

Accepted Manuscript

Use of prior knowledge and extrapolation in paediatric drug development: A case study with deferasirox

Elisa Borella, Sean Oosterholt, Paolo Magni, Oscar Della Pasqua



PII: S0928-0987(19)30186-1
DOI: <https://doi.org/10.1016/j.ejps.2019.05.009>
Reference: PHASCI 4931

To appear in: *European Journal of Pharmaceutical Sciences*

Received date: 29 September 2018
Revised date: 9 March 2019
Accepted date: 13 May 2019

Please cite this article as: E. Borella, S. Oosterholt, P. Magni, et al., Use of prior knowledge and extrapolation in paediatric drug development: A case study with deferasirox, *European Journal of Pharmaceutical Sciences*, <https://doi.org/10.1016/j.ejps.2019.05.009>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Use of prior knowledge and extrapolation in paediatric drug development: a case study with deferasirox

Elisa Borella, (1), Sean Oosterholt (2), Paolo Magni (1), Oscar Della Pasqua (2,3)

(1) Dipart. Ingegneria Industriale e dell'Informazione, Università degli Studi di Pavia, Pavia, Italy

(2) Clinical Pharmacology & Therapeutics Group, University College London, London, UK

(3) Clinical Pharmacology Modelling & Simulation, GlaxoSmithKline, Uxbridge, UK

Corresponding author:

Prof. Oscar Della Pasqua

Clinical Pharmacology & Therapeutics

BMA House, Tavistock Square

WC1H 9JP London

United Kingdom

Tel: +44207 874 1544

Email: o.dellapasqua@ucl.ac.uk

Word count: 3512

Number of figures and tables: 9

Number of references: 55

Keywords: Model-based meta-analysis; rare diseases; optimal design; clinical trial simulations; thalassaemia; deferasirox; paediatrics.

INTRODUCTION

Whereas guidelines are available and decision trees have been developed to describe the requirements for the generation and extrapolation of efficacy and safety of medicines in the paediatric population,^{1, 2} there are many cases in which efficacy and safety data cannot be easily generated. Equally, there are cases where requirements cannot be readily applied to children. Irrespective of similarities or differences in efficacy, safety or in the underlying exposure-response (PKPD) relationships between adults and children, it should be clear that the characterisation of pharmacokinetics (PK) is critical to ensure the appropriate dose rationale and dosing regimens in paediatric diseases.³ Such an assessment is full of challenges, especially in the context of rare diseases, for which the dose rationale for novel interventions is primarily determined by empirical evidence and the relevance of clinical pharmacology principles is often overlooked. Moreover, practical and ethical constraints further limit the opportunities to collect information on the overall benefit-risk ratio of a medicine, as individual data is sparse both in terms of sampling frequency and sampling intervals.^{4,5} These constraints can have clinical implications (e.g. dosing recommendations) if the quality and accuracy of the data cannot be warranted.

Historically, paediatric doses have been derived by dividing the adult dose linearly by body weight, i.e., doses are expressed in mg/kg. Across different therapeutic indications, paediatric doses for many products on the market have been obtained by empirical formulas based on age or other demographic characteristics.⁶ This approach ignores the nonlinear processes which take place due to maturation (e.g., ontogeny, metabolic activity) and developmental growth (e.g., renal function and size) and are known to occur particularly within the first years of life.⁶⁻⁹ Currently, however, different extrapolation approaches can be considered depending on the degree of similarity between adult and paediatric disease progression and response to intervention, including complete, partial or no extrapolation. In fact, evidence arising from pharmacokinetic and safety data may be deemed sufficient only when it is reasonable to assume that the aforementioned requirements can be met.

In spite of the increasing use of PK modelling in paediatric drug development, little attention has been paid to more extreme cases such as rare diseases, where extrapolation techniques along with the use of historical data become critical to ensure accuracy, precision and reduce the risk of bias and/or misspecification.^{10,11} In fact, previous publications have highlighted the limitations of extrapolations based on allometric principles when using adult or a different population as reference group for the dose rationale in another subgroup, without considering data collection in the target population or group of interest.^{12,13} This is partly explained by inaccurate estimation of covariate-parameter correlations and in some cases because of missing or misspecified covariates on key parameters such as clearance and volume of distribution.^{14,15} Consequently, it is rather surprising that despite the availability of different methods to integrate prior knowledge for the analysis of such sparse and unbalanced data,^{16,17} and for the optimisation of the study design,¹⁸⁻²⁰ their uptake in clinical research, and, particularly, in case of rare diseases, remains very low.

Using a case study in which deferasirox is given to paediatric patients affected by beta thalassaemia and other transfusion-dependent haemoglobinopathies, we expand on the principles introduced by Cella et al.² and evaluate the feasibility of applying Bayesian concepts and prior historical data in conjunction with extrapolation principles to improve parameter estimation. We focus on rare diseases, for which evidence synthesis is essential to overcome the difficulties in data generation. Conceptually, we evaluate whether model parameter estimates in adults along with prior distributions can be treated as informative and consequently exchangeable across distinct groups of the patient population or across treatments, after correction for putative covariate effects. In addition, we explore the advantages of the combined use of priors and optimisation techniques (ED-optimality) to support data generation itself, allowing a more robust experimental protocol design.^{21,22}

METHODS

Case study: prospective evaluation of the pharmacokinetics of deferasirox in patients aged >1 to 18 years

The DEferiprone Evaluation in Paediatrics (DEEP) consortium was set up under the auspices of the FP7 program, to evaluate the efficacy and safety of chelation therapy in paediatric patients affected by transfusion-dependent haemoglobinopathies. In this context, the DEEP-2 study was proposed to assess the non-inferiority of deferiprone relative to deferasirox in the paediatric population.²³⁻²⁵ PK data collection was included as an objective to ensure further characterisation of the concentration-effect relationship of the chelating agents in children and adolescents from >1 to 18 years of age. Due to feasibility reasons, only one blood sample per patient could be collected at the end of treatment (i.e., at steady state) in a subgroup of patients (n = 19). The sampling schedule was based on a sampling window of up to four hours after the last dose. Each subject was randomly assigned to one of ten different sampling times: 15 min pre-dose, and 15, 30, 45, 60, 75, 90, 105, 120, 240 min post-dose. A deviation of +/- 10 min was allowed around each sampling time.

Given the availability of a single blood sample per patient, we evaluated the implications of the proposed study design and explored opportunities to deal with data sparseness, uncertainty and (poor) precision in the estimation of parameters describing drug disposition. The possibility of estimating individual primary and secondary pharmacokinetic parameters with sufficient precision, using only one sample per child represents, therefore, a key aspect of the methodology proposed in the subsequent paragraphs.

In addition, a simulation-estimation workflow was adopted to explore the possible limitations of the DEEP-2 protocol design and assess the precision of parameters estimates from this specific trial (Figure 1, panel A). First, a population PK model was developed from historical data appropriately scaled, using an allometric approach based on body weight (Step 1, panel A). Then, deferasirox concentration vs. time data were simulated for a virtual population of paediatric patients using the aforementioned model, taking into account the DEEP-2 study protocol design (Steps 2-3, panel A).

Finally, primary and secondary PK parameters, including area under the concentration-time curve (AUC) and the maximum (or peak) concentration (C_{max}), were estimated to assess the expected precision and accuracy of the estimates from the proposed study protocol design (Step 4, panel A).

Prior knowledge

A model-based meta-analysis was performed using data extracted from five published PK studies with deferasirox.^{24,26-29} Studies were included in the analysis if the following criteria were met: (i) single or multiple oral doses of deferasirox were administered within the study period; (ii) study population consisted of healthy adult subjects and adult or paediatric patients with transfusion-dependent haemoglobinopathies; (iii) the reported data included mean or individual time-concentration data. An overview of the clinical protocols and patient population demographics (age, sex, weight, race, disease) is shown in Table 1. Deferasirox concentration vs. time profiles can be found in Figure S1 (**supplementary material**). All studies were conducted in adult patients except for the study reported in Chirnomas et al.²⁴, where also few paediatric patients aged between 3 and 18 years old were enrolled.

To ensure appropriate data aggregation and subsequent pooling, data from the various sources had to be normalised based on the assumption of linear pharmacokinetics across the range of doses included in the studies. First, all doses were converted into micromole (μmol) using a molecular weight of deferasirox of 373.362 g/mol.³² While accurate information about the dose (amount) is critical for the characterisation of PK parameters, assumptions had to be made when individual patient-level data were not reported. Since the dose is often expressed in mg/kg, when reported mean body weight was used to calculate the actual dose in mg. If body weight was missing, mean or individual body weight was imputed from reported demographic details such as age, sex, and race. Missing body weight for Japanese patients for the study published by Myazawa et al.²⁶ was assumed to be 57.7 kg based on the mean weight for Asians, as reported by Walpole et al.³⁰ A mean body weight of 70 kg for adult subjects was used for the studies published by Galanello et al.²⁹ and

Chirnomas et al.²⁴. By contrast, individual body weight for each patient aged ≤ 18 years old was calculated using the growth charts provided by the Disabled World website³¹, taking into account age, sex, and race. All assumptions along with an explanation for eventual data exclusion are listed in Table 1.

Pharmacokinetic modelling

A population PK model was developed for deferasirox using the available historical data (Step 1 in Figure 1, panel A). Further details on the model evaluation procedures along with the main diagnostic criteria can be found in the **supplementary material**.

Simulation of the pharmacokinetics of deferasirox in a virtual paediatric population

The population PK model was then used as reference for the evaluation of the impact of prior knowledge on the precision and accuracy of PK parameters obtained from the analysis of very sparse data in children. Since the PK model of deferasirox includes the effect of body weight as a covariate on clearance and volume of distribution, a virtual patient population with a representative covariate distribution was simulated (Step 2 in Figure 1, panel A). In this step, $> 10,000$ patients were generated (aged from > 1 to >18 years old with a 1:1 sex ratio in line with the study protocol inclusion criteria) to ensure appropriate covariate distribution across the population. Corresponding weights were simulated using an appropriate demographic model,³⁹ including correlations between postmenstrual age (PMA), sex, and body weight. To simulate body weight, PMA was extracted from a uniform distribution based on age ranging between >1 year and >18 years (plus a gestational time of 40 weeks). Two scenarios (Scenario 1 and 2) were subsequently evaluated to explore the most suitable method for the analysis of clinical trial data (Step 2 in Figure 1, panel A). Details of each scenario are summarised in Table 2. More specifically, in Scenario 2, five different situations were considered (a-e). In Scenario 2.a, 2.b and 2.c, we accounted for the possibility that allometric scaling might not be adequate to fully explain the PK differences between adults and children,

whereas in Scenario 2.d and 2.e we considered the impact of different allometric exponents on clearance. Consequently, the deferasirox concentration vs. time profiles of were simulated using population parameters values, which were significantly lower than the ones estimated in adults or with the same values but scaled with a different allometric exponent (only the allometric exponent on clearance was changed). It should be noted that as normalised renal function and liver metabolic activity in children are usually lower compared to adults, only scenarios in which population PK parameters are reduced relatively to the adult estimates were summarized. Scenarios in which the population PK parameters in the paediatric population are higher than those computed from adult studies were also tested but were not included here. The conditions mentioned above were identified among a range of possible perturbations, which were deemed sufficiently robust to assess the sensitivity of different methods for the analysis of very sparse pharmacokinetic data. Finally, six different simulated data sets were created, i.e., one for each of the aforementioned scenarios, including simulated deferasirox concentrations and corresponding demographic characteristics of the virtual patients (Step 2 in Figure 1, panel A).

Analysis of sparse pharmacokinetic data

For each of the aforementioned scenario, PK data from 19 subjects (i.e., the same number of patients available for pharmacokinetic assessment in the DEEP-2 study) were then randomly extracted from the simulated dataset, with only one sample per patient randomly selected across 10 sampling intervals (with a possible deviation of +/- 10 minutes), as defined in the DEEP-2 protocol (Step 3 in Figure 1, panel A).

The extracted samples were used as data input for the estimation of structural parameters for CL, V2, V3, k_a , and Q, and interindividual variability for CL, V2, V3, and k_a (Step 4 in Figure 1, panel A). This step was performed using the conditional estimation with interaction (FOCE-I). Differently from the model building phase with adult data where different model structures have been tested, the analysis with simulated paediatric data was performed under the assumption that

model structure, as defined by drug disposition in adults, can be used as starting point for the analysis of PK data in the paediatric population. Three conditions were investigated, namely estimation without priors, with highly informative priors or weakly informative priors. Both highly and weakly informative priors were implemented using the \$PRIOR option in NONMEM. Details of the implementation can be found in the **Supplementary material**.

Optimisation of the pharmacokinetic sub-study design

A secondary objective of the present investigation was to identify opportunities for best practice in the evaluation of PK data in rare diseases. Of interest was the need to demonstrate to what extent feasibility considerations can lead to biased estimates of the parameters of interest, and most importantly, how efficient a PK sub-study protocol can be if optimal design principles are applied at the design stage. Currently, these principles remain overlooked and are undervalued by the clinical research community. A graphical representation of the workflow steps adopted for this second objective is represented in Figure 1 (panel B). First, optimal sampling time windows have been identified by means of an ED-optimisation procedure (Step 1, panel B). Primary and secondary PK parameters were derived using the simulated pharmacokinetic data in Scenario 1 (Steps 2-3, panel B) to assess the possible benefits of different study designs, including more samples per subject within the optimal sampling time windows identified in Step 1. Additional details of the optimisation procedures can be found in the **supplementary material**.

Comparing original and optimised protocols

All scenarios were compared in terms of the probability of successful convergence of the NLME algorithm. Results were calculated based on the ratio between 100 and the number of runs necessary to obtain 100 successful runs. In addition, the precision of the estimates was assessed by the probability distributions of the ratios of each individual parameter relative to its 'true' value (i.e., the one used for simulating the concentration vs. time data), calculating for each of them the

proportion of the area under the probability distribution between 0.8 (1/1.25) and 1.25. This range was used under the assumption that a variation of 25% from the nominal value falls within the expected biological variability.⁴¹ The probability distribution of the ratio of estimated individual parameter values relative to their true values is expected to be centred around 1 if individual estimates are accurate (see Figure S3, panel A in the **supplemental material**). A deviation from 1 indicates the presence of bias. By contrast, the more the probability distribution is centred around 1, the higher the precision of the parameter estimates (see Figure S3, panel B, in the **supplemental material**). Four different comparisons have been performed, which are listed in Table 3.

RESULTS

Pharmacokinetic modelling

The pharmacokinetics of deferasirox in adults was best described by a two-compartmental model with first-order absorption and elimination. An overview of the model diagnostics and final estimates of the pharmacokinetic parameters used in the evaluation of priors are presented in Table 4 and Figure 2. Further details are summarized in the **supplementary material**.

Evaluation of the advantages of using priors for the analysis of sparse PK data (Comparison I-II)

The first two comparisons have been performed to demonstrate in a quantitative manner the advantages of using priors for the analysis of sparse data (see Comparison I and II in Table 3).

Comparison I: The convergence of the algorithm was chosen as main criteria to compare the performance in three different conditions: no priors, with weakly or highly informative priors (Table 4). Only results relative to the analysis of simulated data of Scenario 1 are shown. As reported in Table 5, the use of priors increases dramatically the probability to obtain a successful convergence of the NLME algorithm in case of sparse sampling from only 12% with no priors up to

56% and 75% for weakly and highly informative priors, respectively. These results indicate that the use of priors derived from adult appears to address some of the limitations due to the sparseness of data.

Comparison II: From the results in comparison I, it becomes clear that when historical data reflect the characteristics of the paediatric population correctly, highly informative priors should be considered for the evaluation of sparse data in children. However, this might not always be the case. To take into account cases where differences may exist, in Comparison II, PK parameter estimates obtained from the analysis of sparse data were compared to the corresponding ‘true’ values used in the simulation. All six scenarios were considered: from the most straightforward Scenario 1, in which the same population parameters and allometric scaling function obtained with historical data were used, to the more challenging Scenarios 2.a to 2.e, where parameter distributions describing drug disposition were assumed to deviate significantly from adults. In Scenarios 2.d and 2.e, the concentration-time profiles of the paediatric population were simulated only with a different allometric exponent compared to the one found from historical data. In Scenarios from 2.a to 2.c, more extreme situations were considered: the paediatric population was simulated assuming clearance and/or volume values that are half of the adult population estimates. For the first two scenarios, the use of weakly and highly informative priors gave comparable results in terms of ratios of posterior individual estimates. For scenarios from 2.a to 2.c, highly informative priors led to more imprecise estimates of AUC and C_{max} compared to weakly informative priors (Figure 3 and Figure 4, upper panels, and Table S1 in **supplementary material**). As it can be seen in Table S1, the percentages of AUC estimates deviating < 25% relative to their ‘true’ values are 26, 23, and 23% for scenarios 2.a, 2.b, and 2.c, respectively, when using weakly informative priors. By contrast, highly informative priors introduced a larger bias (i.e., 43, 37, and 36% for scenarios 2.a, 2.b, and 2.c, respectively). A graphical comparison between the ‘true’ PK profile for a specific subject and its corresponding PK profiles estimated using weakly, highly or no priors reveals that

weakly informative priors perform better in terms of closeness to the ‘true’ PK profile in most scenarios or conditions (Figure S8 in **supplementary material**, Panel A). On the other hand, no substantial improvement in the estimation is observed when using weakly informative priors for a subject whose PK parameters belong to the tails of their distributions (Figure S8 in **supplementary material**, Panel B). Consequently, the use of weakly informative priors should be the preferred option relatively to highly-informative priors. Weakly-informative priors can increase the probability of successful convergence and, simultaneously, are less likely to introduce bias in the paediatric PK parameters estimates.

Evaluation of the impact of optimised protocol designs on the precision of parameter estimates for extrapolation of data in children (Comparison III-IV)

The last two comparisons reported in Table 4, namely Comparison III and IV, focus on how historical information can be used together with optimisation techniques to guide the design of more informative trials in the paediatric population, especially in case of rare diseases.

ED-optimality showed that four blood samples, collected at optimised sampling times, were sufficient to ensure the precision of the parameters of interest. In contrast to common practice in pharmacokinetic studies, more frequent sampling did not yield further increase in parameter precision. Four optimal sampling windows were identified for deferasirox: -30 to 0 min pre-dose, 15 to 30 min, 90 to 150 min, and 225 to 240 min post-dose (Figure S9).

Comparison III: In Comparison III, the original study design (1 sample/subject within the sampling times defined in the study protocol) was compared to an ‘optimised’ design (1 sample/subject within the selected optimal sampling windows). Both the probability of successful minimisation and ratio between the estimated and ‘true’ parameter values were used to quantify the differences between the two designs.

The result of this comparison (Figure 3 and Figure 4, left lower panel) shows that optimising the sampling protocol when only one sample per subject is collected does not significantly improve the precision or accuracy of the estimates or the probability of having a successful minimisation (Table 5). The use of priors gives the necessary support to parameter estimation irrespective of optimisation procedures.

Comparison IV: In Comparison IV, several ‘optimised’ study designs have been compared to the original one. Notably, four optimised designs including 1, 2, 3, and 4 samples/subject, respectively, all selected within the optimal sampling windows. Both the probability of successful minimisation and ratio between the estimated and ‘true’ parameter values were used as metrics for his comparison. The emerging results show to what extent the precision of the parameter estimates increases when more samples are taken, suggesting what should be the minimum number of samples to balance feasibility and validity of the study (Figure 3 and Figure 4, lower right panel). With two samples the probability of successful minimisation dramatically increases to about 90% (Table 5), but the probability of overestimating or underestimating C_{\max} and AUC of more than 25% remains less than desirable (still around 60%, Table S1 in **supplementary material**).

This last comparison shows that increasing the number of samples to three or four results in significant reduction of the probability of having exposure-related parameter estimates outside the boundary of acceptability, namely < 20% with three samples and < 10% with four samples. Given other practical constraints, our analysis suggests that when optimal sampling windows are identified, the use of three samples per subject is sufficient to ensure accurate and precise estimates for most PK parameter and their variabilities.

DISCUSSION

Historically, limited attention has been given to pharmacokinetics and PKPD relationships as the basis for the dose rationale in children. The change in the legislation, which now requires drug developers to provide evidence of efficacy and safety of novel treatments for paediatric diseases, has not eliminated some important challenges associated with data generation, in particular in rare diseases.

To date, limitations in data generation are being addressed by increased acceptance of extrapolation methodologies,^{42,43} in which pharmacokinetics, pharmacodynamics and efficacy are inferred from a reference patient population, and eventually from animals, another compound or disease.⁴⁴⁻⁴⁶ Whereas inferential methods are extremely important, and formal extrapolation approaches provides transparency for empirical dose selection, which is often observed during off-label use of a medicinal product, a key issue remains, in that any data oncoming from the target population is likely to be sparse and need to be integrated with existing knowledge, whether or not previously framed under an extrapolation exercise. It is therefore essential to maximise the usefulness of the data obtained with the minimum number of subjects enrolled. It is equally important to understand the potential bias in parameter estimation from sparse data, especially for compounds which show complex disposition, or which are affected multiple covariate factors. In fact, it should be clear that evidence of efficacy and safety does not imply optimal dosing. This represents an often-overlooked aspect of paediatric development in that many treatments are associated with chronic conditions. Consequently, dose optimisation can have major implications for the long-term efficacy and safety profile of the product.

Here we used a case study for which the assumption of comparable disease and PKPD relationships across populations appears to be valid and biologically plausible. Chelation treatment of transfusion-dependent iron overload is determined primarily by the total iron intake from blood transfusions. Hence, it has been assumed that the dose rationale for children correlates or corresponds to the efficacious levels of deferasirox in the source population (adults). However, even

when relevant data are available and can be used as prior information, the contribution of covariate factors, such as age, body weight, or other clinical parameters that are determinants of differences in drug disposition in children may not have been identified in the adult population. These covariates need to be considered when defining dosing recommendations.

Another important feature of historical data, beyond extrapolations, are its role and relevance for trial design optimisation. Despite the availability of different techniques, protocol designs in rare diseases continue to be developed and implemented without careful consideration of these instruments.⁴⁷⁻⁴⁹ Recently, several cases of failed completed trials have been reported,⁵⁰⁻⁵¹ where a poor dose selection contributed to a trial failure. In Benjamin et al.⁵⁰ fixed doses across a wide range of body weights have been used; in Momper et al.⁵¹ no dose response was seen in the study, and it was stated that, if higher doses had been evaluated, efficacy might have been demonstrated. This reinforces the importance of PK assessments in age groups where drug disposition cannot be reliably predicted. These failures also highlight the need to collect pharmacokinetic and biomarker data in efficacy studies to assess PKPD relationships when efficacy may not be determined empirically in a controlled clinical trial.

From a methodological perspective, we have shown how the use of priors may have an ‘anchoring’ effect, supporting the minimisation procedures, improving convergence and precision of the parameter estimates of interest. However, adding priors could be a concern for clinicians and drug regulators. For this reason, a sensitivity analysis in which different prior distributions are tested should be done to ensure that the conclusions are not heavily weighted on prior beliefs, as illustrated by the evaluation of both weakly and highly informative priors. In addition, the simulation-estimation modelling exercise supported the views that, in the presence of limited data, the use of priors always has an advantage compared to those situations where no prior information is used. In this regard, there is increasing evidences in the literature on the impact of priors on parameter estimation, especially when (i) model complexity causes identifiability problems, (ii)

data are very noisy, and (iii) data are too scarce or the experiment design is far from optimal.^{52,53} In support of these concepts, we refer the reader to a range of examples in other therapeutic areas.⁵⁴⁻⁵⁶ Despite its application in Bayesian statistics, the integration of a priori information has not been evaluated systematically in the context of paediatric drug development where situations (ii) and (iii) are still commonplace. This investigation does not intend to provide the readers with a final solution for known issues related to data sparseness in paediatric trials. We hope that our findings are sufficient to highlight the advantages and disadvantages of applying Bayesian concepts in conjunction with optimal sampling. In fact, in comparisons I and II, we demonstrated in a quantitative manner to what extent weakly informative priors increase the robustness of the estimates, allowing integration of prior knowledge from historical data without dominating the parameter space during the estimation procedures. On the other hand, the use of highly informative priors led to slightly better performances in terms of convergence but resulted in a bias when drug disposition in children differs significantly from adults (Scenarios 2.a-2.c). Such a situation may be illustrated by cases in which maturation of enzymes involved in drug metabolism (i.e., ontogeny) or when other physiological processes (e.g. immune competence in neonates) are unique to the paediatric population as compared to adults. Nevertheless, even if weakly informative priors are used, the availability of only one single sample per patient prevents the distinction between intra- and inter-individual variability. At best, with such kind of data, sufficiently accurate population estimates of CL/F and its IIV can be obtained, and therefore, individual AUC estimates can be derived. Conscious that the number of samples and time of sampling are also determined by practical constraints, our results also highlight the limitations of common protocol designs, and in particular the possible drawbacks of collecting only one sample per patient, leading to a probability of more than 60% of over/underestimating the exposure by $> 25\%$ (Table S1).

The advantages of using optimisation techniques based on priors have been shown in comparisons III and IV. Whereas it is known that PK parameter estimates and corresponding interindividual

variability are strongly dependent on the number of samples collected per subject, our analysis reveals that with at least three samples per individual, it is possible to obtain accurate and precise estimates for most PK parameter and their variabilities. In this regard, the DEEP-2 study has been designed and conducted with only one PK sample per patient. Even though a different sampling scheme could not be implemented due to the nature of the intervention (out-patients) and other practical limitations, the data collected in the study will be analysed based on the use of the priors defined in in this work.

Limitations

Although several methods are available in the literature for the integration of prior knowledge and trial design optimisation, in this work only \$PRIOR in NONMEM have been explored. This choice was based on the fact that estimation methods based on the maximum likelihood represent the most widely used approach in nonlinear mixed effects modelling. It is also the most straightforward procedure to integrate prior knowledge into a model. Fully Bayesian hierarchical modelling is another way to integrate prior knowledge into the parameter estimation process.⁵³ Several software tools, such as \$BAYES in NONMEM, WinBUGS,⁵⁷ Stan,⁵⁸ can be used to encode Bayesian models and to carry out parameter estimation via Markov Chain Monte Carlo (MCMC) algorithms. However, it was not our objective to compare the performance of different estimation methods. Regardless the limitations, we have shown that the use of adult data should not be restricted to extrapolation exercises. Priors from adults can be used in conjunction with ED-optimal design to ensure sampling schemes are truly informative. This concept can be readily generalised to other drugs for which data generation in children is a potential issue. One should also keep in mind that, our findings may not be generalised to situations where the assumption of comparable exposure-response relationship, efficacy and safety between reference and target population is not biologically or clinically plausible.

CONCLUSIONS

The use of priors increases the probability of successful convergence and higher precision of the estimates as compared to the no-priors case when dealing with very sparse data. Here, priors were derived from PK parameter estimates in adults, taking into account their uncertainties. We have shown that the choice of highly informative priors should be based on evidence supporting the comparability of the drug disposition processes across populations (e.g., children and adults) to prevent biased estimates. Our analysis also highlighted the limitations of collecting only a single sample per subject, which may result in over/underestimation of the exposure in a fraction of patients. This bias can be overcome by more frequent sampling with three samples per subject.

Disclosure:

E Borella is currently an employee of Menarini Ricerche, Florence, Italy. O Della Pasqua is also Senior Director Clinical Pharmacology at GlaxoSmithKline, Uxbridge, United Kingdom. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Funding:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

1. ICH Expert Working Group. Addendum to ICH E11: Clinical investigation of medicinal products in the paediatric population E11 (R1). http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E11/E11-R1EWG_Step4_Addendum_2017_0818.pdf
2. Cella M, Gorter de Vries F, Burger D, Danhof M, Della Pasqua O. A model-based approach to dose selection in early paediatric development. *Clin Pharmacol Ther.* 2010;87(3):294-302.
3. Cella M, Knibbe C, Danhof M, Della Pasqua O. What is the right dose for children? *Br J Clin Pharmacol.* 2010;70(4):597-603.
4. European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP). Guideline on clinical trials in small populations. Doc. Ref. CHMP/EWP/83561/2005. 2006. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003615.pdf
5. US Food and Drug Administration. Report: complex issues in developing drugs and biological products for rare diseases and accelerating the development of therapies for paediatric rare diseases. 2014. <https://www.fda.gov/downloads/RegulatoryInformation/LawsEnforcedbyFDA/SignificantAmendments/totheFDCA/FDASIA/UCM404104.pdf>
6. Bellanti F, Della Pasqua O. Modelling and simulation as research tools in paediatric drug development. *Eur J Clin Pharmacol.* 2011;67(1):75–86.
7. Anderson, Brian J., et al. "Vancomycin pharmacokinetics in preterm neonates and the prediction of adult clearance." *British journal of clinical pharmacology* 63.1 (2007): 75-84.
8. Liu, Tao, et al. "Mechanistic population pharmacokinetics of morphine in neonates with abstinence syndrome after oral administration of diluted tincture of opium." *The Journal of Clinical Pharmacology* 56.8 (2016): 1009-1018.
9. Gonzalez, Daniel, et al. "Clindamycin pharmacokinetics and safety in preterm and term infants." *Antimicrobial agents and chemotherapy* (2016): AAC-03086.
10. De Cock RF, Piana C, Krekels EH, Danhof M, Allegaert K, Knibbe CA. The role of population PK–PD modelling in paediatric clinical research. *Eur J Clin Pharmacol.* 2011;67(1):5–16.
11. Vermeulen E, van den Anker JN, Della Pasqua O, Hopppu K, van der Lee JH, Global Research in Paediatrics (GRiP). How to optimise drug study design: pharmacokinetics and pharmacodynamics studies introduced to paediatricians. *J Pharm Pharmacol.* 2017;69(4):439-447.
12. Cella M, Knibbe C, de Wildt SN, Van Gerven J, Danhof M, Della Pasqua O. Scaling of pharmacokinetics across paediatric populations: the lack of interpolative power of allometric models. *Br J Clin Pharmacol.* 2012;74(3):525-535
13. Cella M, Zhao W, Jacqz-Aigrain E, Burger D, Danhof M, Della Pasqua O. Paediatric drug development: are population models predictive of pharmacokinetics across paediatric populations? *Br J Clin Pharmacol.* 2011;72(3):454-464.
14. Behrman RE, Field MJ. Ethical conduct of clinical research involving children. National Academies Press; 2004.

15. Piana C, Danhof M, Della Pasqua O. Influence of covariate distribution on the predictive performance of pharmacokinetic models in paediatric research. *Br J Clin Pharmacol*. 2014;78(1):145-57.
16. Bellanti F, Del Vecchio GC, Putti MC, et al. Population pharmacokinetics and dosing recommendations for the use of deferiprone in children younger than 6 years. *Br J Clin Pharmacol*. 2017;83(3):593–602.
17. Gamalo- Siebers M, Savic J, Basu C, et al. Statistical modeling for Bayesian extrapolation of adult clinical trial information in paediatric drug evaluation. *Pharmaceutical statistics*. 2017;16(4):232-249.
18. D'Argenio DZ. Optimal sampling times for pharmacokinetic experiments. *J Pharmacokinet Pharmacodyn*. 1981;9(6):739–756.
19. D'Argenio DZ. Incorporating prior parameter uncertainty in the design of sampling schedules for pharmacokinetic parameter estimation experiments. *Math Biosci*. 1990;99(1):105–118.
20. Petit C, Jullien V, Samson A, et al. Designing a paediatric study for an antimalarial drug by using information from adults. *Antimicrob Agents Chemother*. 2016;60(3):1481–1491.
21. Nyberg J, Ueckert S, Strömberg EA, Hennig S, Karlsson MO, Hooker AC. PopED: an extended, parallelized, nonlinear mixed effects models optimal design tool. *Comput Methods Programs Biomed*. 2012;108(2):789–805.
22. Foracchia M, Hooker AC, Vicini P, Ruggeri A. POPEd, a software for optimal experiment design in population kinetics. *Comput Methods Programs Biomed*. 2004;74(1):29–46.
- 23: <https://clinicaltrials.gov/ct2/show/NCT01825512>
24. Chirnomas D, Smith AL, Braunstein J, et al. Deferasirox pharmacokinetics in patients with adequate versus inadequate response. *Blood*. 2009;114(19):4009–4013.
25. “DEferiprone Evaluation in Paediatrics” Project reference: 261483
26. Miyazawa K, Ohyashiki K, Urabe A, et al. A safety, pharmacokinetic and pharmacodynamic investigation of deferasirox (Exjade®, ICL670) in patients with transfusion-dependent anemias and iron-overload: a Phase I study in Japan. *Int J Hematol*. 2008;88(1):73.
27. Sechaud R, Robeva A, Belleli R, Balez S. Absolute oral bioavailability and disposition of deferasirox in healthy human subjects. *J Clin Pharmacol*. 2008;48(8):919–925.
28. Waldmeier F, Bruin GJ, Glaenzel U, et al. Pharmacokinetics, metabolism, and disposition of deferasirox in β -thalassemic patients with transfusion-dependent iron overload who are at pharmacokinetic steady state. *Drug Metab Dispos*. 2010;38(5):808–816.
29. Galanello R, Piga A, Cappellini MD, et al. Effect of food, type of food, and time of food intake on deferasirox bioavailability: recommendations for an optimal deferasirox administration regimen. *J Clin Pharmacol*. 2008;48(4):428–435.
30. Walpole SC, Prieto-Merino D, Edwards P, Cleland J, Stevens G, Roberts I. The weight of nations: an estimation of adult human biomass. *BMC Public Health*. 2012;12(1):439.
31. <https://www.disabled-world.com/artman/publish/height-weight-teens.shtml>

32. Wishart DS, Knox C, Guo AC, et al. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* 2006;34(suppl_1):D668–D672.
33. Figg WD, McLeod HL. *Handbook of anticancer pharmacokinetics and pharmacodynamics.* Springer Science & Business Media; 2004.
34. Miyagi SJ, Milne AM, Coughtrie MW, Collier AC. Neonatal development of hepatic UGT1A9: implications of paediatric pharmacokinetics. *Drug Metab Dispos.* 2012;40(7):1321–1327.
35. Strassburg CP, Strassburg A, Kneip S, et al. Developmental aspects of human hepatic drug glucuronidation in young children and adults. *Gut.* 2002;50(2):259–265.
36. Tayman C, Rayyan M, Allegaert K. Neonatal pharmacology: extensive interindividual variability despite limited size. *J Paediatr Pharmacol Ther.* 2011;16(3):170–184.
37. Walton K, Dorne JL, Renwick AG. Uncertainty factors for chemical risk assessment: interspecies differences in glucuronidation. *Food Chem Toxicol.* 2001;39(12):1175–1190.
38. Bruin GJ, Faller T, Wiegand H, et al. Pharmacokinetics, distribution, metabolism, and excretion of deferasirox and its iron complex in rats. *Drug Metab Dispos.* 2008;36(12):2523–2538.
39. Sumpter AL, Holford NH. Predicting weight using postmenstrual age—neonates to adults. *Paediatr Anesth.* 2011;21(3):309–315.
40. Gisleskog PO, Karlsson MO, Beal SL. Use of prior information to stabilize a population data analysis. *J Pharmacokinet Pharmacodyn.* 2002;29(5):473–505.
41. US Food and Drug Administration. Center for Drug Evaluation and Research (CDER). Draft Guidance: Drug interaction studies: study design, data analysis, implications for dosing, and labeling recommendations. 2012. <https://www.fda.gov/downloads/drugs/guidances/ucm292362.pdf>
42. European Medicines Agency. Reflection paper on extrapolation of efficacy and safety in paediatric medicine development. Doc. Ref. EMA/199678/2016. 2016. http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2016/04/WC500204187.pdf
43. Dunne J, Rodriguez WJ, Murphy MD, et al. Extrapolation of adult data and other data in paediatric drug-development programs. *Paediatrics.* 2011;ped5–2010.
44. Zhao W, Biran V, Jacqz-Aigrain E. Amikacin maturation model as a marker of renal maturation to predict glomerular filtration rate and vancozacin clearance in neonates. *Clin Pharmacokinet.* 2013;52(12):1127–1134.
45. Kechagia IA, Kalantzi L, Dokoumetzidis A. Extrapolation of enalapril efficacy from adults to children using pharmacokinetic/pharmacodynamic modelling. *J Pharm Pharmacol.* 2015;67(11):1537–1545.
46. Bradley JS, Peacock G, Krug SE, et al. Paediatric anthrax clinical management. *Paediatrics.* 2014.
47. European Medicines Agency (EMA). Committee for Medicinal Products for Human Use (CHMP). Guideline on the role of pharmacokinetics in the development of medicinal products in the paediatric population. Doc. Ref. EMA/CHMP/EWP/147013/2004 Corrigendum. 2006. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003066.pdf

48. Manolis E, Pons G. Proposals for model-based paediatric medicinal development within the current European Union regulatory framework. *Br J Clin Pharmacol*. 2009;68(4):493–501.
49. Bellanti F, Di Iorio VL, Danhof M, Della Pasqua O. Sampling optimisation in pharmacokinetic bridging studies: Example of the use of deferiprone in children with β -thalassaemia. *J Clin Pharmacol*. 2016;56(9):1094–1103.
50. Benjamin DK, Smith PB, Jadhav P, et al. Paediatric antihypertensive trial failures. *Hypertension*. 2008;51(4):834–840.
51. Momper JD, Karesh A, Green DJ, et al. Drug development for paediatric neurogenic bladder dysfunction: dosing, endpoints, and study design. *J Clin Pharmacol*. 2014;54(11):1239–1246.
52. Carson E, Cobelli C. Parameter Estimation. In: *Modelling methodology for physiology and medicine*. Newnes; 2013.
53. Carson E, Cobelli C. Population Modelling. In: *Modelling methodology for physiology and medicine*. Newnes; 2013.
54. Magni P, Bellazzi R, Nauti A, Patrini C, Rindi G. Compartmental model identification based on an empirical Bayesian approach: the case of thiamine kinetics in rats. *Med Biol Eng Comput*. 2001;39(6):700–706.
55. Pillonetto G, Sparacino G, Cobelli C. Numerical non-identifiability regions of the minimal model of glucose kinetics: superiority of Bayesian identification. *Math Biosci*. 2003; 184(1):53–67.
56. Magni P, Sparacino G, Bellazzi R, Cobelli C. Reduced sampling schedule for the glucose minimal model: importance of Bayesian estimation. *Am J Physiol Endocrinol Metab*. 2006;290(1):E177–184.
57. <http://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-winbugs/>
58. <http://mc-stan.org/>

TABLES

Table 1: Overview of the clinical studies available in the published literature in which pharmacokinetic data has been reported for deferasirox.

Table 2: Simulation scenarios implemented for the evaluation of the impact of historical data (priors) on model parameter estimation for paediatric rare diseases.

Table 3: Overview of the comparisons included in this analysis.

Table 4: Population pharmacokinetic parameter estimates for deferasirox in adult subjects.

Table 5: Probability of successful run for Comparison I, III, and IV.

FIGURES

Figure 1: Simulation-estimation workflow.

Figure 2: In panel (a): goodness-of-fit plots (GOF) for the estimation procedures in step (2) with individual data where population parameters were fixed to values obtained in step (1) and inter-individual variabilities were estimated. Plot of observed concentrations vs. individual (top-right panel) and population predicted (top-left panel). Identity and regression lines are shown in red and blue, respectively. Plot of conditional weighted residuals (CWRES) vs. time (bottom-right panel) and individual predicted (bottom-left panel). Zero and regression lines are shown in red and blue.

In panel (b): histogram of probability density (red line) of the normalised predictive distribution error (NPDE) with the normal distribution (black line) superimposed (top-left panel), scatter plot of NPDE vs. time after dose (top-right panel) and scatter plot of NPDE vs. individual predicted concentration (bottom panel) for the estimation in step (2). Horizontal dashed lines in the scatter plots represent zero line, 90% and 95% confidence intervals.

In panel (c): Population predicted (red), individual predicted values (black), and observed data (solid circles) vs. time after dose (2).

In panel (d): visual predicted check (VPC) plots for the estimation procedures in step (2). The observed data (black circles) were overlaid with predicted median (dashed black line), and 95% prediction interval (PI) (shaded grey area).

Figure 3: Probability distributions of the ratio between the estimated area under the concentration-time curve (AUC) and the “true” AUC (i.e., from the simulated data). Upper panels refer to Comparison II, lower left panel to Comparison III and lower right panel to Comparison IV.

Figure 4: Probability distributions of ratios of the estimated maximum concentration (C_{max}) to the true C_{max} (i.e., from the simulated data). Upper panels refer to Comparison II, lower left panel to Comparison III and lower right panel to Comparison IV.

Reference	Daily dose	Sampling times	N° of patients	Age (years)	Weight (kg)	Sex (male:female)	Race	Disease	Exclusions	Additional Comments
[26]	5 mg/kg 10 mg/kg 20 mg/kg 30 mg/kg	Samples on Day 1 and Day 14	6 7 6 7	71.5 68.0 66.0 75.0	Not reported. The mean weight for Asian countries was assumed to be 57.7 kg, as reported in Walpole et al., 2012 [30]	1:5 3:4 1:5 3:4	Japanese	3 MDS, 3 AA, 5 MDS, 1 AA, 1 other 4 MDS, 1 AA, 1 other 4 MDS, 1 AA, 2 other	Data relative to 30 mg/kg (day 14) were excluded	Data excluded because of high noise in the data
[27]	375 mg	Samples on Day 1 at pre-dose (0), 15, 30, 45 min, and at 1, 1.5, 2, 2.5, 3, 3.5, 4, 6, 8, 12, 24, 36, and 48 h post-dose	17	30.5	79.05	17:0	22.2% Caucasian, 16.7% Black, 61.1% other race.	Healthy subjects	-	-
[28]	1000 mg (~20 mg/kg)	Samples on Day 7 at pre-dose (0), 0.5, 1, 1.5, 2, 3, 4, 6, 8, 12, 24, 48, 72, 96, 120, 144, 168 h post-dose	5	20-38	50-81	3:2	Not reported	5 THA	Study data were excluded	Data excluded because discrepancies between the dose reported in mg/kg, the mean body weight of the study population and the actual dose administered in mg
[29]	20 mg/kg	Samples on Day 1 at pre-dose (0), 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 6, 8, 12, 24, 36, 48 h post-dose	28	18-45	A mean weight of 70 kg was assumed for this population	28:0	Not reported	Healthy subjects	Data relative to arm C were excluded	Difficulties in extracting data due to overlap of the data with other treatment arms
[24]	34.7 mg/kg	Samples on Day 1 at pre-	15	9-38 3-36	For patients < 18 years,	3:2 7:3	1 Asian, 1 Black, 3 White	1 SCD, 4 THA 1 SCD,	-	-

Reference	Daily dose	Sampling times	N° of patients	Age (years)	Weight (kg)	Sex (male:female)	Race	Disease	Exclusions	Additional Comments
		dose (0), 1, 2, 4, 6, 12, 24 h post-dose			the weight was calculated from individual age, sex, and race using the growth charts in [31]. A mean weight of 70 kg was assumed for patients > 18 years old.		4 Asian, 1 Black, 5 White	9 THA		

Table 1: Overview of the clinical studies available in the published literature in which pharmacokinetic data has been reported for deferasirox.

AA: aplastic anaemia; MDS: myelodysplastic syndromes; SCD: sickle-cell disease; THA: thalassaemia

Table 2: Simulation scenarios implemented for the evaluation of the impact of historical data (priors) on model parameter estimation for paediatric rare diseases.

Scenario	Assumption	Parameters
Scenario 1	The pharmacokinetics of deferasirox in the paediatric population was assumed to be accurately described by allometric principles, as described by the proposed reference PK model;	Deferasirox concentration vs. time profiles of the virtual paediatric patients were simulated using the population parameter distributions summarised in Table 3
Scenario 2	The pharmacokinetics of deferasirox in the paediatric population was assumed to vary significantly from the proposed reference PK model. For this scenario, five different conditions were considered.	a. CL was significantly lower than that computed from adult studies (i.e., 50% of the predicted value);
		b. Both CL and V2 were significantly lower than those computed from adult studies (i.e., 50% of the predicted value);
		c. All disposition parameters (CL, V2, Q, and V3) were significantly lower than those computed from adult studies (i.e., 50% of the predicted value);
		d. The exponent of CL and Q was set to 0.85 (rather than 0.75) in the allometric equation;
		e. The exponent of CL and Q was set to 2/3 (rather than 0.75) in the allometric equation.

Table 3: Overview of the comparisons included in this analysis.

Type of sampling	N° of samples	Scenario	Notes	Priors
Comparison I*				
Empirical sampling	1	1	Parameters allometrically scaled	Weakly informative
				Highly informative
				No priors
Comparison II*				
Empirical sampling	1	1	Parameters allometrically scaled	Weakly informative Highly informative
		2.a	$CL=CL_{adult}/2$	
		2.b	$CL=CL_{adult}/2,$ $V2=V2_{adult}/2$	
		2.c	$CL=CL_{adult}/2,$ $V2=V2_{adult}/2,$ $Q=Q_{adult}/2,$ $V3=V3_{adult}/2$	
		2.d	Allometric exponent of CL and $Q=0.85$	
		2.e	Allometric exponent of CL and $Q=2/3$	
Comparison III*				
Empirical sampling	1	1	Parameters allometrically scaled	Weakly informative
Optimised sampling				
Comparison IV*				
Optimised sampling	1	1	Parameters allometrically scaled	Weakly informative
	2			
	3			
	4			

* In *Comparison I*, the results obtained with and without the use of priors from the sparse data generated using the original protocol were compared. With this comparison, we aimed to demonstrate the added value of the use of priors when no other approaches for protocol optimisation are feasible. In *Comparison II*, the results obtained with the use of highly and weakly informative priors from the sparse data generated using the original protocol were compared assuming differences in the disposition of deferasirox in the paediatric population, as compared to adults. The aim of this comparison was to assess situations in which highly informative priors can lead to biased conclusions and, in such cases, quantify it. In *Comparison III*, an optimised pharmacokinetic sampling protocol with the same number of samples/patient was compared to the original clinical trial design including weakly informative priors for the pharmacokinetic analysis. The goal here was to demonstrate to what extent optimised PK sampling times can contribute to improved parameter precision, even though the total number of samples is the same and cannot be modified. In *Comparison IV*, different optimised protocols with an increasing number of samples/patient have been compared. The aim of this comparison was to demonstrate the informative value of additional, but yet sparse samples per patient, i.e. from 1 to 2, or from 1 to 3 or from 1 to 4 samples.

Table 4: Population pharmacokinetic parameter estimates for deferasirox in adult subjects.

Parameter	Description	Unit	Population estimate (%RSE)	Bootstrap median (90% CI)	IIV ^a (%RSE)	Bootstrap median (90% CI)
k_a	Rate of absorption	h ⁻¹	0.956 (27.1%)	1.02 (0.51-1.40)	1.63 (33%)	1.62 (0.53-2.72)
CL	Clearance	L/h	1.81 (8.1%)	1.77 (1.59-2.04)	0.45 (65%)	0.45 (0.15-0.76)
Q	Intercompartmental clearance	L/h	1.85 (31.3%)	1.73 (0.86-2.84)	-	-
V2	Central volume of distribution	L	20.80 (14.6%)	21.38 (15.83-25.70)	0.412 (38%)	0.43 (0.17-0.66)
V3	Peripheral volume of distribution	L	15.10 (19%)	14.79 (10.03-20.21)	2.32 (>100%)	2.44 (89.97-94.2)
F	Bioavailability	-	0.70 (FIXED)	0.70 (FIXED)	-	-
σ_{PROP}^b	Residual error (proportional)	-	0.018 (18.9%)	-	-	-

RSE: relative standard error; IIV: inter-individual variability; CI: confidence interval.

^aReported as OMEGA that is the NONMEM output for IIV.

^bReported as SIGMA that is the NONMEM output for the variance of the residual error.

Table 5: Probability of successful run for Comparisons I, III, and IV.

Type of sampling	N° of samples	Scenario	Notes	Priors	Median of probability of successful run (90% CI)
<i>Comparison I</i>					
Empirical sampling	1	1	Parameters allometrically scaled	Weakly informative	56.50 (50.28-62.71)
				Highly informative	75.19 (69.17-81.20)
				No priors	12.22 (10.51-14.18)
<i>Comparison III</i>					
Empirical sampling	1	1	Parameters allometrically scaled	Weakly informative	56.50 (50.28-62.71)
Optimised sampling					51.28 (45.64-57.43)
<i>Comparison IV</i>					
Optimised sampling	1	1	Parameters allometrically scaled	Weakly informative	51.28 (45.64-57.43)
	2				89.96 (81.74-92.17)
	3				92.59 (88.89-93.30)
	4				94.34 (90.57-98.11)

CI: confidence interval.

^aBootstrap median and 90% CI were calculated sampling N-times with replacement from the pool of N runs necessary to have reached 100 successful runs.

ACCEPTED MANUSCRIPT

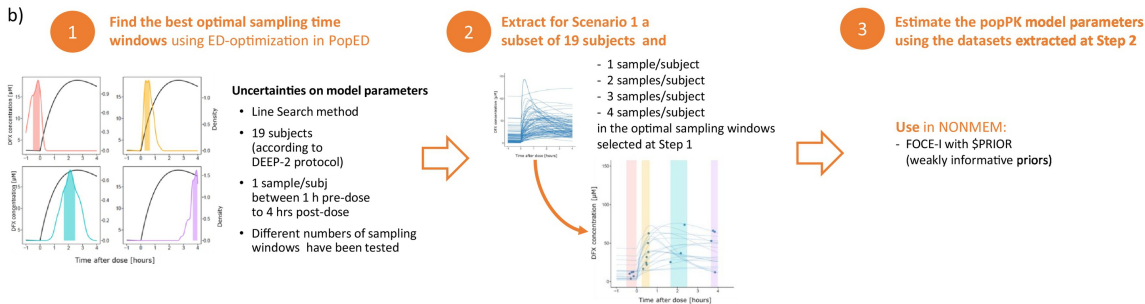
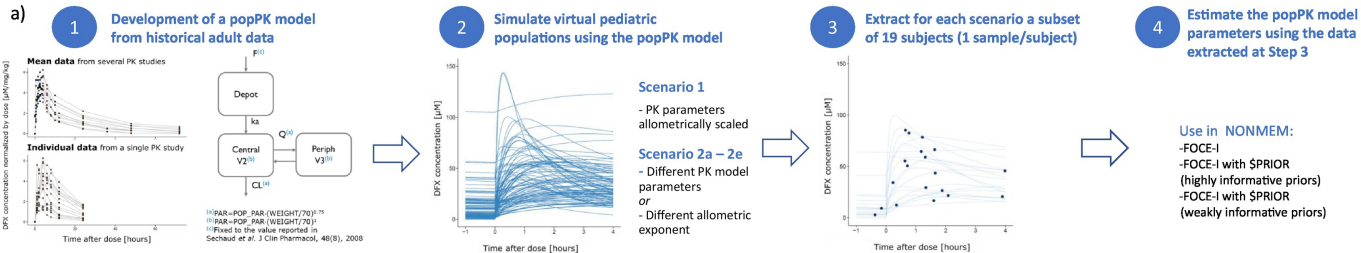


Figure 1

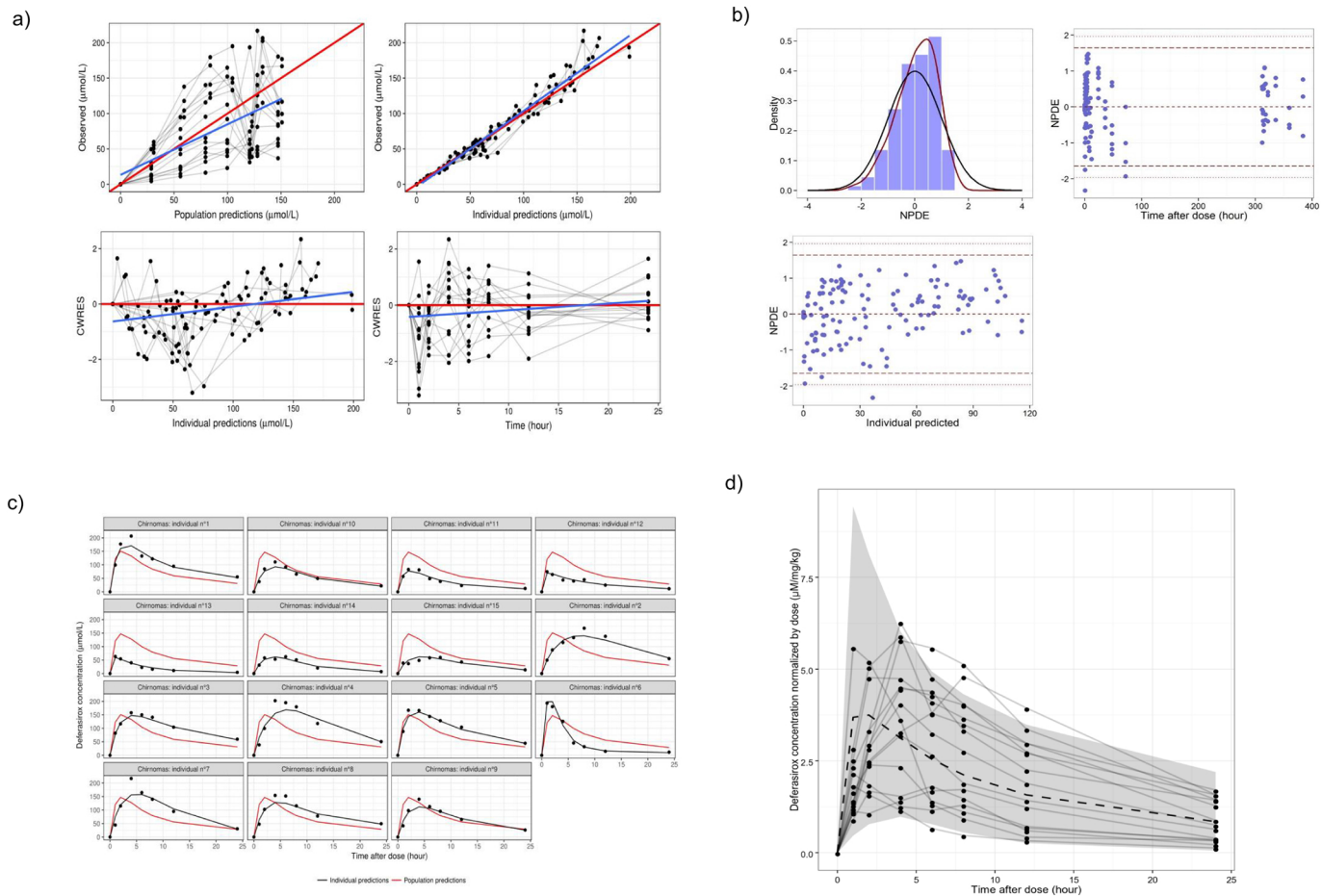


Figure 2

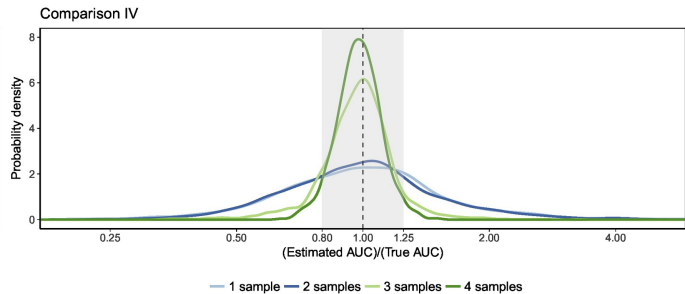
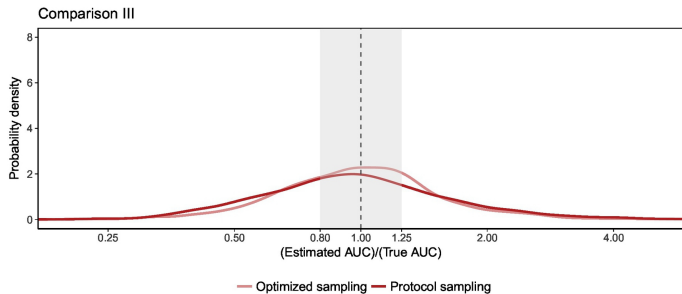
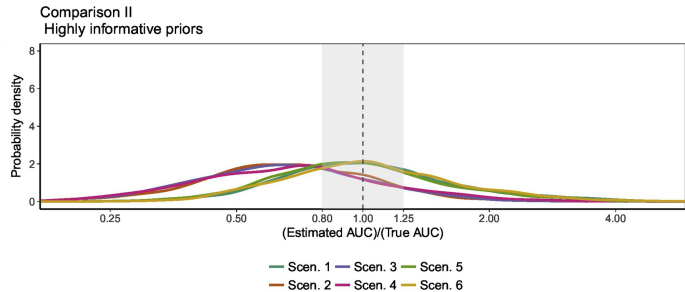
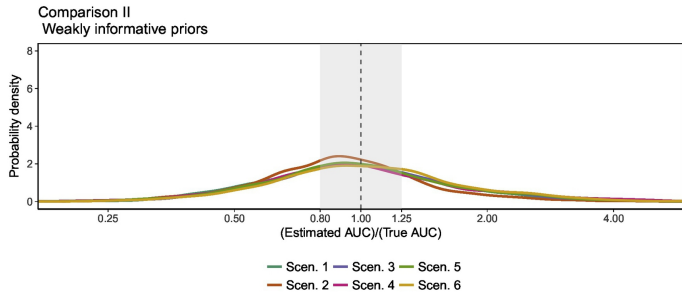


Figure 3

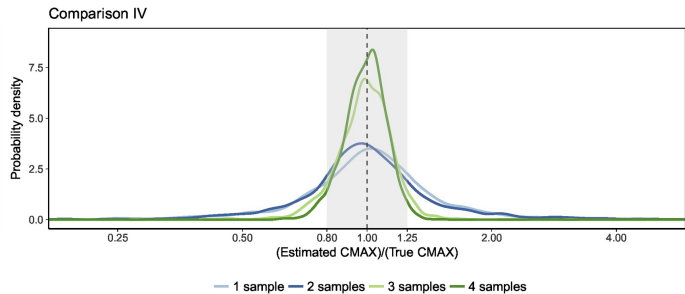
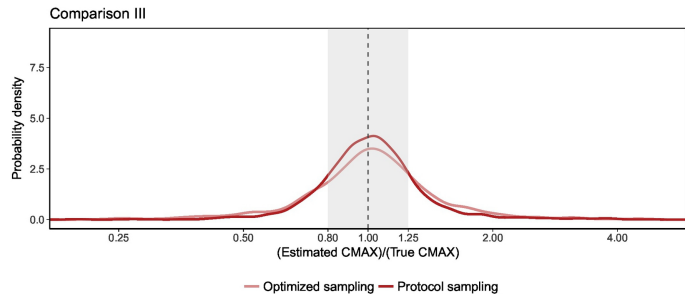
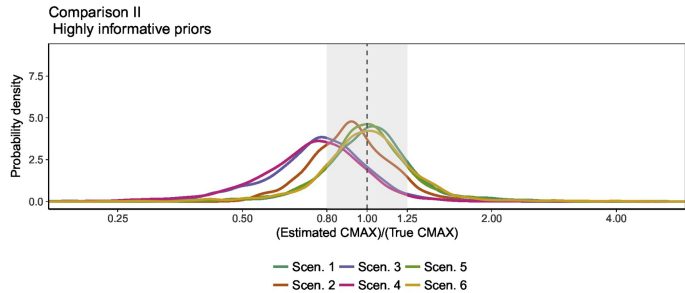
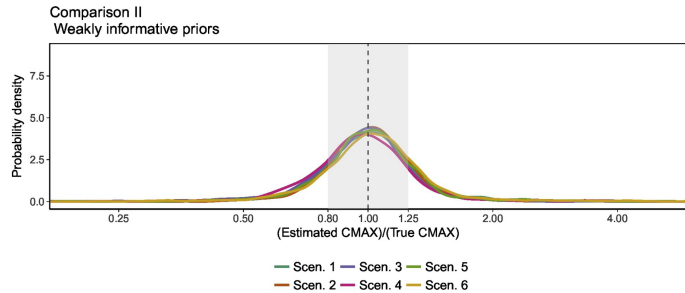


Figure 4