Investigating the contribution of residual unexplained variability components on bias and imprecision of parameter estimates in population pharmacokinetic mixed-effects modeling

Mutaz M. Jaber[1,2] · Richard C. Brundage[1,3]

Abstract

In a nonlinear mixed-effects modeling (NLMEM) approach of pharmacokinetic (PK) and pharmacodynamic (PD) data, two levels of random effects are generally modeled: between-subject variability (BSV) and residual unexplained variability (RUV). The goal of this simulation-estimation study was to investigate the extent to which PK and RUV model misspecification, errors in recording dosing and sampling times, and variability in drug content uniformity contribute to the estimated magnitude of RUV and PK parameter bias. A two-compartment model with first-order absorption and linear elimination was simulated as a true model. PK parameters were clearance 5.0 L/h; central volume of distribution 35 L; inter-compartmental clearance 50 L/h; peripheral volume of distribution 50 L. All parameters were assumed to have a 30% coefficient of variation (CV). One hundred in-silico subjects were administered a labeled dose of 120 mg under 4 sample collection designs. PK and RUV model misspecifications were associated with relatively larger increases in the magnitude of RUV compared to other sources for all levels of sampling design. The contribution of dose and dosing time misspecifications have negligible effects on RUV but result in higher bias in PK parameter estimates. Inaccurate sampling time data results in biased RUV and increases with the magnitude of perturbations. Combined perturbation scenarios in the studied sources will propagate the variability and accumulate in RUV magnitude and bias in PK parameter estimates. This work provides insight into the potential contributions of many factors that comprise RUV and bias in PK parameters.

Keywords Residual unexplained variability · Nonlinear mixed-effects · Population pharmacokinetics · Sources of variability

Introduction

In nonlinear mixed-effect modeling (NLMEM) of pharmacokinetic (PK) and pharmacodynamic (PD) data, we generally refer to two levels of random effects: between subject variability (BSV), and residual unexplained variability (RUV). The former quantifies the variability of PK/PD parameters within the population. The latter, RUV, quantifies the residual variability in the model after accounting for the PK structural model, the covariate model, and the between-subject variability model.

We generally assume when we enter PK data into a model that most measurements are absolute with no variability, or at least an ignorable amount. That is rarely the case even in a well-controlled clinical trial. Probably the easiest variability to understand is variability due to the imprecision of our analytical techniques and is considered a given in the remainder of this paper. Other potential errors can occur, and we categorize those errors into three types: (i) Manufacturing, (ii) clinical trial execution, and (iii) technical execution errors.

During the manufacturing process, certain amount of deviation from the labeled dose is allowed. Manufacturing error includes this content uniformity variability that, in general, the US Pharmacopeia recommends to be less than 15% [1]. Clinical trial execution errors can occur in

✉ Richard C. Brundage
brund001@umn.edu

1 Department of Experimental and Clinical Pharmacology, College of Pharmacy, University of Minnesota, Minneapolis, MN, USA

2 Clinical pharmacology and Pharmacometrics, Gilead Sciences, Inc., Foster City, USA

3 Metrum Research Group, Tariffville, CT, USA

recording the time the dose is given as well as in recording the time of sampling blood [2, 3]. Technical errors include misspecification of the pharmacostatistical model components including the PK structural model and RUV statistical model. Other technical errors include misspecification of covariate models or BSV models, but they are not addressed in this analysis.

To date, the degree to which these sources of variability contribute to the estimated value of RUV is not well characterized. The goal of this study is to quantify the extent to which pharmacostatistical model (PK and RUV) misspecification, errors in recording dose time, errors in recording blood sampling time, and drug content uniformity contribute to the value of RUV and bias and imprecision of PK parameter estimates. We further evaluate the impact of sample collection density on the results.

Methods

Literature review

To get a sense of RUV magnitude in recently published models, we evaluated studies published from 2019 to 2021 from any journal indexed in PubMed and Google Scholar. Keywords included: "Population pharmacokinetics", "NONMEM", and "First-order conditional estimation with interaction". Our inclusion criteria were a pharmacokinetic structural model, first-order conditional estimation, and the usage of NONMEM software. For simplicity, we captured the magnitude of RUV from proportional error models and the proportional term from combined error models. Papers reported the RUV as a coefficient of variation (CV) calculated as the square root of the variance or using a more exact equation [4]. When RUV was reported as a variance, values were converted to a CV using the square root of the variance estimate. An analytical methods section was sometimes included and the CV of the assay was recorded when available.

Pharmacostatistical model

A single 120 mg dose study with a two-compartment model, first-order absorption, and first-order elimination was assumed as the true simulated model. This provided a rapid distributive (alpha) half-life of 0.28 h and a slower terminal elimination (beta) half-life of 12.2 h. Individual PK parameters were assumed to have lognormal distributions with population typical values and CVs as reported in Table 1. A 5% CV for proportional RUV magnitude served as a somewhat optimistic baseline for analytical variability.

Table 1 Parameter values used in the simulation

Parameter	Value	BSV (CV%)
CL (L/h)	5	30
V (L)	35	30
Q (L/h)	50	30
V_P (L)	50	30
K_a (1/h)	0.7	30
σ_{prop} (%)	5	

Drug clearance, *CL* Central volume of distribution, *V* Intercompartmental clearance, *Q* Peripheral Volume, *VP* Absorption rate constant, *Ka* Coefficient of variation for residual unexplained variability, σ BSV Between-subject variability

Sample collection design

We conducted our simulations with 4 levels of sample collection designs; (i) SD9: 9 samples per subject with nominal sampling times of 0.5, 1, 2, 4, 6, 8, 12, 24 and 48 h; (ii) SD7: 7 samples per subject with nominal sampling times of 2, 4, 6, 8, 12, 24, and 48 h; (iii) SD5: 5 samples per subject with nominal sampling times of 2, 8, 12, 24, and 48 h; and (iv) SD4: 4 samples per subject with nominal sampling times of 2, 12, 24, and 48 h. Sample collections were selected based on a typical pharmacokinetic study design for the given half-life and optimal design was not evaluated in this study.

Perturbations to the model

Starting with the reference pharmacostatistical model, various perturbations from the nominal values were made to the simulation dataset (*vide infra*). In addition, several levels of perturbations were included. For each perturbed dataset, 500 replicates of 100 *in-silico* subjects were simulated. These datasets included the perturbed item of interest and simulated concentrations based on those perturbations. These datasets were then altered to include the original nominal value while maintaining the perturbed concentrations for estimation creating a mismatch between the times and concentrations. Table 1S details the simulation and estimation models used in this study.

Manufacturing perturbation (MP)

Let us assume that the protocol labeled dose (P_D) is 120 mg. Since the amount of drug in any individual dosing unit can vary from that, we model the true individual dose (I_D) for the ith subject as:

$$I_{D,i} = P_D + \kappa_i$$

where κ_i is a normally distributed random variable with mean of 0 and standard deviation of K. K was chosen to represent two levels of MP variability, CV = 10% and 20% of D_L. Therefore, the model was independently simulated using $I_{D,i}$ under the two scenarios.

Clinical trial execution perturbations (CTEP)

Given the protocol dosing time (P_{DT}), the individual true dosing time (I_{DT}) for the ith subject was simulated as:

$$I_{DT,i} = P_T + \gamma_i$$

where γ_i is a normally distributed random variable with mean of 0 and standard deviation of G. G was chosen to have two levels of variability associated with a standard deviation of SD = 5 min, and 10 min. Similarly, given the protocol sample collection times for the ith subject and jth point ($P_{ST,i,j}$), the true sample collection times at the ith point and jth subject ($I_{ST,i,j}$) is:

$$I_{ST,i,j} = P_{ST,i,j} + \rho_{i,j}$$

where $\rho_{i,j}$ is a normally distributed random variable with mean of 0 and standard deviation of R. Four levels of variability were selected for R associated with standard deviations of SD = 5, 10, 15, and 30 min.

Technical execution perturbations (TEP)

For PK model misspecification, we chose to fit the data using a one-compartment model with first-order absorption and first-order elimination instead of the true two-compartment model mentioned above.

A combined proportional and additive RUV was used as the perturbation. An additive component was added to the 5% proportional component and simulated with 3 different magnitudes. The approximate lower 10th percentile of simulated concentrations was 0.1 mg/L. Specific perturbations of 10% (σ_{Add} = 0.01 mg/L), 20% (σ_{Add} = 0.02 mg/L) and 40% (σ_{Add} = 0.04 mg/L) of this relatively low concentration were included as the additive component in the simulation. In these scenarios, data were fit using a proportional error model instead of combined error model.

Combined perturbations

Additional simulations were done using the true two-compartment model with the combined perturbations $I_{D,10\%}$, $I_{DT,5}$, $I_{ST,5min}$, and σ_{Add} = 0.04 while fitting the data using a one-compartment model.

Software and evaluation

The R package mrgsolve version 1.04 [5] was used to generate the simulated data. All generated datasets were carried over to the estimation process using NONMEM version 7.5 (ICON plc Development Inc.). Parameter estimation was obtained using first-order conditional estimation with interaction (FOCE-I). ADVAN4/TRANS4 was used for two-compartment model fitting and ADVAN2/TRANS2 was used to fit a one-compartment model. Python (Python Software Foundation. Python Language version 3.9) was used to automate the assigning, creating control streams and executing NONMEM for estimation. All programs/scripts used in this project can be found in the MMJ GitHub repository (https://github.com/Mutaz94/residual) under open-source creative commons attribution 4.0 license with a detailed documentation for research reproducibility.

Deviations of parameter estimates ($\widehat{\theta}$) from the true parameter value (θ_T) were computed as fractional deviations (ψ) :

$$\psi = \frac{\widehat{\theta}}{\theta_T}.$$

We wish to define a term that captures deviations of parameter estimates that is caused by variabilities introduced by components of real world RUV. Although the terms bias and imprecision may seem appropriate, it must be kept in mind that bias and imprecision are not a function of the data themselves. If bias and imprecision are strictly related to an estimation method, perhaps erroneous data and subsequent erroneous parameter estimates should not be labeled as bias. Nonetheless, we refer to bias as the deviation from original parameter estimates and imprecision as the size of typical deviation. Relative bias (rBias), and relative root mean squared error (rRMSE) were calculated to assess the bias and imprecision of estimation, respectively.

$$rBias(\%) = \frac{1}{n}\sum_{i=1}^{N}(\psi_i - 1) \cdot 100\%$$

$$rRMSE(\%) = \sqrt{\frac{1}{n}\sum_{i=1}^{N}(\psi_i - 1)^2} \cdot 100\%$$

For combined perturbations, the median RUV magnitude for the kth sample collection design ($mCV_{RUV,k}$) is calculated as:

$$mCV_{RUV,k} = CV_{RUV,true} \cdot median(\Psi_{RUV,k})$$

and a corresponding interquartile range (25th and 75th percentiles).

Results

Literature review

In this study, 330 published studies from the literature were evaluated. In 239 of the 330 studies, authors used a proportional error model and 89 used a combined error model. Two used an additive error model which was not included in the analysis. The median and interquartile range [25th–75th] of proportional RUV CV was 25% [17.6%–37.1%]. The median reported analytical assay (n = 72) CV was 8.9% [5%–11%]. Additional information is found in the supplementary material.

Manufacturing perturbation (MP)

MP based on 4 different study designs is demonstrated in Fig.1. Dosage content uniformity that reflects on relative bioavailability of the drug had negligible effect on RUV magnitude and it contributes more to the value of rBias in random effects (BSV) on drug clearance (CL) and central volume of distribution (V). BSV on V had the larger effect by dosage content perturbations with bias up to 140% compared to CL that had bias up to 20%. A sparse study design had more rBias than an intense design on V.

Clinical trial execution perturbations (CTEP)

With CTEP, Figs. 2 and 3 demonstrate the contribution of sampling and dose times on RUV magnitude and bias and imprecision on PK parameter estimates, respectively. We saw a small effect on inaccurate dosing time data with the 5-minute perturbation on RUV magnitude which monotonically increases with higher perturbations. However, bias is noticeable with some estimated parameters. For instance, the typical value for V and corresponding BSV are sensitive to 5-minute perturbation and even larger with 10-minute perturbation. On the other hand, perturbations did not have much of effect on CL and intercompartmental clearance (Q). While dosing time perturbations have a small effect on RUV, inaccurate sampling time data has a more noticeable rBias. The magnitude of RUV will increase with increasing the magnitude of inaccurate data. This pattern is more pronounced with intense designs compared to sparse design.

Technical execution perturbations (TEP)

Figures 4 and 5 present the TEP deviation from the baseline RUV magnitude for both RUV model misspecifications. Residual error misspecification contributed the largest inflation in RUV magnitude compared to other
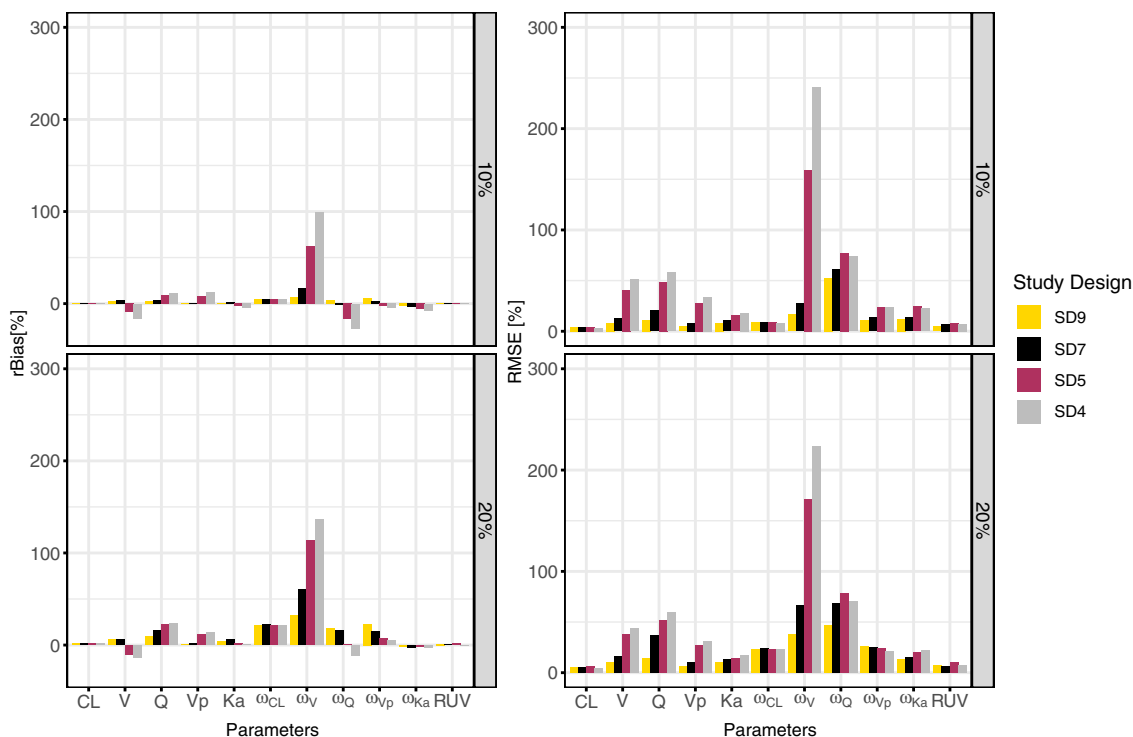

Fig. 1 Contribution of manufacturing perturbation on RUV and PK parameters; left-hand panel displays the relative bias; right-hand panel displays the relative mean squared error. The Greek symbol Omega present between-subject variability magnitude. The grey title boxes represent the perturbation magnitude
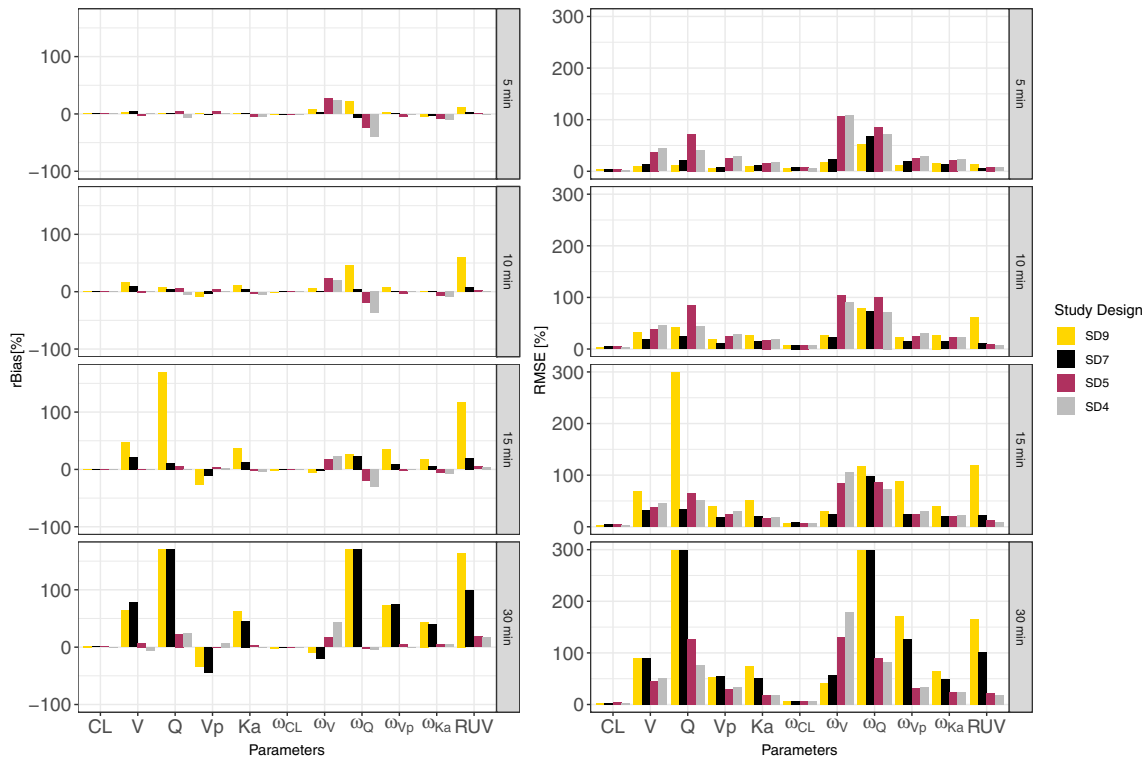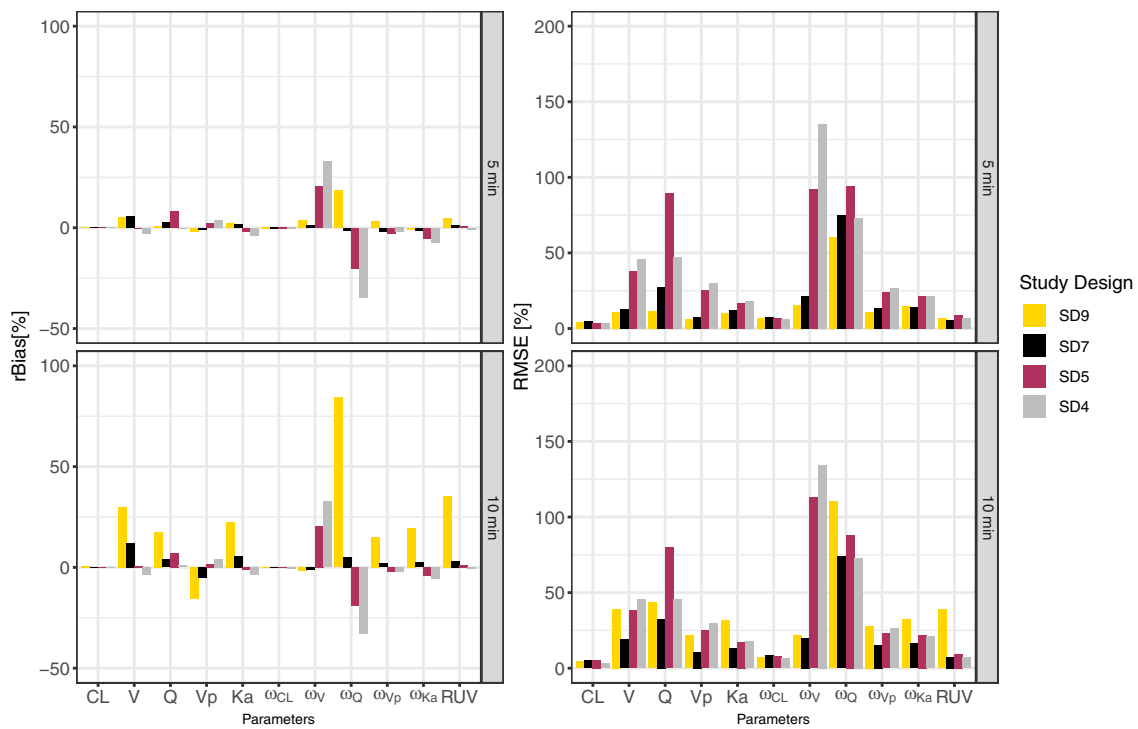
Fig. 2 Contribution of inaccurate PK sampling time data on RUV and PK parameters; left-hand panel displays relative bias; right-hand panel displays the relative mean squared error. The Greek symbol Omega present between-subject variability magnitude. The grey title boxes represent the perturbation magnitude
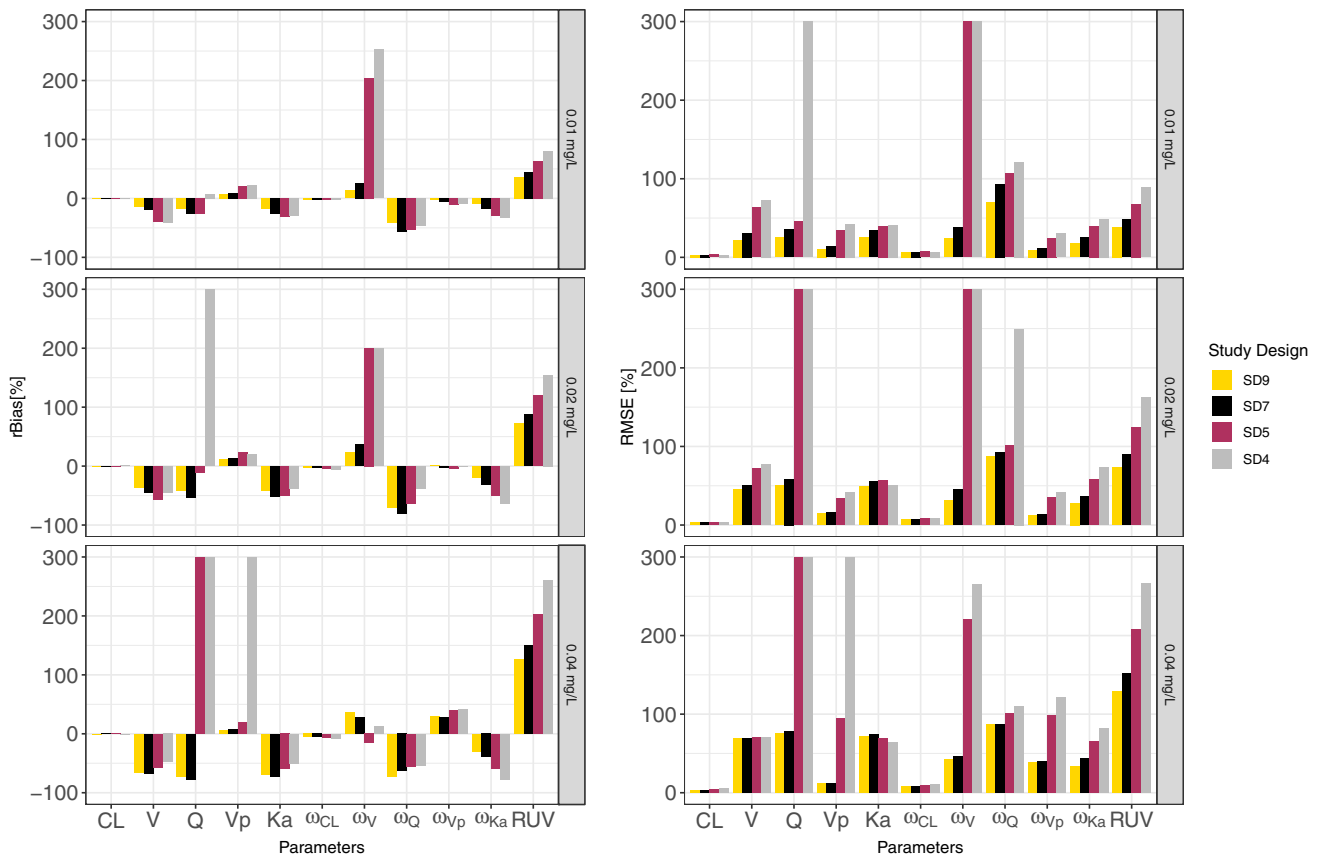

Fig. 3 Contribution of inaccurate dosing time data on RUV and PK parameters; left-hand panel displays relative bias; right-hand panel displays the relative mean squared error. The Greek symbol Omega present between-subject variability magnitude. The grey title boxes represent the perturbation magnitude

Fig. 4 Contribution of residual model misspecification on RUV and PK parameters; left-hand panel displays absolute bias; right-hand panel displays the relative mean squared error. The Greek symbol Omega present between-subject variability magnitude. The grey title boxes represent the additive standard deviation magnitude
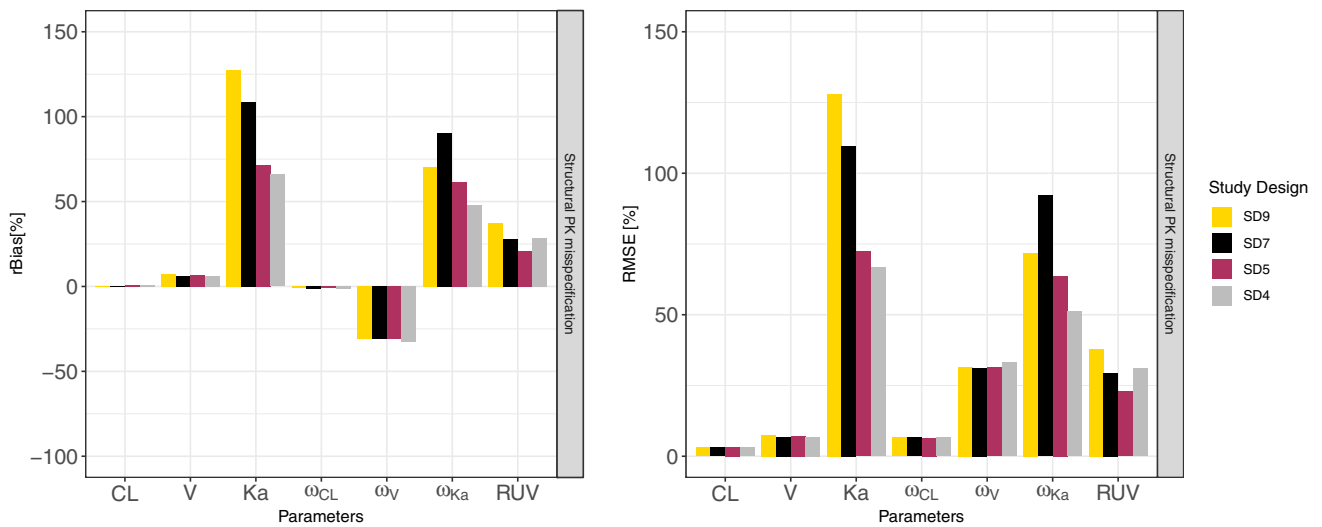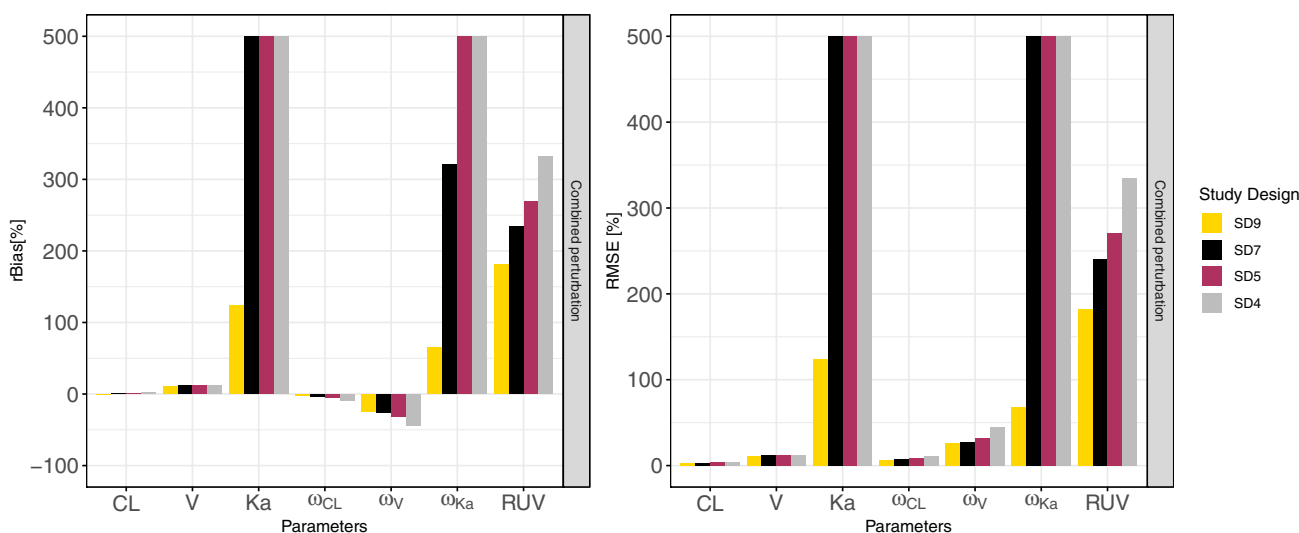

Fig. 5 Contribution of structural PK model misspecification on RUV and PK parameters; left-hand panel displays absolute bias; right-hand panel displays the relative mean squared error. The Greek symbol Omega present between-subject variability magnitude

Fig. 6 Contribution of combined perturbations on RUV and PK parameters; left-hand panel displays relative bias; right-hand panel present the relative mean squared error. The Greek symbol Omega present between-subject variability magnitude. The figure was truncated to a reasonably high value (500%), maximum observed value > 10^{50}

Combined perturbations

Combined perturbations are presented in Fig. 6. The absorption rate constant, Ka, has high bias and imprecision with combined perturbation for sample collection designs SD7, SD5, and SD4 thus, the figure was truncated to a reasonably high value (500%).

The median RUV under the combined perturbations for different sample collection designs is shown in Table 2.

Discussion

Residual unexplained variability is an outcome of the NLMEM analysis. It is most often reported without much regard as to what constitutes it. Assay variability is the obvious component and information regarding this is often reported in publications as a % CV of the assay. In our literature review, we found the median % CV to be 8.9% over a three-year review of published articles. This can be compared to the reported RUV in those same papers. The proportional component of RUV in these papers was 25%. Clearly, there are other components that contribute to RUV. The reader is referred to the supplemental material for a complete listing of the analytical variability and reported RUV.

The selection of 10%–20% variability in dose content was based on the criteria for US pharmacopeia on batch variability with a guideline of uniformity should be less than 15% variability on average [1]. Thus, it is reasonable that the produced batch has a variability in the content corresponding to a 10% CV. We also examined 20% to better understand the behavior of more extreme variability in the content uniformity. The PK parameter V was more affected by dosage content perturbations with bias in BSV up to 140% compared to CL that had bias up to 20% with extreme 20% CV. The large bias can be explained due to the high imprecision in relative bioavailability; thus, the random effect could not be precisely estimated. A sparse study design had more bias than an intense design on V. The reason is likely due to the greater number of samples in the early distributive phase compared to the former. The low density sample design has more bias than intense design with Q and Vp BSV parameters which is expected

Table 2 Median RUV magnitude for combined sources based on different sample collection designs

Sample collection design	Median RUV magnitude (CV)	Range, CV [25th–75th]
SD9	14.1	(13.8–15.2)
SD7	16.7	(15.5–17.0)
SD5	18.4	(17.5–19.6)
SD4	21.6	(20.4–22.8)

SD Sample collection design with corresponding number of samples, *CV* Coefficient of variation

due to practical identifiability related to study design [6]. The estimation of these parameters requires intense sampling and they are sensitive to the times of sampling.

Data quality in a clinical trial is an important factor to obtain reliable PK/PD model estimates [7]. In regard to CTEP, we saw a small effect of inaccurate dosing time data on RUV magnitude, and it was noted to monotonically increase with higher perturbations. In an agreement with our results, Alihozdic et al. [8] studied the effect on inaccurate infusion time data on meropenem and caspofungain and they conclude even 5 min inaccuracy can lead to biased and imprecise estimation of typical population value of V and corresponding BSV. While dosing time perturbations had a small effect on RUV, inaccurate sampling time data have a more noticeable effect. The resulting bias is more evident with intense designs compared to sparse design. The reason for this observation likely relates to the curvature of the structural PK profile. When the rate of change (dC/dt) is higher, the perturbations in time will cause greater deviations in concentrations compared to the later time points when dC/dt is usually smaller. Our results agreed with Karlsson et al., Choi et al. and Ludden et al. [9–11] and supports the consensus that the sensitivity of parameter estimates to the deviation of recorded time would be affected by the curvature of PK profile and the magnitude of perturbed recorded time. As a clinical consequences of inaccurate sampling time data, Santalo et al [3] studied the deviation in sampling time with vancomycin trough concentrations and report sub-therapeutic dosing as a consequence of early measurement. Similarly, Wang et al. [2] in a simulation study, reported inaccurately recorded tacrolimus trough concentrations would lead to inappropriate dose tailoring in up to 36% of cases.

Technical execution perturbations resulted in higher bias and imprecision in RUV magnitude compared to other sources. The magnitude of RUV increases with the misspecified structural PK model. Structural model misspecification did not have a noticeable effect on CL or BSV associated with CL, but that isn't the case with other parameters. Overall, misspecification that cannot be identified by residual plots can have a consequence on secondary PK parameters such as half-life due to a biased V. Study design also played a role here; with more data, model inadequacy is more readily apparent. With that said, times of sample collection are important in the estimation of some parameter values more than others. For example, with limited absorption data points, Ka estimates would have high imprecision and it is common to see this value fixed to a reasonable value.

Residual error misspecification contributed the largest inflation in RUV magnitude compared to other sources we studied. Choosing the wrong error model can mask the additivity from analytical assay and lead to biased estimates in the studied situations, including CL. The ignored error from the additive component will propagate to the parameters leading to the inaccuracy in parameter estimates. Silber et al. [12] found misspecification in the residual error model can inflate Type I error and induce bias in covariate inclusion.

Combined perturbations propagate to the magnitude of RUV and bias in PK parameter estimates. Recently, Al-Sallami et al. [13] published a commentary discussing the relationship of BSV from dose to response and the authors argued that variability in the system propagates linearly or nonlinearly based on the assumption of independency. This can be also observed in the current work. The propagation of all perturbations is not additive (the sum of variances squared is not equal to the RUV magnitude). For example, under the combined perturbations, the square root sum of variances is 17% (Table 2S) without taking into consideration the study sample collection design, which is larger than the median RUV shown in Table 2. The most likely explanation is due to the spread of variability between BSV and RUV.

It is noteworthy that median RUV value under the combined perturbations for all designs had smaller values compared to the median values captured in the published population pharmacokinetic analyses. Therefore, it might be that combined errors in this simulation study were underestimated compared to the literature. Despite the predominant source of variability in the combined sources might be model misspecification (residual and PK structural models) one should embrace this and carefully review model evaluation steps and follow good modeling practices.

In most of the studied scenarios, population typical values for primary PK parameters were robust for small perturbations in manufacturing, clinical trial execution, and technical errors, and resulted in relatively unbiased estimations. Random effects were more sensitive than fixed effects in the presence of perturbations, and that's due to the way the first-order approximation handles fixed-effect parameters (nonlinear approximation) and random-effect parameters (linear approximation) [11, 14]. Even in the absence of error in data, the first-order approximation might lead to bias in random-effects with sparse designs. Al-Banna et al. [15] evaluated the impact of sample collection design on parameter estimation using a one-compartment model. It was concluded that one early and one late time point would be sufficient to estimate primary fixed-effect parameters such as CL and V but resulted in imprecise random-effects estimation.

Limitations of this study include that we did not study the inadequacy of covariate models nor inaccurate covariate data. We also did not assess the unexplained variability in scenarios with multiple dosing or

misspecification in the covariate or BSV models. The extent to which the results in this manuscript would differ if other structural models or data (e.g., sampling time points) were used cannot be determined. They may give a general guidance but must not be over-interpreted in other contexts. Another limitation is that we did not consider other estimation algorithms such as SAEM, IMP, IMPMAP, or Bayesian. It is possible for example that SAEM would provide somewhat different results and that parts of the bias observed in the parameter estimation are in fact caused by the Taylor series expansion used in FOCE-I. We did not examine this issue and any practical relevance is unknown.

Finally, our misspecifications were chosen under an assumption of a well-controlled trial as might be seen in Phase 1 and Phase 2 clinical studies. Bias and imprecision findings are likely to be higher in studies that include outpatient dosing and multiple dosing studies that assume steady-state conditions as might be seen in Phase 3 clinical trials. It also needs to be considered that when parameter estimates used for simulating future concentration-time profiles and pharmacodynamic effects, imprecise and/or biased values could result in model-informed decisions that are imprecise and/or biased. The impact of these biases and imprecisions on decision making resulting from simulated exposure metrics such as area-under-the-curve or maximum concentrations is beyond the scope of this manuscript.

In conclusion, we assessed the contribution of hypothetical sources of variability in RUV that occur during NLMEM and offer these takeaway points:

- Pharmacostatistical model misspecifications were associated with relatively large increases in the magnitude of RUV compared to other sources for all levels of study design.
- The contribution of dose misspecification, and dosing time misspecifications have negligible effects on RUV but result in biased PK parameter estimates.
- Inaccurate sampling time data results in biased RUV magnitude which is sensitive to the magnitude of perturbations, and this effect was greater with more intense sampling designs.

Errors of the type studied here are real and do occur in PK studies. It is important to consider the question of what one wants to answer from the analysis. For example, if one is primarily interested in drug clearance, then some of these perturbations have less consequence. Nonetheless, as pharmacometricians, we don't have control over the perturbations and our challenge is to choose a pharmacostatistical model that adequately explains the data provided to us. This work provides information that can be used to understand and give insight for the interpretation of RUV magnitude. It may lead to recommendations from the pharmacometrics community that favor results from studies with lower RUV and minimize the "believability" of studies with large RUV.

Declarations

References

1. US Pharmacopeia (2011) <905> Uniformity of dosage units. https://www.usp.org/sites/default/files/usp/document/harmonization/gen-method/q0304_stage_6_monograph_25_feb_2011.pdf. Accessed Aug 2020
2. Wang J, Gao P, Zhang H et al (2020) Evaluation of concentration errors and inappropriate dose tailoring of tacrolimus caused by sampling-time deviations in pediatric patients with primary nephrotic syndrome. Ther Drug Monit 42:392–399. https://doi.org/10.1097/FTD.0000000000000717
3. Santalo O, Baig U, Poulakos M, Brown D (2016) Early vancomycin concentrations and the applications of a pharmacokinetic extrapolation method to recognize sub-therapeutic outcomes. Pharmacy 4:37. https://doi.org/10.3390/pharmacy4040037
4. Proost JH (2019) Calculation of the coefficient of variation of log-normally distributed parameter values. Clin Pharmokinet 58:1101–1102
5. Baron KT (2022) mrgsolve: simulate from ODE-based models
6. Lavielle M, Aarons L (2016) What do we mean by identifiability in mixed effects models? J Pharmacokinet Pharmacodyn 43:111–122. https://doi.org/10.1007/s10928-015-9459-4
7. Irby DJ, Ibrahim ME, Dauki AM et al (2021) Approaches to handling missing or "problematic" pharmacology data: pharmacokinetics. CPT Pharmacometr Syst Pharmacol 10:291–308. https://doi.org/10.1002/psp4.12611
8. Alihodzic D, Broeker A, Baehr M et al (2020) Impact of inaccurate documentation of sampling and infusion time in model-informed precision dosing. Front Pharmacol 11:172. https://doi.org/10.3389/fphar.2020.00172
9. Karlsson MO, Jonsson EN, Wiltse CG, Wade JR (1998) Assumption testing in population pharmacokinetic models: illustrated with an analysis of moxonidine data from congestive heart failure patients. J Pharmacokinet Biopharm 26:207–246

10. Choi L, Crainiceanu CM, Caffo BS (2013) Practical recommendations for population PK studies with sampling time errors. Eur J Clin Pharmacol 69:2055–2064. https://doi.org/10.1007/s00228-013-1576-7
11. Sun H, Ette EI, Ludden TM (1996) On the recording of sample times and parameter estimation from repeated measures pharmacokinetic data. J Pharmacokinet Biopharm 24:637–650
12. Silber HE, Kjellsson MC, Karlsson MO (2009) The impact of misspecification of residual error or correlation structure on the type I error rate for covariate inclusion. J Pharmacokinet Pharmacodyn 36:81–99
13. Al-Sallami HS, Wright DF, Duffull SB (2022) The propagation of between-subject variability from dose to response. Br J Clin Pharmacol 88:1414–1417
14. Ette EI, Kelman AW, Howie CA, Whiting B (1993) Interpretation of simulation studies for efficient estimation of population pharmacokinetic parameters. Ann Pharmacother 27(9):1034–1039
15. Al-Banna MK, Kelman AW, Whiting B (1990) Experimental design and efficient parameter estimation in population pharmacokinetics. J Pharmacokinet Biopharm 18:347–360