Response to:

"Guidance Document: Population Pharmacokinetics: July 2019"

Docket No. FDA 2019-D-2398

Submitted by: Justin Wilkins on behalf of the International Society of Pharmacometrics (ISoP)	
With contributions from, and on behalf of:	
	Roberta Bursi (InSilicoTrials) **Julie Desrochers (Syneos Health)** **Immanuel Freedman (Freedman Patent)** **Leonid Gibiansky (QuantPharm)** **Rakesh Gollen (Synteract Inc.)** **Nick Holford (University of Auckland)** **Chuanpu Hu (Johnson & Johnson)** **Mark Lovern (Certara)** **Pooja Manchandani (AbbVie Inc.)** **Krina Mehta (qPharmetra)** **France Mentre (INSERM/University of Paris)** **Samer Mouksassi (Certara)** **Nikolas Onufrak (ICPD; Clinical Pharmacometrics SIG)** **Teun Post (LAP&P)** **Mark Sale (Nuventra)** **Stacey Tannenbaum (Astellas)** **Max Tsai (Eli Lilly & Co)** **Janet Wade (Occams)** **Justin Wilkins (Occams; ISoP Board of Directors)** **…and others from the ISoP Population PK Guidance Working Group**

Table of Contents

1 INTRODUCTION AND SUBJECT OF COMMENTS 3
2 KEY POINTS FOR CONSIDERATION 4
2.1 Design Considerations 4
2.2 Using Population PK to Inform Drug Use 4
2.3 Planning 5
2.4 Model development 5
2.5 Data exclusions 6
2.6 Covariates 6
2.7 Model evaluation 7
2.8 Simulations 8
2.9 Reporting 8
2.10 Electronic File Submissions 9
2.11 Terminology 9
3 LINE BY LINE COMMENTS 11
3.1 Background 11
3.2 Application of Population PK Analysis 11
3.3 Data Used for Population PK Analysis 12
3.4 Data Analysis 12
3.5 Labeling based on the Results of Population PK Analysis 13
3.6 Population PK Study Reporting 13
4 REFERENCES 15

1 INTRODUCTION AND SUBJECT OF COMMENTS

On July 12, 2019, the U.S. Food and Drug Administration (FDA) announced the availability of a public docket to receive public comments on the draft guidance for industry entitled “Population Pharmacokinetics; Revised Draft Guidance for Industry”, through Federal Register (FR) notice (83 FR 14856).

This document represents the consensus view of the International Society of Pharmacometrics (ISoP) working group, which reviewed the draft document; we thank the FDA for the opportunity to provide feedback. Our key points for consideration are discussed in Section 2. We have provided additional detailed line-by-line comments and feedback in Section 3.

Throughout the document, the current draft guidance will be referred to as the “new guidance” and the 1999 Population PK guidance will be referred to as the “old guidance”. While the new guidance is less detail oriented and more ‘application oriented’ the FDA should ensure that the guidance still provides enough substantive information to guide sponsors in establishing population PK (PPK) approaches that are scientifically sound, value added and incorporate agency expectations for regulatory review so that information requests are reduced during the NDA or BLA review cycle.

2 KEY POINTS FOR CONSIDERATION

2.1 Design Considerations

The new guidance differs from the old guidance, where design considerations were outlined in detail in their own section "Study Design and Execution". In contrast, the new guidance discusses (in more general terms and in several places throughout the document) the influence of design (number of patients, sampling times, covariates...) on the results and the importance of its evaluation beforehand using simulations or optimal design tools. We think the more general approach is preferable, but we wish to make some comments for consideration to improve upon some of the design topics:

- Lines 115 and 259: The guidance suggests performing simulations to evaluate the power of covariate analysis or for drug-drug interaction (DDI) evaluation. This can also be done using an optimal design software tool (1). In addition, power calculations based on PPK models can be helpful and often require fewer patients than static statistical analyses because of the assumptions made by the PPK model. Informative priors can often dramatically reduce the number of patients required for a clinical trial. Sufficient rigor should be undertaken to provide evidence that statements made are supported.
- PK sampling schedule (Section IV. B.):
  - Although the practical reasons for using trough-only PK are clear, we believe that a more representative metric such as average concentration is a better empirical strategy and, if possible, should be encouraged in the guidance.
  - The statement on line 316 regarding "collect PK data from all patients" is vague. Additional details (for example, which development phases, how much data, which patient populations) would be more informative in assisting us to convince clinical and operational study teams the importance of PK sampling. Often, clinical and study operational teams will not endorse anything beyond very sparse sampling (i.e., trough only) and/or be from a subset of patients. Adding discussion on the limitations of subset PK sampling and applicability of trough-only data would be valuable. In general, support for avoiding favoring operational convenience over good science would be welcome.
  - It is typically not possible to make reliable observations of C_{max} because the true time of maximum concentration is not known, and all drug effects are necessarily delayed with respect to plasma concentrations. If the time of maximum response is the goal of the analysis then we propose samples should be based on both concentration and effect measurements.

2.2 Using Population PK to Inform Drug Use

The concept of "concentration boundaries" (also known as the "therapeutic range") is not described clearly in the guidance. The boundaries for dose adjustment may need to be individualized and may depend on ethnic and local factors in the practice of medicine (2).

2.3 Planning

A data analysis plan is considered an essential component of any analysis. Agency suggestions for what should be contained in analysis plans are welcome in addition to timing or pathway for agency review for analysis plans (e.g., with EOP2 or pre-NDA meetings or as a separate submission). We suggest that analysis plans should be required as appendices to PPK reports.

Some considerations for approaching analysis plans that could be considered for inclusion into the guidance:

- We propose that lines 268-270 should include that: 1) the studies included and 2) the rationale for studies or data to be excluded should be outlined clearly in analysis plans.
- We believe that extensive exploratory analysis results are more likely to incur selection bias and should be viewed as hypothesis-generating. In light of the learn-confirm paradigm (3), it is important to understand the extent of the "learn" and "confirm" aspects of the analysis findings (4,5).
- We suggest that a (brief) review of existing class and compound knowledge, existing models for the compound or related compounds, different analyses and other relevant information be provided in order to better ensure consistency.
- It would be helpful to encourage the use of informative prior information from prior modeling including PPK, physiologically based PK (PBPK), quantitative systems pharmacology (QSP) or epidemiological models to inform possible PPK model structures, parameters and associated variabilities.
- Lines 268-270: "The sponsor should … prespecify such omissions in the data analysis plan or study protocol." Generally, it is not possible to pre-specify all excluded data or individuals as some can only be identified after the dataset is available and exploratory data analysis is conducted. It is, however, feasible and recommended that the methods and criteria of exclusion of outlier records and individuals be specified in the analysis plan.

An important section explaining the requirements for assay description, which was in the old guidance, is missing in the new guidance. We propose that this be included with suggestions regarding defining the limits of quantitation and how to treat measurements outside the limits of quantitation in the PK analysis.

2.4 Model development

There are two options to compute shrinkage, based on standard deviation or based on variance. Shrinkage based on variance is approximately twofold larger than shrinkage based on standard deviation. Thus, any mention of specific numbers for shrinkage parameters should specify whether shrinkage of variance or shrinkage of standard derivations is being used. In addition, the guidance needs to distinguish between shrinkage of random effects and shrinkage of residual error.

Early access to concentration data from clinical trials may be needed to enable timely PPK and E-R analysis to support drug labeling and dose recommendations. Similar language was present in the old guidance but does not appear in the new guidance.

2.5 Data exclusions

Use of weighted residuals as the only criterion for data exclusion may be questionable as weighted residuals change with the model. A model with inappropriate structural and/or residual error parts may result in high weighted residuals for some points that are informative rather than outliers. With each model change, weighted residuals change, resulting in the need to exclude more and more points (and requirement to keep intermediate models and datasets available for examination).

While residuals are a good diagnostic tool to identify suspicious points, we do not believe this should be a universal tool for data exclusion. It should be used in extreme cases when no other reasons for exclusion can be found but the model fit is substantially influenced by outliers. The example of using absolute value of weighted residuals > 5 as a way to detect outliers is given; we suggest that care be taken to avoid giving the reader the impression that this is the only acceptable approach.

Lines 375-377: Outliers are often related to data errors, and requirement to investigate their influence by model refit needs to be qualified. Examples of circumstances in which refitting the model with outliers would not be appropriate include positive pre-first-dose observations, pre-dose observations mistakenly assigned post-dose times and vice versa, samples with very large times after dose caused by doses mistakenly omitted from the database, etc. The guidance should explicitly mention at least some of these cases so that the wording is not misinterpreted as the requirement to use all excluded data under all circumstances.

2.6 Covariates

Special consideration should be afforded to the role of body size and allometry throughout (6).

The division between intrinsic and extrinsic factors is somewhat arbitrary and we propose removing it.

Physiological plausibility should be considered when planning covariate scope.

Thresholds for clinical relevance/significance should be considered. Inclusion of small or poorly supported covariate effects can result in degraded precision of key model parameters and complicate subsequent simulations.

The number of subjects with a covariate is less important than the distribution of the covariate. For example, all subjects will have a weight and sex but if the weight and sex is the same in all subjects then nothing can be learned about the influence of these covariates. The range of a covariate is similarly not much use because it only describes the extreme outliers (two individuals) in the distribution. A 95% observation interval is recommended as a simple statistic that is superior to the range when describing continuous covariate distributions.

Drawing conclusions about the lack of a covariate effect should recognize the possibility of a Type 1 statistical error. This may occur, for example, when weight is tested as a covariate on clearance when the weight distribution is not sufficiently broad.

Generally, it is biologically highly implausible that clearance is not related to body size, and the failure to reject the null hypothesis for a covariate effect of weight should always be considered

suspicious. Indeed, it makes more biological sense to assume that weight is always a covariate for clearance and other PK model parameters such as volume of distribution (7).

The example of weight vs. creatinine clearance with respect to avoiding correlations between covariates is suboptimal: although there is undeniably a correlation between them in most subjects, they are separate physical descriptors, which are useful for different things. We propose making this argument with a less ambiguous example, such as weight vs. body surface area.

2.7 Model evaluation

We believe that "model validation" is not an appropriate term (8), and that the new guidance should instead use "model evaluation". One can never fully "validate" a nonlinear mixed-effects model, only evaluate it given its purpose.

An ISoP working group composed of internationally recognized experts in pharmacometrics published a tutorial covering model evaluation for longitudinal models with continuous data (9), but this appears to have escaped attention when preparing the new guidance. We recommend that due consideration be made of its contents, which we believe represent the current best practice in model evaluation.

Additional observations and suggestions regarding model evaluation in the updated guidance are below:

- Residual-based diagnostics are useful under most circumstances, but appropriate and relevant simulation-based diagnostics such as visual predictive checks (VPCs) are equally essential components of any PPK analysis, and are often more informative than standard goodness-of-fit (GOF) plots. We suggest that their use be strongly recommended by this guidance.
- There are different ways to compute a population prediction (PRED) and it is not always the prediction with the typical parameters that should be used for GOF plotting purposes. This has been extensively described in the ISoP tutorial quoted above. We used xPRED where x could be nothing, C, or P (PPRED is called EPRED in NONMEM). Particularly, CWRES should be plotted versus CPRED not PRED.
- There is no universally accepted definition of acceptable shrinkage (20% versus 30%) or of what is excessive. In addition, it should be noted that models that have high shrinkage might still be valid. This is also discussed in the ISoP guidance: "A shrinkage value of 30% or 50%, if calculated from SD or variance, respectively, has been suggested as a threshold for high shrinkage, but whether this threshold should be applied for all models and population parameter values remains to be evaluated." We agree that shrinkage should always be reported since it can result in loss of power (10,11).
- Using normalized prediction distribution errors (NPDE) as a simulation-based model evaluation tool is not mentioned, although it is now available in most PPK software. In the ISoP tutorial, some graphs based on NPDE vs. time and xPRED are part of the core set of recommended diagnostics.
- Not all estimation methods provide eigenvalues to determine condition number. In addition, it should be noted that models with high condition number might still be appropriate and useful.

2.8 Simulations

We agree that parameter uncertainty should be included under some circumstances (consideration of parameter uncertainty is appropriate when simulating new studies, for example, but usually not necessary when preparing simulation-based diagnostics such as VPCs).

The purpose of simulations is to answer questions of key concern. The choice of variability and uncertainty sources considered in simulations is a natural consequence of the nature of the question to be answered. Rather than having subsections based on what variability/uncertainty is incorporated, we think it would be more intuitive to have subsections by the question type:

1. Simulations that answer questions of central tendency
2. Simulations that explore the plausible range of true parameter values
3. Simulations that answer questions about exposure ranges in populations of interest
4. Simulations that evaluate clinical trial designs

2.9 Reporting

We suggest the Agency considers recommending the inclusion of a table of assumptions, as defined by Marshall and colleagues, in PPK reports (12).

Having both an executive summary and synopsis seems excessive. We suggest that only one summary should be necessary (the name is not very important, it could be either "executive summary" or "synopsis"). More thought needs to be put in this section to provide a meaningful content. One option might be to include:

- A plain language summary of objectives, data, and methodology;
- Summary of results and conclusions;
- The key findings that affect approval or labeling decisions;
- Any other recommendations based on the PPK analysis.

The discussion section should include a critical evaluation of how well the model-predicted distribution of concentrations is matched by the observed distribution of concentrations (as shown in VPCs).

Only relevant runs should be included in the run record; there is no need to include all model runs.

Scripts such as user-defined R code should be made available upon request. PPK reports are already very lengthy.

The data section appears to mix methodology and specific results. It would be more logical to split it into a methodology section that describes the studies used in the analysis (the first 5 bullet points of the data section of Table 1) and a results section that describes specific characteristics of the data set (the last 2 bullet points of the data section of Table 1).

We suggest highlighting that the language used in this section is guidance and not mandatory under all circumstances, with reports focusing on the specific questions being addressed. Several points from this section are excessive for a technical report (clinical relevance, alternative dosing

strategies and development/regulatory decisions are not always straightforward to provide and may be more appropriate to cover elsewhere).

2.10 Electronic File Submissions

This requirement represents a significant increase in the scope of files that need to be submitted. It is understandable that the FDA asks for this additional information in order to complete a thorough audit of PPK analyses. It will, however, be somewhat onerous for the sponsor to package this information, and for the FDA to unpack it considering the current electronic submission system, which is archaic and burdensome. The request to submit these model files, which are typically in file formats not accepted by the gateway, may incur not only significant cost to sponsors but also delays to submission due the resource capacity and limited availability of programmers to convert the file formats to 'submission ready' .xpt and .txt formats.

Before implementing this request, we suggest a more modern, secure, compliant data submission system be considered for adoption.

2.11 Terminology

There are some instances where terminology in the guidance could be better described and/or more consistently used. Some noted terms are discussed below:

- 'Clinically relevant' seems to be defined in a manner consistent with the ICH E5 guideline, i.e., as clinically significant. Greater clarity on what is meant by this term would be helpful (2).
- What is meant by "PK properties"? This should be clearly defined in order to understand what factors support the intended objective.
- How is "acceptable bias" for describing data defined?
- Biological parameters such as clearance are not known. They can only be estimated. Unless the true value is known then it is not possible to determine the bias of a parameter.
- The "fit for purpose" concept needs to be defined a priori in the analysis plan, in order to avoid inconvenient results being ignored when they do not agree an a posteriori definition of "fit for purpose".
- Page 16, "Methods":
  - It is not clear what "transformations of parameters" might be referred to here. We propose this be clarified.
  - The choice and justification of the 'model-estimation method' is a vague term: multiple estimation methods may have been appropriate and justified by practical considerations including the evaluation and/or fidelity criteria.
- Page 16, "Results": The statement at the bottom of the page about variability being reported as CV% and precision reported RSE% could be made clearer by giving an unambiguous definition of variability and uncertainty. The use of CV% to report the apparent coefficient of variation of a log-normal random effect distribution would often be reasonable but could be confusing if the random effect distribution was normal.

- Page 15, “Synopsis”: The term ‘sufficiency of the data” is not well defined. Does this relate to the required accuracy of PK parameters or concentration predictions?
- There are a number of subjective terms in the document (e.g., “adequate”, “sufficient”). More precise descriptions would be helpful.

3 LINE BY LINE COMMENTS

3.1 Background

- The response to 'extrinsic patient factors' such as food consumption or concomitant medication may depend on 'intrinsic patient factors' such as enzyme and/or transporter expression and polymorphism or hepatic impairment: e.g., food effect in subjects taking Proton Pump Inhibitors or diabetic subjects with low enzyme expression.

3.2 Application of Population PK Analysis

- Lines 83-84 (also 116-123, 207-210, 257-258): FDA may encourage Phase 3 study protocols to pre-specify adequate numbers of patient subgroups representative of the real-world population to be enrolled.
- Based on knowledge gained in Phases 1-2, allowing confirmation/refinement of the effects of influential covariates on dosing and enhancing model generalizability.
- Line 94-96: Suggest to add "clinically meaningful" as statistically significant effect may have no clinical relevance.
- Line 113 Section III.A.2: This section confounds the power to detect a covariate effect (which will depend on the study design and distribution of the covariates) and the precision of a PK parameter such as clearance.
- A parameter describing a covariate effect is not a PK parameter.
- Line 115 and line 259: the guidance suggests performing simulations to evaluate the power of covariate analysis or for DDI evaluation. This can be done using optimal design software tool also (1).
- Line 116-120: Power calculations based on PPK models can be helpful and often require fewer patients than static statistical analyses because of the assumptions made by the PPK model. Furthermore, informative priors can often dramatically reduce the number of patients required for a clinical trial.
- Line 131-133: Add "if appropriate" as exposure metrics listed may only represent a single snapshot in time and may not be appropriate in some situations where there are time-dependent effects.
- Line 132: C_{min} cannot describe the average steady state drug exposure. Necessarily C_{min} must be below the average. This statement should be changed.
- Line 163: The term sampling windows should be defined. There is no a priori reason why a sampling window should be larger in children. The sampling window may be large in adults if the study design only allows sparse sampling. In our experience, it is more often the case that sampling is sparse in adult studies rather than in studies in children.
- Line 171-175: Considerations of ontogeny disease similarity and E-R may be also adequately handled by PBPK models provided by tools that provide suggestions as the functional form of enzyme maturation for scaling to pediatric physiology.
- Line 173: A reference to Rhodin et al (13) would be suitable here to justify the assertion about physiology in those less than 2 years old.
- Line 175: "General Clinical Pharmacology Considerations for Pediatric Studies for Drugs and Biological Products" referenced twice (5 and 6).

- Line 180: The use here of "late-stage" may suggest exclusion of relevant Phase 1 data in therapeutic areas like, for instance, oncology and/or orphan diseases. Suggestion could be: "using data from patients."
- Line 229: The use of the wording "could conceivably be used" in this example is in contrast with the other examples where "can be used" was chosen. Further, the combination of the conditional "could" with "conceivably" sounds exaggerated. Wording like "can be used" should also apply to this example.

3.3 Data Used for Population PK Analysis

- Line 268: Including a preponderance of HV studies could potentially bias results in the population of interest. Perhaps "all available data" could be changed to "all available and relevant data".
- Line 292: Biological parameters such as clearance are not known. They can only be estimated. Unless the true value is known then it is not possible to determine the bias of a parameter.
- Line 301: In the list of references for optimal design of the PPK model only Nyberg et al 2015 is a review, the others are specific articles and are not the most recent ones. Perhaps the guidance should retain only Nyberg et al?
- Line 320: Missing data e.g., due to patient dropout, has to be dealt with after the fact. It is not possible a priori to require that patients with missing data do not have adverse effects. We suggest this statement should be reworded.

3.4 Data Analysis

- Line 332: The citation "Bonate and Steimer 2013" should be "Bonate PL (2011). Pharmacokinetic-Pharmacodynamic Modeling and Simulation, 2nd edition. New York, Springer."
- Line 342: What is the Agency's position in regards to what correlation between covariates would be considered high? (>50%?)
- Line 365: Change "or allometric principles" to "and allometric principles" and add a reference to theory based allometry such as West & Brown (14) and its experimental confirmation in a PK context (15,16).
- Line 367: One issue we have encountered frequently is that analysts will only look at the overall proportion of BQL values and conclude that if they are less than 10%, likelihood based methods need not apply. However, if %BQL values by time since last dose are assessed, the fraction at later time points is often much higher and merits accounting for. Therefore, adding a sentence here would be helpful.
- Line 371: "The sponsor should justify their methodological approach with regard to missing data and outliers and provide a sensitivity analysis." Please be more specific on what 'sensitivity analysis' means in this context (method comparison? With and without certain data? Comparing imputation methods?).
- Line 373-383: Although abs(CWRES > 5) is a well-known criterion for outlier detection, other mature approaches include Random Sample Consensus (RANSAC), which iteratively locates inliers and is often reliable to locate and fit inliers including up to ~50% outliers from long-tailed distributions such as the Cauchy distribution.

- Lines 393-394: Should the sponsor propose the acceptable level of bias and precision to demonstrate their case? Does FDA mean the precision and bias in terms of model residuals and the selection of modeling algorithms according to simulations that demonstrate such precision and bias? Alternatively, does FDA mean the precision and bias of parameters estimated from models simulated and evaluated during the development of a prospective PPK analysis plan?
- Lines 393-394: the sentence "An appropriate model should be biologically plausible, consistent with current knowledge, and have mathematically identifiable parameters" does not belong to section C, Model Validation; it should be moved to section B, Model Development.
- Line 486-487: Why does the model always need to be revalidated (or re-evaluated) for a new purpose? An example would be a Phase 2 model to translate across populations or diseases in Phase 3.
- Line 518-520: The parameter uncertainty can be addressed if it can be reliably estimated. Although the parametric bootstrap based on refitting models is widely used for this purpose, let us note that close to an optimum, it may be appropriate to estimate the parameter distribution based on bootstrapping the individual model-based NCA estimates or Empirical Bayes Estimates as a reasonable and computationally inexpensive approximation to a full parametric bootstrap based on refitting the data. Does this simply refer to including parameter uncertainty as well as variability in simulations?

3.5 Labeling based on the Results of Population PK Analysis

- Lines 532-536 (see also Lines 190-198): FDA may encourage that information facilitating dose optimization/adaptation, as derived from PPK analyses, be contained within the product package insert:
  - Provision of model-derived target exposure ranges (e.g., C_{max}, AUC_{24}), suggested PK sampling windows, and guidance for PK parameter derivation will assist clinicians in tailoring prescription to the individual patient.

3.6 Population PK Study Reporting

- Table 1, Line 613: The Discussion section should include a critical evaluation of how well the model predicted distribution of concentrations is matched by the observed distribution of concentrations (as shown in VPCs).
- Table 1: The term "sufficiency of the data" is not well defined. Does this relate to the required accuracy of PK parameters or concentration predictions?
  - What are 'clinically relevant' drug exposures? Does this term refer to the likely range of exposures given the regimens and the planned dose or approved dose range?
  - The term 'final model' implies that there will never be any others. 'Best model' or 'selected model' may be appropriate for the model that was eventually selected.
  - The choice and justification of the 'model-estimation method' is a vague term: multiple estimation methods may have been appropriate and justified by practical considerations including the evaluation and/or fidelity criteria.
  - 'Physiological plausibility' is a vague term. Plausibility depends on current scientific knowledge and should be treated with caution (a point already made by Bradford

Hill) since discoveries (such as previously unknown or unrecognized transporters) are quite possible.

- Table 1, Line 613: Only relevant runs should be included in the run record; there is no need to include all model runs.
- Table 1, Line 613: Scripts such as user-defined R codes should be made available upon request. PPK reports are already very lengthy.
- Lines 627-628: FDA may consider suggesting/requiring the public release of PPK models at the time of drug approval:
  - Immediate availability of models allows for their rapid uptake into model-informed precision dosing platforms and fosters academic/collaborative efforts at model refinement using real-world evidence.

4 REFERENCES

1. Retout S, Comets E, Samson A, Mentré F. Design in nonlinear mixed effects models: Optimization using the Fedorov-Wynn algorithm and power of the Wald test for binary covariates. Stat Med. 2007 Dec 10;26(28):5162–79.

2. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. Ethnic Factors in the Acceptability of Foreign Clinical Data: ICH E5(R1) [Internet]. Geneva; 1998. Available from: https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E5_R1/Step4/E5_R1__Guideline.pdf

3. Sheiner LB. Learning versus confirming in clinical drug development. Clin Pharmacol Ther. 1997 Mar;61(3):275–91.

4. Ribbing J, Nyberg J, Caster O, Jonsson EN. The lasso - A novel method for predictive covariate model building in nonlinear mixed effects models. J Pharmacokinet Pharmacodyn. 2007;34(4):485–517.

5. Hu C, Zhang J, Zhou H. Confirmatory analysis for phase III population pharmacokinetics. Pharm Stat. 10(1):14–26.

6. Holford NHG, Anderson BJ. Allometric size: The scientific theory and extension to normal fat mass. Eur J Pharm Sci. 2017 Nov 15;109:S59–64.

7. European Medicines Agency. How should results/predictions of pharmacokinetic analyses be presented to facilitate decision making about the adequacy of the proposed dosing regimen in paediatric patients? [Internet]. Modelling and simulation: questions and answers. 2018. Available from: https://www.ema.europa.eu/en/human-regulatory/research-development/scientific-guidelines/clinical-pharmacology-pharmacokinetics/modelling-simulation-questions-answers

8. Mentré F, Escolano S. Prediction discrepancies for the evaluation of nonlinear mixed-effects models. J Pharmacokinet Pharmacodyn. 2006 Jun;33(3):345–67.

9. Nguyen THT, Mouksassi M-S, Holford N, Al-Huniti N, Freedman I, Hooker AC, et al. Model Evaluation of Continuous Data Pharmacometric Models: Metrics and Graphics. CPT Pharmacometrics Syst Pharmacol. 2017 Feb;6(2):87–109.

10. Combes FP, Retout S, Frey N, Mentré F. Powers of the likelihood ratio test and the correlation test using empirical bayes estimates for various shrinkages in population pharmacokinetics. CPT pharmacometrics Syst Pharmacol. 2014 Apr 9;3:e109.

11. Xu XS, Yuan M, Yang H, Feng Y, Xu J, Pinheiro J. Further Evaluation of Covariate Analysis using Empirical Bayes Estimates in Population Pharmacokinetics: the Perception of Shrinkage and Likelihood Ratio Test. AAPS J. 2017 Jan 1;19(1):264–73.

12. Marshall S, Burghaus R, Cosson V, Cheung S, Chenel M, DellaPasqua O, et al. Good Practices in Model-Informed Drug Discovery and Development: Practice, Application, and Documentation. CPT Pharmacometrics Syst Pharmacol. 2016;5(3):93–122.

13. Rhodin MM, Anderson BJ, Peters AM, Coulthard MG, Wilkins B, Cole M, et al. Human renal function maturation: a quantitative description using weight and postmenstrual age. Pediatr Nephrol. 2009 Jan;24(1):67–76.

14. West GB, Brown JH. The origin of allometric scaling laws in biology from genomes to ecosystems: Towards a quantitative unifying theory of biological structure and organization. Vol. 208, Journal of Experimental Biology. 2005. p. 1575–92.

15. Holford NHG, Ma SC, Anderson BJ. Prediction of morphine dose in humans. Paediatr Anaesth. 2012 Mar;22(3):209–22.

16. McCune JS, Bemer MJ, Barrett JS, Baker KS, Gamis AS, Holford NHG. Busulfan in infant to adult hematopoietic cell transplant recipients: A population pharmacokinetic model for initial and bayesian dose personalization. Clin Cancer Res. 2014;20(3):754–63.