational Models in Life Science," *Nucleic Acids Research*, p. gkz1055, Nov. 2019, doi: 10.1093/nar/gkz1055.

[71] V. V. Tomas Radivoyevitch, "SBMLR." Accessed: Mar. 07, 2026. [Online]. Available: https://bioconductor.org/packages/SBMLR

[72] J. L. Hellerstein, L. P. Smith, L. T. Tatka, S. S. Andrews, M. A. Kochen, and H. M. Sauro, "Discovering Subnetworks in SBML Models," *Bioinformatics*, vol. 41, no. 9, p. btaf482, Sept. 2025, doi: 10.1093/bioinformatics/btaf482.

[73] P. F. Lang, S. Shin, and V. M. Zavala, "SBML2Julia: Interfacing SBML with Efficient Nonlinear Julia Modelling and Solution Tools for Parameter Optimization." Accessed: Mar. 07, 2026. [Online]. Available: https://arxiv.org/abs/2011.02597

[74] "Sys-Bio/Roadrunner." Accessed: Mar. 08, 2026. [Online]. Available: https://github.com/sys-bio/roadrunner

[75] C. Welsh, J. Xu, L. Smith, M. König, K. Choi, and H. M. Sauro, "libRoadRunner 2.0: A High Performance SBML Simulation and Analysis Library," *Bioinformatics*, vol. 39, no. 1, p. btac770, Jan. 2023, doi: 10.1093/bioinformatics/btac770.

[76] L. Schmiester *et al.*, "PEtab—Interoperable Specification of Parameter Estimation Problems in Systems Biology," *PLOS Computational Biology*, vol. 17, no. 1, p. e1008646, Jan. 2021, doi: 10.1371/journal.pcbi.1008646.

[77] Y. Schälte *et al.*, "pyPESTO: A Modular and Scalable Tool for Parameter Estimation for Dynamic Models," *Bioinformatics*, vol. 39, no. 11, p. btad711, Nov. 2023, doi: 10.1093/bioinformatics/btad711.

[78] F. Fröhlich and P. K. Sorger, "Fides: Reliable Trust-Region Optimization for Parameter Estimation of Ordinary Differential Equation Models," *PLOS Computational Biology*, vol. 18, no. 7, p. e1010322, July 2022, doi: 10.1371/journal.pcbi.1010322.

[79] S. G. Johnson, "Stevengj/Nlopt." Accessed: Mar. 08, 2026. [Online]. Available: https://github.com/stevengj/nlopt

[80] A. Wächter and L. T. Biegler, "On the Implementation of an Interior-Point Filter Line-Search Algorithm for Large-Scale Nonlinear Programming," *Mathematical Programming*, vol. 106, no. 1, pp. 25–57, Mar. 2006, doi: 10.1007/s10107-004-0559-y.

[81] P. Virtanen *et al.*, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, no. 3, pp. 261–272, Mar. 2020, doi: 10.1038/s41592-019-0686-2.

[82] B. Miasojedow, E. Moulines, and M. Vihola, "An Adaptive Parallel Tempering Algorithm," *Journal of Computational and Graphical Statistics*, vol. 22, no. 3, pp. 649–664, July 2013, doi: 10.1080/10618600.2013.778779.

[83] B. Ballnus, S. Hug, K. Hatz, L. Görlitz, J. Hasenauer, and F. J. Theis, "Comprehensive Benchmarking of Markov Chain Monte Carlo Methods for Dynamical Systems," *BMC Systems Biology*, vol. 11, no. 1, p. 63, June 2017, doi: 10.1186/s12918-017-0433-1.

[84] M. K. Łącki and B. Miasojedow, "State-Dependent Swap Strategies and Automatic Reduction of Number of Temperatures in Adaptive Parallel Tempering Algorithm," *Statistics and Computing*, vol. 26, no. 5, pp. 951–964, Sept. 2016, doi: 10.1007/s11222-015-9579-0.

[85] H. Huang, A. Handel, and X. Song, "A Bayesian Approach to Estimate Parameters of Ordinary Differential Equation," *Computational Statistics*, vol. 35, no. 3, pp. 1481–1499, Sept. 2020, doi: 10.1007/s00180-020-00962-8.

[86] C. Margossian, L. Zhang, S. Weber, and A. Gelman, "Solving ODEs in a Bayesian Context: Challenges and Opportunities." Accessed: Mar. 08, 2026. [Online]. Available: https://www.page-meeting.org/Abstracts/solving-odes-in-a-bayesian-context-challenges-and-opportunities/

[87] T. Matsuda and Y. Miyatake, "Estimation of Ordinary Differential Equation Models with Discretization Error Quantification," *SIAM/ASA Journal on Uncertainty Quantification*, vol. 9, no. 1, pp. 302–331, Jan. 2021, doi: 10.1137/19M1278405.

[88] S. Minsker, S. Srivastava, L. Lin, and D. B. Dunson, "Robust and Scalable Bayes via a Median of Subset Posterior Measures," *Journal of Machine Learning Research*, vol. 18, no. 124, pp. 1–40, 2017.

[89] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, "Basics of Markov Chain Simulation," *Bayesian Data Analysis*. Chapman and Hall/CRC, 2013.

[90] A. Sokal, "Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms," *Functional Integration: Basics and Applications*. Springer US, Boston, MA, pp. 131–192, 1997. doi: 10.1007/978-1-4899-0319-8_6.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to Dr. Matthias König, whose invaluable feedback and exceptional guidance made this research and thesis possible. His unwavering support and extensive expertise were instrumental in the successful completion of this work.

I would also like to extend my heartfelt appreciation to Prof. Dr. Frank Konietschke for serving as the second examiner of this thesis and for his important contribution to the evaluation process.

Additionally, I wish to acknowledge the use of draw.io (https://drawio.com) and BioRender (https://biorender.com) for the creation of the diagrams and illustrations included in this thesis.

Lastly, I am deeply grateful to my friends and family for their constant support and encouragement throughout this journey, which was essential to bringing this thesis to completion.

Declaration of academic honesty

I hereby declare in lieu of an oath that I completed this master thesis independently and used only these sources and aids that are listed. All materials used, from published as well as unpublished sources, whether directly quoted or paraphrased, are duly reported.

Furthermore, I declare that the master thesis I have submitted, neither in its entirety nor in substantial parts, has been or is the subject of any other examination procedure.

I also assure that the digital version submitted by me corresponds in form and content to the printed version of the master thesis.

Erklärung zur akademischen Integrität

Hiermit erkläre ich an Eides statt, dass ich diese Masterarbeit selbstständig verfasst und ausschließlich die aufgeführten Quellen und Hilfsmittel verwendet habe. Sämtliche verwendeten Materialien, sowohl aus veröffentlichten als auch aus unveröffentlichten Quellen, ob wörtlich zitiert oder paraphrasiert, sind ordnungsgemäß angegeben.

Darüber hinaus erkläre ich, dass die von mir eingereichte Masterarbeit weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahrens war oder ist.

Ich versichere außerdem, dass die von mir eingereichte digitale Version in Form und Inhalt mit der gedruckten Version der Masterarbeit übereinstimmt.