



**HAL**  
open science

## **Better Dosing Through Better Error: Residual Error as a Hidden Lever in Model-Informed Precision Dosing**

Racym Berrah, Iris Minichmayr, Jean-Baptiste Woillard

► **To cite this version:**

Racym Berrah, Iris Minichmayr, Jean-Baptiste Woillard. Better Dosing Through Better Error: Residual Error as a Hidden Lever in Model-Informed Precision Dosing. *Therapeutic Drug Monitoring*, 2025, Online ahead of print. <10.1097/FTD.0000000000001413>. <hal-05414757>

**HAL Id: hal-05414757**

**<https://unilim.hal.science/hal-05414757v1>**

Submitted on 13 Dec 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

1 **Better Dosing Through Better Error: Residual Error as a**  
2 **Hidden Lever in Model-Informed Precision Dosing**

3 Racym Berrah PharmD<sup>1</sup>, Iris Minichmayr PharmD, PhD<sup>2</sup>, Jean Baptiste Woillard  
4 PharmD, PhD<sup>1,3</sup>. On behalf of the IATDMCT Pharmacometrics group

5 1. Pharmacology & Transplantation, INSERM U1248, Limoges University, Limoges,  
6 France

7 2. Department of Clinical Pharmacology, Medical university of Vienna, Vienna,;  
8 Austria

9 3. Department of Pharmacology, Toxicology and Pharmacovigilance, CHU Limoges,  
10 Limoges, France

11

12 Corresponding author:

13 Jean-Baptiste Woillard

14 Title: PharmD, PhD

15 ORCID: 0000-0003-1695-0695

16 Mail: jean-baptiste.woillard@unilim.fr

17 Address:

18 Univ. Limoges, Inserm U1248 P&T,

19 2 rue du Pr Descottes,

20 F-87000 Limoges, France.

21 Phone: +33 5 55 05 61 40

22 Fax: +33 5 55 05 61 62

23

24 Conflict of interest and source of fundings: This work is part of the DIGPHAT project  
25 which was supported by a grant to Jean-Baptiste WOILLARD (# ANR-22-PESN-  
26 0017) from the National Research Agency (ANR), France, under the France 2030  
27 program. For the remaining authors none were declared.

28

29

30 **Abstract**

31 Background: Model-Informed Precision Dosing is increasingly integrated into  
32 Therapeutic Drug Monitoring to optimize individualized pharmacotherapy, with  
33 Maximum A Posteriori Bayesian Estimation (MAPBE) serving as a key tool for  
34 personalized pharmacokinetic parameter estimation. However, the influence of  
35 residual error specification on MAPBE imprecision remains underexplored. We  
36 hypothesized that reducing residual error in MAPBE reduces the imprecision of  
37 individual area under the curve (AUC) prediction.

38 Methods: Using rich pharmacokinetics datasets comprising 321 profiles for  
39 tacrolimus, iohexol, and mycophenolate mofetil, we applied MAPBE based on sparse  
40 sampling (three time points) using population pharmacokinetics models implemented  
41 from literature while varying proportional residual error settings, including only a  
42 near-zero proportional error ( $10^{-8}$  %), a low proportional error (1%, “Flat1”), and the  
43 original published error model.

44 Results: Results showed that employing a low proportional error reduced the overall  
45 root mean square error (%RMSE) of individual AUC predictions by 30–40% relative to  
46 the original residual error configuration. For instance, for tacrolimus, %RMSE  
47 decreased from 28.5% under the published error model to 16.3% with the 1% error  
48 setting, whereas for iohexol, a near-zero residual error achieved up to a 40%  
49 %RMSE relative reduction. At the individual level, the near-zero error scenario  
50 provided the most accurate AUC estimates in 45–62% of patients across the  
51 evaluated drug models.

52 Conclusions: These findings indicate that a strategic reduction of residual errors in  
53 MAPBE significantly enhances the influence of observed data on posterior  
54 calculations, thereby improving AUC estimation precision and consequently dosing  
55 accuracy without necessitating new data acquisition or model redevelopment.

56 **Keywords:** model-informed precision dosing; residual error; population  
57 pharmacokinetics; tacrolimus; MMF

58

## 60 1. Introduction

61 The integration of model-informed precision dosing (MIPD) into routine therapeutic  
62 drug monitoring (TDM) has opened new avenues for individualized  
63 pharmacotherapy. Central to this approach is maximum a posteriori Bayesian  
64 estimation (MAPBE), which enables the derivation of patient-specific  
65 pharmacokinetic (PK) parameters by combining sparse clinical data with prior  
66 information from validated population PK models.

67 Despite growing interest, the implementation of Model- Informed Precision Dosing  
68 (MIPD) faces challenges including inconsistent reporting of model assumptions and  
69 residual error structures <sup>1</sup>. One such critical yet underexplored element is the  
70 handling of residual error. Typically modelled as a combination of proportional and  
71 additive components, residual error accounts for random unexplained variability in  
72 measured concentrations. However, it also directly influences the balance between  
73 prior and observed data in MAPBE and might thus affect the quality of individual  
74 predictions. When residual error is high, observed concentrations are down-weighted  
75 in the individual predictions in favor of the prior (i.e., the underlying population  
76 model), potentially increasing the imprecision of individualised exposure estimates  
77 such as the area under the concentration-time curve (AUC). Conversely, minimizing  
78 residual error increases the influence of observed data, but may bear the risk of  
79 overfitting if noise is misinterpreted as true signal.

80 In current practice, residual error parameters are often retained unchanged from the  
81 original population model during clinical implementation, even when applied to  
82 different populations or sampling conditions. Informal discussions among  
83 pharmacologists reveal heterogeneity in these practices, with some teams modifying  
84 error parameters empirically and others using them as published—despite the lack of  
85 systematic evaluation of these choices.

86 Another promising strategy to enhance individualization within MAPBE involves  
87 reducing the weight of the prior distribution, as proposed by Hughes et al <sup>2</sup>. By  
88 inflating the variance of interindividual variability terms ( $\omega^2$ ), the influence of the prior  
89 is down-weighted, allowing the model to rely more heavily on observed data. This  
90 approach, known as “flattening the prior,” was shown to improve prediction accuracy  
91 in patients poorly represented by the population model, especially when  
92 implemented dynamically via machine learning.

93 In MAP-BE, residual error represents the short-term noise around each measured  
94 concentration, i.e. the deviation between the measured concentration and individual  
95 model prediction, potentially arising from assay imprecision, sampling/handling  
96 variability, documentation errors, or small model–data mismatches. Operationally,  $\sigma$   
97 is the standard deviation of the measurement noise (RUV for residual unexplained

98 variability) around each observed concentration. The MAP-BE objective combines  
99 two forces: (i) a data term that penalizes squared prediction errors (scaled by  $\sigma^2$ ),  
100 and (ii) a prior term that penalizes individual deviations  $\eta$  from the typical profile  
101 according to the population covariance  $\Omega$ . Thus, a smaller  $\sigma$  gives more weight to the  
102 observations (the model is pulled closer to the measured points), whereas a larger  $\sigma$   
103 gives more weight to the prior (individual parameters stay closer to the population).  
104 In this work we vary only  $\sigma$  in the RUV; the structural model and interindividual  
105 variability ( $\eta, \Omega$ ) are kept exactly as published. (While one could alternatively inflate  $\Omega$   
106 to “flatten” the prior<sup>2</sup> and move closer to the data, we did not do so here in order to  
107 isolate the effect of  $\sigma$ .)

108 In this study, we investigate how varying the magnitude of residual error influences  
109 the accuracy of AUC estimation when literature-derived pharmacokinetic models are  
110 implemented using MAPBE.

111 To this end, we used three representative case studies, tacrolimus, iohexol, and  
112 MMF, for which complete pharmacokinetic profiles were available. These rich  
113 datasets allow direct comparison between AUC estimates obtained via MAPBE with  
114 sparse sampling and reference AUCs based on rich data calculated using the  
115 trapezoidal rule. These three drugs, which differ in population pharmacokinetic  
116 models serve as representative test cases for generalizing methodological insights.

## 117 **2. Materials and Methods**

### 118 **2.1 Drugs and Data Sources**

119 This study focused on 2 drugs commonly used in transplant pharmacotherapy,  
120 tacrolimus, and MMF and one used to estimate the glomerular filtration rate  
121 (iohexol), selected for their diverse pharmacokinetic profiles and the availability of  
122 published population PK models suitable for MAPBE.

123 For tacrolimus, we used data from the PCCP study<sup>3</sup>, which originally informed the  
124 development of the Woillard et al. model. To assess the impact of dataset  
125 provenance on MAPBE performance, we also analysed an external validation cohort  
126 (AADAPT<sup>4</sup>). Both once-daily (QD) and twice-daily (BID) formulations were included  
127 when available. The Woillard model, which accommodates both formulations, was  
128 applied to the AADAPT and PCCP cohorts, whereas the Campagne model  
129 (developed on BID formulation) was used only with PCCP. For iohexol and MMF,  
130 data were drawn from independent, previously published clinical studies<sup>5-7</sup>. All MMF  
131 patients were co-treated with tacrolimus; none received cyclosporine or everolimus.  
132 The variables available in each dataset are summarised in Table 1.

133 To ensure robust reference AUC values, only profiles with at least eight observed  
134 concentration measurements per patient were included in the analysis. These rich

135 datasets provided the necessary granularity to evaluate AUC estimates derived from  
136 sparse sampling under different proportional residual error conditions.

137 An additional dataset of 500 patients based on the Woillard tacrolimus population  
138 model was simulated to also investigate the effect of decreasing the residual error in  
139 simulated patients from a model used to perform MAPBE. In simulations, data were  
140 generated using the Woillard tacrolimus model including its original residual error  
141 (proportional 11.3% and additive 0.71 µg/L). To eliminate aberrant cases, we first  
142 removed the lowest and highest 5% of trough concentrations (5<sup>th</sup>–95<sup>th</sup> percentile  
143 trimming). This yielded a final subset of 450 patients with trough concentrations  
144 ranking between 2 and 22 µg/L.

## 145 2.2 Sampling Design for MAPBE

146 MAPBE was applied using only three concentration time points per profile to emulate  
147 sparse clinical sampling, as typically encountered in routine MIPD. For tacrolimus,  
148 samples were taken at 0, 1, and 3 hours post-dose; for iohexol, at approximately 5  
149 minutes, 30 minutes, and 3 hours; and for MMF, at 20 minutes, 1 hour, and 3 hours.  
150 These time points were chosen to reflect common clinical sampling schedules. We  
151 applied a ±10% tolerance to each nominal time, but bounded the absolute window to  
152 avoid extremes: a minimum width of ±5 minutes (to prevent overly tight windows at  
153 early times) and a maximum of ±30 minutes (to avoid excessively wide windows at  
154 late times). Thus, for the 20-min sample, observations from 15-25 min (rather than  
155 18–22 min) were accepted, for the 1-h sample 54-66 min, and the 9-h sample was  
156 capped at 8 h 30 min–9 h 30 min (instead of 8 h 06 min –9 h 54 min).

## 157 2.3 Residual error scenario for MAPBE

158 Individual pharmacokinetic parameters were estimated using MAPBE. In this  
159 approach, the goal is to minimize the negative log-likelihood of the posterior  
160 distribution, which combines the deviation between observed and predicted  
161 concentrations and the deviation of individual parameters from population means:

$$Criteria = -2\log\left(\sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{\sigma_i^2} + \sum_{j=1}^m \frac{\eta_j}{\omega_j^2}\right)$$

162

163  $y_i$  represents the observed concentration,  $\hat{y}_i$  is the predicted concentration based on  
164 the pharmacokinetic model and estimated parameters,  $\sigma^2$  is the variance of the  
165 residual error for observation  $i$ ,  $\eta_j$  is the individual deviation of the  $j$ th parameter from  
166 the population mean, and  $\omega^2$  is the variance reflecting the interindividual variability  
167 for the  $j$ th pharmacokinetic parameter.

168 Reducing  $\sigma^2$  increases the weight of the observed data in the estimation process,  
169 thereby forcing the posterior prediction to more closely follow the individual  
170 measurements. In practice, a smaller residual error term magnifies the penalty  
171 associated with deviations between observed and predicted concentrations,  
172 effectively compelling the estimation algorithm to prioritize data fitting over prior  
173 conformity.

174 MAPBE was conducted under a range of predefined residual error conditions to  
175 evaluate their effect on AUC estimation accuracy. These conditions included a nearly  
176 negligible proportional error level ( $10^{-8}$  %, referred to as "Null") to approximate  
177 maximum weighting of the observed concentrations, as well as a low proportional  
178 error of 1% ("Flat1") and a moderate proportional error of 5% ("Flat5"), reflecting  
179 increasing degrees of measurement noise. In addition, we tested scaled versions of  
180 the original published error models, with residual error components reduced to 80%,  
181 60%, 40%, and 20% of their original values. The final configuration employed the full  
182 original error model as reported in the respective population PK publications and  
183 served as a benchmark. Details of all error settings are provided in Supplementary  
184 Table 1.

## 185 2.4 Pharmacokinetic Models

186 For each drug, MAPBE was based on previously published population PK models  
187 selected for their differences in terms of structure (Supplementary Table 2).

188 Tacrolimus exposure was estimated using the model developed by Woillard et al.<sup>3</sup>,  
189 originally constructed from the AADAPT cohort and widely applied in routine MIPD.  
190 To assess the robustness of MAPBE across model structures and datasets, an  
191 additional tacrolimus model by Campagne et al.<sup>8</sup>, incorporating CYP3A5 genotype  
192 effects, was also evaluated on the external PCCP dataset.

193 For iohexol, the model by Zwart et al.<sup>9</sup> was implemented, which was developed for  
194 GFR estimation in renal transplant patients using a population-based clearance  
195 approach.

196 MMF refers to the administered prodrug, whereas the pharmacokinetic model and  
197 exposure metrics here pertain to mycophenolic acid (MPA), the active metabolite of  
198 MMF. MPA exposure was estimated using the model published by Rong et al.<sup>10</sup>,  
199 which characterizes MPA pharmacokinetics in corticosteroid-free adult kidney  
200 transplant recipients co-treated with tacrolimus.

201 All models were encoded in *mrgsolve*<sup>11</sup> and implemented without structural  
202 modification. Parameter values, residual error structures, and covariate relationships  
203 were applied as reported in the original publications. This approach reflects real-  
204 world use of literature-derived models in MIPD applications, where re-estimation or  
205 reparameterization is typically not feasible.

## 206 2.5 Bayesian Estimation Process

207 MAPBE was performed in R using the *mapbayR*<sup>12</sup> package. For each patient and  
208 each residual error condition, individual pharmacokinetic parameters were computed  
209 by fitting the corresponding population model to the three selected sampling time  
210 points. To evaluate the influence of residual error settings, simulated concentration-  
211 time profiles were generated from the individual estimates obtained under each  
212 scenario and compared to the full reference profiles for AUC imprecision  
213 assessment.

## 214 2.6 Evaluation Criteria

215 Reference AUCs were calculated based on the trapezoidal rule, applied to the full  
216 set of observed concentration–time points for each patient. As a sensitivity analysis  
217 for the tacrolimus case study, we also computed a model-predicted reference AUC  
218 by integrating the model-predicted concentration–time profile conditional on all  
219 observations from that profile.

220 The primary endpoint was the percentage root mean square error (%RMSE) across  
221 the study population, defined as the mean percentage difference between the AUC  
222 estimated by each MAPBE scenario and the corresponding reference AUC allowing  
223 to facilitate comparisons across drugs with different AUC scales and reflects method  
224 performance at the population level.

225 As a secondary endpoint, we assessed method performance at the individual level.  
226 For each patient, we identified the MAPBE scenario that produced the AUC estimate  
227 with the lowest root mean square error (%RMSE) relative to the reference AUC. The  
228 distribution of patients according to their best-performing method was then analyzed  
229 to capture inter-individual variability in the accuracy of AUC estimation across  
230 residual error settings. In addition to %RMSE, we report P15 (percentages of patients  
231 with absolute percentage AUC error  $\leq 15\%$ ) as clinically interpretable accuracy  
232 thresholds.

## 233 3. Results

### 234 3.1 Dataset Composition

235 A total of 321 real patients after filtering out patients with less than 8 samples and  
236 450 simulated pharmacokinetic profiles were analyzed. Most of the profiles  
237 comprised a minimum of 10 concentration points, thereby ensuring the accurate  
238 determination of AUCs as shown in Supplemental Table 3.

### 239 3.2 %RMSE and Model imprecision

240 Across the entire range of drugs, the %RMSE exhibited a consistent decrease as the  
241 residual error was reduced (Figure 1a). The smallest %RMSE values for AUC  
242 estimation were obtained under the 1% proportional error condition. The  
243 intermediate error settings (1–5%) were shown to enhance performance in  
244 comparison with the original model error. It is important to note that these  
245 enhancements were successfully replicated in external test sets. Only for iohexol,  
246 the smallest %RMSE was for a negligible error.

247 As for the model-predicted reference AUC, the conclusions regarding the effect of  $\sigma$   
248 were similar to those obtained with the trapezoidal reference; the comparative plot is  
249 provided as Supplementary Figure 1. Furthermore, the concentration-time profiles of  
250 tacrolimus exhibited similar trends, regardless of which model was applied (PCCP  
251 using the Woillard et al. model vs. the Campagne et al. model) as seen in more  
252 detailed in Figure 1b.

253 As for the training and external datasets, the increase in %RMSE in the null scenario  
254 is due to three cases of data overfitting. However, the decrease in %RMSE as the  
255 proportional error decreases is still observable. P15 summaries mirrored the  
256 %RMSE trends: Flat1 and Null were most often within 15% at the individual level.  
257 Figure 2 displays the P15– $\sigma$  curves for each dataset.

### 258 3.3 Posterior Profile Fidelity

259 The application of MAPBE with minimal residual error resulted in posterior profiles  
260 that closely followed the individual's observed data (Figure 3a). In contrast, models  
261 incorporating higher residual error yielded smoother posterior profiles that deviated  
262 from the observed data and tended to regress toward the population prior. A small  
263 subset of patients –3 out of 321– with marked variability in their measured  
264 concentrations showed overfitting when the residual error was set to near-zero.  
265 However, such cases were infrequent, as illustrated in Figure 3b and overfitting was  
266 easy to recognize on inspection, with posterior curves tightly tracking assay noise  
267 and yielding implausible parameter or AUC values (e.g., abnormally high tacrolimus  
268 AUC). The magnitude of residual unexplained variability (RUV) affected individual  
269 trends: lower RUV allowed the model to closely follow observed data, whereas  
270 higher RUV resulted in predictions that leaned more heavily on typical population  
271 values, potentially masking true interindividual variability.

272

### 273 3.4 Inter-individual variability in method performance (Individual level)

#### 274 Inter-Individual Variability in Optimal MAPBE Performance

275 While population-level performance metrics, such as average %RMSE, generally  
276 favored the Flat1 scenario (1% proportional error), a deeper analysis at the individual  
277 patient level revealed higher variability in optimal method performance. To  
278 investigate this inter-individual variability, we determined, for each patient, the  
279 MAPBE scenario that yielded the most accurate AUC estimate—defined as the  
280 lowest absolute percentage error relative to the reference AUC.

281 Figure 4 and Supplementary Table 4 presents the distribution of patients according  
282 to the best-performing residual error scenario. Although Flat1 and the null error  
283 setting frequently emerged as top performers, the null scenario—characterized by  
284 the absence of residual error and thus maximal weighting of observed data—proved  
285 to be the most accurate method for most patients across all drug models. In total, it  
286 outperformed all other scenarios in over 300 individual cases, substantially  
287 exceeding the count observed for Flat1. These findings indicate that in many  
288 instances, especially when measurement variability is low, removing residual error  
289 altogether may yield more accurate AUC estimates than individualized MAPBE with  
290 standard residual error assumptions.

#### 291 A Head-to-Head Comparison

292 To further compare the two most promising approaches, we performed a direct  
293 comparison between Flat1 and null residual error configurations (Figure 5). While  
294 Flat1 showed slightly better or equivalent average performance across all patients,  
295 this advantage was not consistently observed at the individual level. Notably, Flat1  
296 only outperformed the null scenario in the MMF dataset, which included just 16  
297 patients. In all other drug models, both strategies yielded comparable or superior  
298 accuracy in favor of the null scenario.

299 Interestingly, the pattern of optimal residual error scenario was consistent across  
300 different pharmacokinetic models for tacrolimus. In the PCCP dataset, 79 out of 136  
301 patients had the same best-performing scenario regardless of the model used: 34  
302 patients consistently benefited from the null error approach, while 45 had better  
303 estimates under Flat1. This suggests that, in many cases, data quality and error  
304 specification may have a greater impact on individual predictions than the choice of  
305 structural PK model.

### 306 3.5 Simulated data

307 The preceding results assume that full-profile AUCs derived via trapezoidal rule  
308 represent an adequate reference standard. While this approach is widely accepted, it  
309 remains sensitive to assay noise and sampling variability—factors that can influence

310 %RMSE calculations and benchmark comparisons. Simulated data are here to  
311 confirm with exact AUC the preceding results shown in Figure 1b.

#### 312 **4. Discussion**

313 This study aimed to evaluate the impact of proportional residual error magnitude on  
314 AUC determination when applying literature-derived population PK models and LSS  
315 using MAPBE. By leveraging full pharmacokinetic profiles for tacrolimus, iohexol,  
316 MMF, we compared MAPBE-derived AUCs from sparse sampling to reference AUCs  
317 calculated using the trapezoidal rule.

318 We selected the rich-profile trapezoidal AUC as the primary reference because it is  
319 model-independent and remains the most accepted empirical standard in clinical PK  
320 when dense sampling is available.

321 Importantly, because our MAP-BE predictions for the LSS scenarios are based on  
322 only three samples, decreasing  $\sigma$  does not guarantee a closer match to the rich  
323 trapezoidal AUC, which integrates the entire profile. In fact, in a small number of  
324 extreme cases, visual inspection shows that the model-predicted AUC (even when  
325 conditioned on all points) can be less plausible owing to shape distortion  
326 (Supplemental Figure 2), whereas the trapezoidal AUC remains consistent with the  
327 observed data. These findings support retaining the trapezoidal AUC as the primary  
328 reference, with the model-predicted reference provided as a corroborative analysis  
329 (Supplemental Figure 1).

330

331 In this study, we specifically focused on the proportional component of the residual  
332 error structure. The additive error was not explored, primarily because it is expressed  
333 in concentration units, which vary across models depending on the drug and assay  
334 used. This heterogeneity limits the comparability of additive error magnitudes  
335 between models. Moreover, proportional error typically has a greater influence on  
336 individual prediction accuracy in MAP-BE, especially for drugs with wide  
337 concentration ranges. Focusing on proportional error thus allows for a more  
338 interpretable and clinically relevant assessment of its impact on individualised  
339 predictions, while minimizing confounding from unit-dependent additive terms.

340

341 Our findings show that reducing residual error systematically improves AUC  
342 prediction accuracy across a range of drugs with differing model structures. These  
343 improvements were observed not only in training datasets but also in independent  
344 test cohorts, suggesting that residual error minimization might be a generalisable  
345 and reproducible strategy for enhancing MAPBE performance, provided that the  
346 underlying data are of sufficient quality.

347 While minimizing residual error can improve the fidelity of posterior predictions,  
348 caution is needed when dealing with noisy or sparse datasets, as illustrated in Figure  
349 2. In such contexts, assuming a null residual error can lead to overfitting, where the  
350 model interprets all observed fluctuations as meaningful biological signal rather than  
351 measurement noise. For example, in Figure 3b, setting RUV to zero resulted in an

352 implausibly high AUC for tacrolimus exceeding 1000  $\mu\text{g}\cdot\text{h}/\text{L}$ , highlighting the risk of  
353 unstable and clinically unrealistic parameter estimates in certain individuals; although  
354 such cases can often be identified by pharmacometricians by simply looking at the  
355 fitted curve.

356

357 The relative performance of MAPBE across the three drug models was also  
358 influenced by the alignment between the sampling times used in this study and those  
359 employed during model development. For tacrolimus, the Woillard et al. model was  
360 built using rich sampling that included 0, 1, and 3 hours post-dose, the same points  
361 used in our sparse MAPBE design. This likely contributed to its high accuracy.

362 In comparison, the iohexol model developed by Zwart et al. was primarily based on  
363 relatively early sampling times (up to 6 hours post-dose) without including late time  
364 points. In contrast, Destere's model incorporated later sampling points (e.g., nine  
365 hours), resulting in improved predictive performance. When the MAPBE method is  
366 applied to Zwart's model using later sampling points, a notably higher %RMSE is  
367 observed compared to Destere's results, highlighting the importance of including late  
368 sampling times for iohexol.

369

370 Interestingly, the baseline %RMSE of the original model appeared to predict the  
371 degree of improvement achievable through residual error reduction. Models with  
372 higher initial %RMSE showed greater benefit when residual error was minimized,  
373 likely because down-weighting the prior and increasing the influence of observed  
374 data helped correct for model imprecision. Conversely, models with already low  
375 %RMSE required more aggressive down-weighting to see further gains, suggesting  
376 that residual error tuning may be particularly beneficial when adapting older or less  
377 precise models for clinical implementation. For instance, if a model was originally  
378 developed in specific subgroups such as ICU or obese patients, and is later applied  
379 to a more general patient population, increasing the influence of the observed data  
380 (by reducing the residual error) can improve individual predictions without requiring  
381 model redevelopment.

382

383 Although aggregate metrics favored the Flat1 scenario, our patient-level analysis  
384 revealed substantial variability in which null residual error setting yielded the most  
385 accurate AUC estimate. Both the Flat1 and null scenarios consistently outperformed  
386 other methods across individuals, but they did so in different ways. The null scenario  
387 was particularly effective at capturing extreme profiles, likely due to its full reliance  
388 on observed data. However, this also rendered it more vulnerable to overfitting in the  
389 presence of noisy or aberrant measurements. In contrast, the Flat1 scenario  
390 maintained a mild constraint through its low residual error assumption, offering a  
391 practical balance between individualization and robustness. While it may not capture  
392 all outliers, it avoids the risks associated with overinterpretation of noise. Therefore,  
393 both methods are viable, but clinical judgment should guide their use, particularly in

394 cases where model output conflicts with expected pharmacological behavior or  
395 clinical context.

396

397 Our findings reinforce the notion that residual error should not be treated as a fixed,  
398 immutable element of population PK models when applied to new clinical contexts.  
399 Instead, it can and should be pragmatically adjusted during MAPBE implementation,  
400 particularly when the original model was developed under different assay conditions,  
401 sampling designs, or population characteristics. A low fixed proportional error (e.g.,  
402 1%) may serve as a suitable default, striking a balance between enabling  
403 individualization and avoiding overfitting. This approach offers a practical alternative  
404 to blindly adopting the original model's residual error, which may no longer reflect  
405 current analytical precision or clinical sampling practices.

406

407 Importantly, residual error adjustment does not require any modification to the  
408 structural model or covariate relationships and can often be implemented directly  
409 within MAPBE software or through wrapper code. To support transparency and  
410 reproducibility, we recommend that pharmacometricians report the rationale and  
411 parameters used for residual error modification when applying literature-based  
412 models in MIPD tools. Incorporating this flexibility into clinical decision-support  
413 platforms would facilitate broader uptake of model-informed dosing while maintaining  
414 methodological rigor.

415

416 Several limitations should be considered when interpreting the findings of this study.  
417 First, although our analysis was based on rich pharmacokinetic datasets and multiple  
418 published models, only three drugs were evaluated. While tacrolimus, iohexol, and  
419 MMF differ in their pharmacokinetic characteristics and thus provide a diverse set of  
420 test cases, broader validation across other drug classes and clinical settings is  
421 warranted to fully assess generalizability. Nonetheless, the consistency of results  
422 across models with differing structural complexity and error behavior suggests that  
423 the observed impact of residual error adjustment may extend beyond these specific  
424 examples. This may be especially relevant for drugs exhibiting nonlinear  
425 pharmacokinetics, where small deviations in concentration may disproportionately  
426 affect model predictions and AUC estimates, making residual error specification  
427 even more critical.

428

429 Second, our implementation focused solely on modifying the residual error  
430 component without adjusting other model elements, such as interindividual variability  
431 (flatten of the priors).

432

433 Third, available MMF profiles ended at 9 h, which may not fully capture late  
434 enterohepatic recirculation (typically ~6–12 h)<sup>13,14</sup> and thus may slightly  
435 underestimate the reference AUC. This limitation is common in published MMF PK  
436 studies because late samples are rarely collected densely between 6 and 12 h.

437 While model-based simulations could be produced, they would not replace empirical  
438 information in this time window and would remain dependent on model assumptions.  
439 Consequently, our MMF findings should be interpreted with caution, and future  
440 prospective work including full 6-12-h profiles is warranted.

441 Fourth, clinical-trial datasets are usually more standardized than routine care  
442 datasets (sampling windows, adherence, handling), so models trained on trial-like  
443 data may more likely encounter a mismatch between development and real-world  
444 conditions.

445  
446 Finally, this study did not evaluate the downstream clinical consequences of  
447 improved AUC estimation, such as target attainment or clinical outcomes, limiting our  
448 ability to directly infer therapeutic benefit. Future prospective studies linking residual  
449 error optimization to clinical endpoints will be important for confirming its relevance in  
450 routine care.

451 In MAP-BE,  $\sigma$  primarily weights the observations relative to the prior and reflects the  
452 total observation-level variability i.e. assay imprecision plus pre-analytical factors  
453 (e.g., timing/documentation, handling) and minor model data mismatches rather than  
454 the assay CV alone.

455 Accordingly, a very small but non-zero  $\sigma$  is most appropriate when overall  
456 observation noise is low (high-precision analytics, such as LC-MS/MS, and tightly  
457 documented sampling/handling), because it prevents the prior from dominating  
458 sparse LSS data and lets local measurements drive individualization while the  
459 literature model provides the structural “shape.” By contrast, in noisier settings an  
460 overly small  $\sigma$  will overweight measurement noise and yield unstable posteriors.  
461 However, even with a precise LC-MS/MS platform, poor TDM practices argues for a  
462 larger  $\sigma$ .

463 Collectively, our results offer practical guidance for centers deploying published  
464 models: when local assay performance, TDM practices, or patient differ from that of  
465 the source model, use a small but non-zero proportional  $\sigma \sim 1\%$  (Flat1) only in high-  
466 quality contexts. In other case, increase  $\sigma$  (e.g., 2–5%) or select it via a brief  
467 sensitivity check ( $\approx 0.5\text{--}3\%$ ) and confirm with routine diagnostics (visual fits, residual)  
468 before dosing decisions. Finally, when timing records are imprecise or sample  
469 handling is variable (e.g., rounded draw times or small handling deviations),  
470 robustness can also be improved by adding a small additive noise term to the RUV  
471 model.

## 472 **5. Conclusion**

473 This study highlights the critical influence of residual error specification on the  
474 accuracy of AUC estimation when applying literature-derived pharmacokinetic  
475 models in MIPD. Our results demonstrate that reducing residual error enhances the  
476 contribution of observed data during MAPBE, often leading to more accurate  
477 individual exposure estimates, particularly when the original model's performance is  
478 suboptimal or sampling design is sparse. However, minimizing residual error must be

479 approached with caution, as it can increase the risk of overfitting in noisy or  
480 unstructured data.

481 From a practical standpoint, residual error represents a flexible implementation  
482 parameter that can be adjusted without modifying the structural model. Our findings  
483 support the use of a low, fixed residual error as a pragmatic and reproducible  
484 strategy when original model assumptions are misaligned with the clinical context.  
485 Moreover, we advocate for greater transparency in the reporting of residual error  
486 adjustments and sampling designs in MIPD studies and tools. These methodological  
487 refinements should become part of standardized guidelines for model reuse in  
488 clinical pharmacology. In parallel, we encourage pharmacometrician to clearly report  
489 their residual error structures, including whether terms are expressed as variances or  
490 standard deviations, to support consistent and correct reuse in external applications.

491  
492 As MIPD continues to expand in scope and clinical impact, establishing clear,  
493 evidence-based practices for model adaptation will be essential. Residual error  
494 tuning, alongside complementary strategies such as prior flattening<sup>2</sup>, represents a  
495 key opportunity to enhance the precision and reliability of model-based therapeutic  
496 decisions—potentially without requiring model redevelopment.

497  
498 Acknowledgment: This work originated from discussions within the PMx committee.  
499 We would like to especially thank M. Hovd and Anders Åsberg, as well as Michael  
500 Neely, for their valuable advices, and T. Mizuno and J. Hughes for the initial  
501 discussions that helped shape the direction of this work.

502 References

- 503 1. Minichmayr IK, Dreesen E, Centanni M, et al. Model-informed precision dosing: State of  
504 the art and future perspectives. *Adv Drug Deliv Rev.* 2024;215:115421.  
505 doi:10.1016/j.addr.2024.115421
- 506 2. Hughes JH, Keizer RJ. A hybrid machine learning/pharmacokinetic approach  
507 outperforms maximum a posteriori Bayesian estimation by selectively flattening model  
508 priors. *CPT Pharmacom & Syst Pharma.* 2021;10(10):1150-1160.  
509 doi:10.1002/psp4.12684
- 510 3. Woillard JB, de Winter BCM, Kamar N, Marquet P, Rostaing L, Rousseau A. Population  
511 pharmacokinetic model and Bayesian estimator for two tacrolimus formulations--twice  
512 daily Prograf and once daily Advagraf. *Br J Clin Pharmacol.* 2011;71(3):391-402.  
513 doi:10.1111/j.1365-2125.2010.03837.x
- 514 4. Benkali K, Rostaing L, Premaud A, et al. Population Pharmacokinetics and Bayesian  
515 Estimation of Tacrolimus Exposure in Renal Transplant Recipients on a New Once-  
516 Daily Formulation: *Clinical Pharmacokinetics.* 2010;49(10):683-692.  
517 doi:10.2165/11535950-000000000-00000
- 518 5. Djebli N, Picard N, Rérolle JP, Meur YL, Marquet P. Influence of the UGT2B7 promoter  
519 region and exon 2 polymorphisms and comedications on Acyl-MPAG production in vitro  
520 and in adult renal transplant patients. *Pharmacogenetics and Genomics.*  
521 2007;17(5):321-330. doi:10.1097/FPC.0b013e32801430f8
- 522 6. Åsberg A, Bjerre A, Almaas R, et al. Measured GFR by Utilizing Population  
523 Pharmacokinetic Methods to Determine Iohexol Clearance. *Kidney International*  
524 *Reports.* 2020;5(2):189-198. doi:10.1016/j.ekir.2019.11.012
- 525 7. Carrier P, Destere A, Giguet B, et al. Iohexol plasma and urinary concentrations in  
526 cirrhotic patients: A pilot study. *WJH.* 2022;14(8):1621-1632.  
527 doi:10.4254/wjh.v14.i8.1621
- 528 8. Campagne O, Mager DE, Brazeau D, Venuto RC, Tornatore KM. Tacrolimus  
529 Population Pharmacokinetics and Multiple CYP3A5 Genotypes in Black and White  
530 Renal Transplant Recipients. *The Journal of Clinical Pharma.* 2018;58(9):1184-1195.  
531 doi:10.1002/jcph.1118
- 532 9. Zwart TC, De Vries APJ, Engbers AGJ, et al. Model-Based Estimation of Iohexol  
533 Plasma Clearance for Pragmatic Renal Function Determination in the Renal  
534 Transplantation Setting. *Clin Pharmacokinet.* 2021;60(9):1201-1215.  
535 doi:10.1007/s40262-021-00998-z
- 536 10. Rong Y, Mayo P, Ensom MHH, Kiang TKL. Population Pharmacokinetics of  
537 Mycophenolic Acid Co-Administered with Tacrolimus in Corticosteroid-Free Adult  
538 Kidney Transplant Patients. *Clin Pharmacokinet.* 2019;58(11):1483-1495.  
539 doi:10.1007/s40262-019-00771-3
- 540 11. Elmokadem A, Riggs MM, Baron KT. Quantitative Systems Pharmacology and  
541 Physiologically-Based Pharmacokinetic Modeling With mrgsolve: A Hands-On Tutorial.  
542 *CPT Pharmacometrics Syst Pharmacol.* 2019;8(12):883-893. doi:10.1002/psp4.12467
- 543 12. Le Louedec F, Puisset F, Thomas F, Chatelut É, White-Koning M. Easy and reliable  
544 maximum a posteriori Bayesian estimation of pharmacokinetic parameters with the

545 open-source R package mapbayr. *CPT Pharmacometrics Syst Pharmacol.*  
546 2021;10(10):1208-1220. doi:10.1002/psp4.12689

547 13. Woillard JB, Debord J, Marquet P. Comment on "Population Pharmacokinetics of  
548 Mycophenolic Acid: An Update." *Clin Pharmacokinet.* 2018;57(9):1211-1213.  
549 doi:10.1007/s40262-018-0687-9

550 14. Bergan S, Brunet M, Hesselink DA, et al. Personalized Therapy for Mycophenolate:  
551 Consensus Report by the International Association of Therapeutic Drug Monitoring and  
552 Clinical Toxicology. *Therapeutic Drug Monitoring.* 2021;43(2):150.  
553 doi:10.1097/FTD.0000000000000871

554  
555  
556

557 **Figure Legends**

558

559 Figure 1. %RMSE trends of individual AUC prediction across residual error scenarios  
560 for each drug model for external datasets (a) and compared to train dataset for  
561 Tacrolimus (b)

562 Figure 2 : Proportion of predicted AUCs within  $\pm 15\%$  of the reference (P15) as a  
563 function of  $\sigma$ , by drug and dataset.

564 Figure 3. Representative individual good fits (a) or overfit (b) obtained using MAPBE  
565 for different scenarios

566 Figure 4: Number of patients for whom each residual error scenario produced the  
567 most accurate AUC estimate in external datasets. (i.e., lowest absolute % error  
568 compared to reference AUC).

569 Figure 5: Proportion of patients for whom the best AUC estimate was obtained using  
570 either the Null, Flat1, or original method, stratified by drug model

571

572 Supplementary data

573 Supplementary table 1. Residual Error Models Evaluated in MAPBE Analysis

<b>Error Model</b>	<b>Proportional Error (CV%)</b>	<b>Additive Error (mg/L or µg/L)</b>	<b>Purpose</b>
<b>Null (minimal error)</b>	<b>10<sup>-8</sup>-8%</b>	<b>0</b>	<b>Maximize weight of observed data</b>
<b>Flat1 (low proportional error)</b>	<b>1%</b>	<b>0</b>	<b>Represent low-level measurement noise</b>
<b>Flat5 (Moderate proportional error)</b>	<b>5%</b>	<b>0</b>	<b>Represent moderate measurement noise</b>
<b>0.8 (Scaled original error)</b>	<b>original × 0.8</b>	<b>original × 0.8</b>	<b>Gradual reduction from original model</b>
<b>0.6 (Scaled original error)</b>	<b>original × 0.6</b>	<b>original × 0.6</b>	<b>Gradual reduction from original model</b>
<b>0.4(Scaled original error)</b>	<b>original × 0.4</b>	<b>original × 0.4</b>	<b>Gradual reduction from original model</b>
<b>0.2(Scaled original error)</b>	<b>original × 0.2</b>	<b>original × 0.2</b>	<b>Gradual reduction from original model</b>
<b>Original published error model</b>	<b>original</b>	<b>original</b>	<b>Baseline for comparison</b>

574

575

576

577 Supplementary Table 2: Overview of the different pharmacokinetic models used

Drug	Model	Compartments	Absorption	Covariates	IIV (ETAs)	RUV Structure
<b>Tacrolimus</b>	Woillard et al.	2 + 3 transit	Transit (3 compartments)	HT, CYP, ST	CL, V1, Q, V2, KTR	Proportional + Additive (0.113, 0.71)
<b>Tacrolimus</b>	Campagne et al.	2	First-order with lag time	CYP	CL, V1, Q, KTR	Proportional only (0.09)
<b>Iohexol</b>	Zwart et al.	2	IV bolus (no absorption)	PT (donor/recipient)	CL, V1, Q, V2	Proportional + Additive (0.0523, 0.0)
<b>Mycophenolate Mofetil (MMF)</b>	Rong et al.	2	First-order with lag time	AcMPAG, AUCMPAG, AUCMPA	Tlag, Ka, CL, Vc, Q, Vp	Proportional + Additive (0.08, 0.32)

578

579

580

581 Supplementary Table 3: Distribution of pharmacokinetic profiles by drug and dataset

582

Drug	Tacrolimus			lohexol	MMF
Dataset	AADAPT	Simulated	PCCP	External	External
<b>Possible timepoints (h)</b>	0 - 0.33 - 0.67 - 1 - 2 - 3 - 4 - 6 - 8 - 12 - 15 - 24	1 sample every 0.05h	0 - 0.33 - 0.67 - 1 - 1.5 - 2 - 3 - 4 - 6 - 8 - 12	0 - 0.16 - 0.33 - 0.5 - 1 - 2 - 3 - 4 - 5 - 6 - 9 - 12 - 18 - 24	0 - 0.33 - 0.67 - 1 - 1.5 - 2 - 3 - 4 - 6 - 9
<b>8 points</b>	0	NA	0	12	0
<b>9 points</b>	3	NA	0	60	1
<b>10 points</b>	30	NA	0	0	15
<b>11 points</b>	1	NA	136	4	0
<b>12 or more points</b>	39	NA	0	20	0
<b>Total</b>	73	450	136	96	16

583

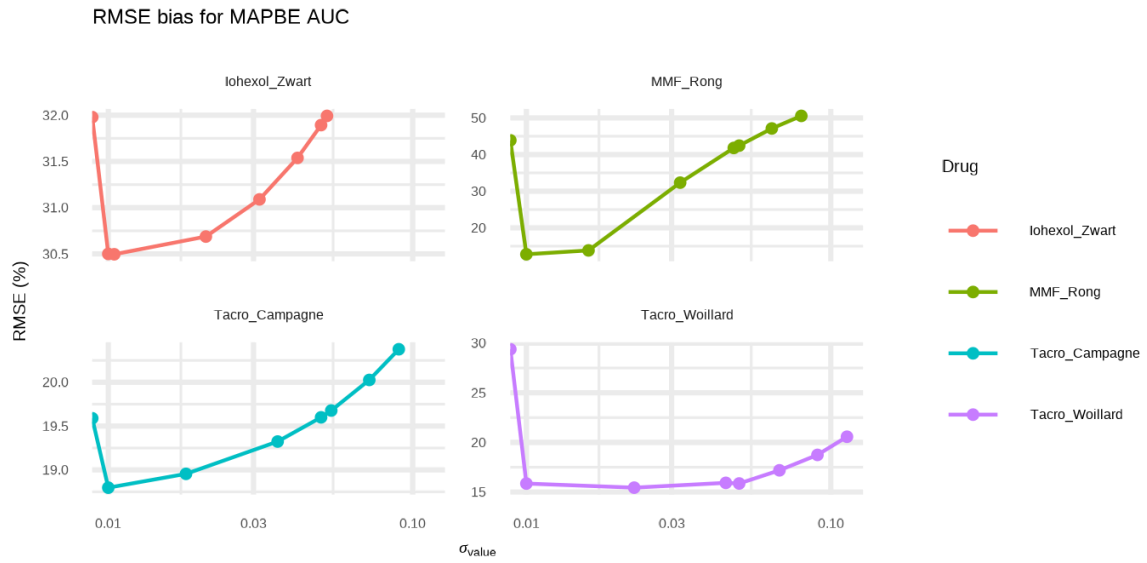
584

585 Supplementary Table 4: percentage of patients) per best-case scenario for each  
 586 drug

Error Model		Tacrolimus			Tacrolimus	Iohexol	MMF
		Woillard		Campagne	Zwart	Rong	
		AADAPT (N=73)	Simulated (N=450)	PCCP (N=136)	PCCP (N=136)	External (N=96)	External (N=16)
Minimal (10 <sup>-8</sup> )	error	40%	68.9%	42.6%	41.2%	62.5%	6.2%
Low	proportional error (1%)	38%	19.1%	22.1%	17.6%	14.6%	56.2%
Moderate	proportional error (5%)	5.3%	2%	5.1%	2.9%	0%	0%
Scaled	original error (80%)	0%	0.4%	2.2%	1.5%	2.1%	0%
Scaled	original error (60%)	1.3%	0.7%	1.5%	2.2%	4.2%	6.2%
Scaled	original error (40%)	2.7%	2.7%	2.2%	4.4%	5.2%	6.2%
Scaled	original error (20%)	13.3%	4.7%	5.9%	8.8%	8.3%	18.8%
Original published	error model	9.3%	1.6%	18.4%	21.3%	3.1%	6.2%

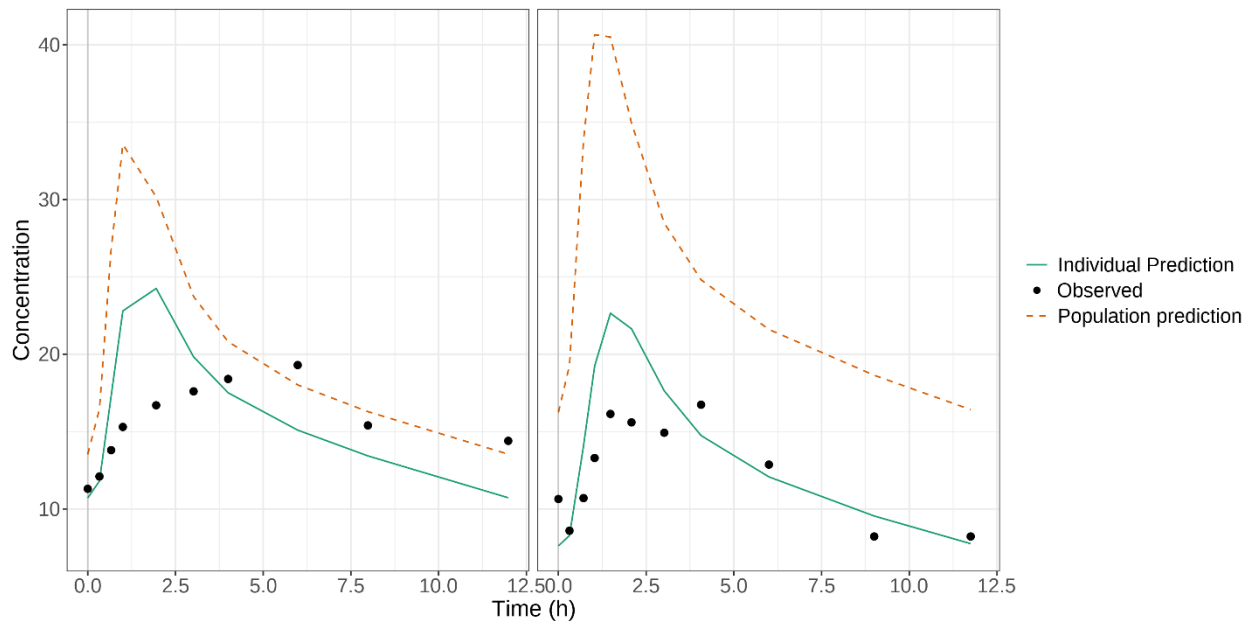
587

588 Supplementary Figure 1 : %RMSE trends of individual AUC prediction across  
589 residual error scenarios for each drug model for external datasets with AUC  
590 estimated by MAPBE integrating the model-predicted concentration–time profile  
591 conditional on all observations from that profile



592  
593

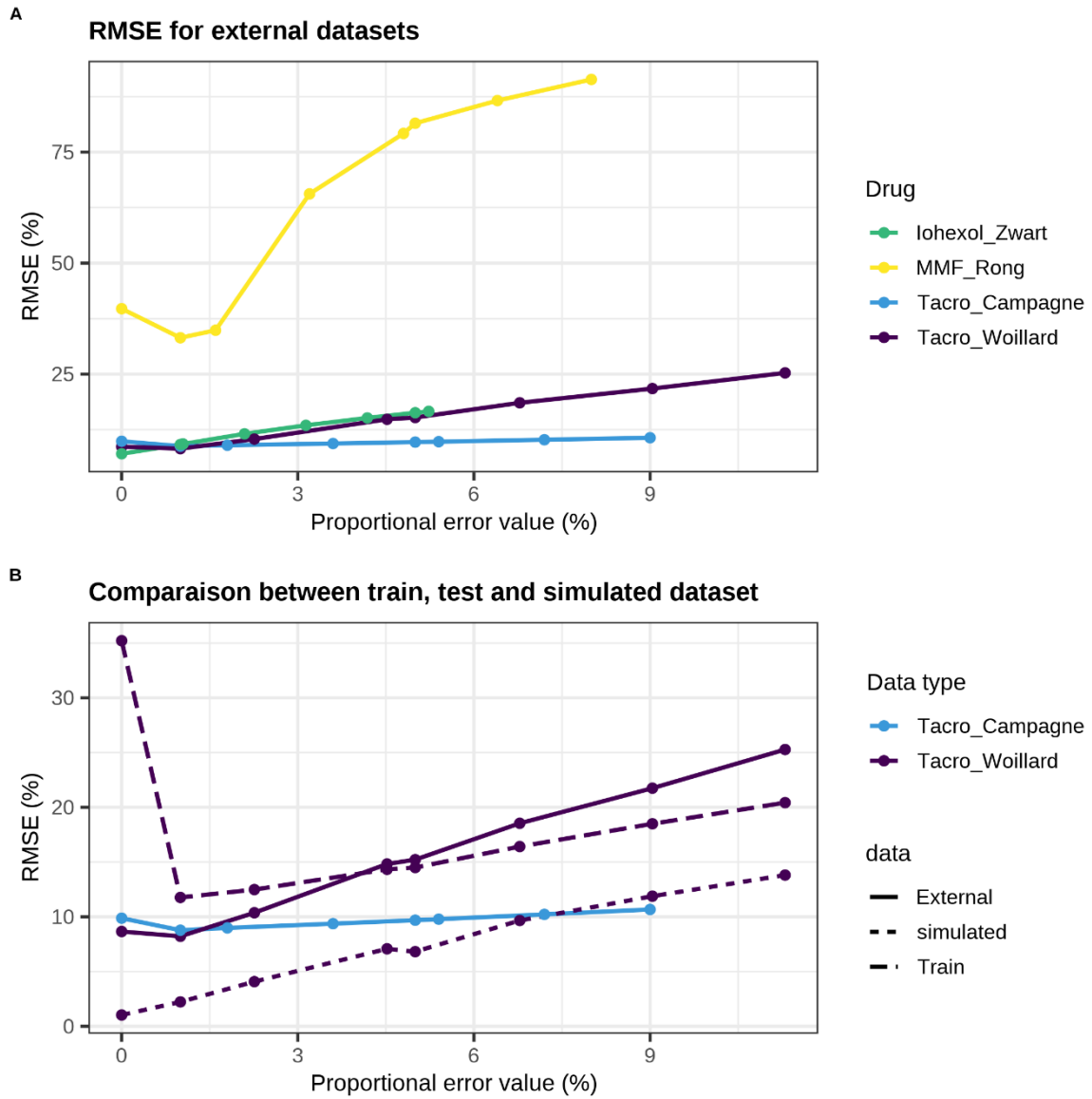
594 Supplementary Figure 2 : Illustration of Shape Distortion in MAP-Bayesian Individual  
595 Predictions for 2 typical patients



596 Each panel shows a different subject. Observed concentrations are represented by black  
597 dots. The Population Prediction (PRED; dashed orange line) shows the typical model  
598 trajectory based on fixed effects and dosing. The Individual Prediction (IPRED; solid green  
599 line) is the MAP-Bayesian estimate, representing a compromise between the population  
600 prior and the likelihood from the subject's own data. This figure illustrates how a dominant  
601 population prior can force the IPRED to adopt its characteristic shape. This "shape  
602 distortion" leads to underfitting of observed peaks and troughs, which in turn biases the  
603 model-predicted AUC. As demonstrated in this work, reducing the residual error gives more  
604 weight to the observed data, allowing the IPRED to better capture the individual's profile and  
605 mitigate this distortion..  
606

607 **Figures:**

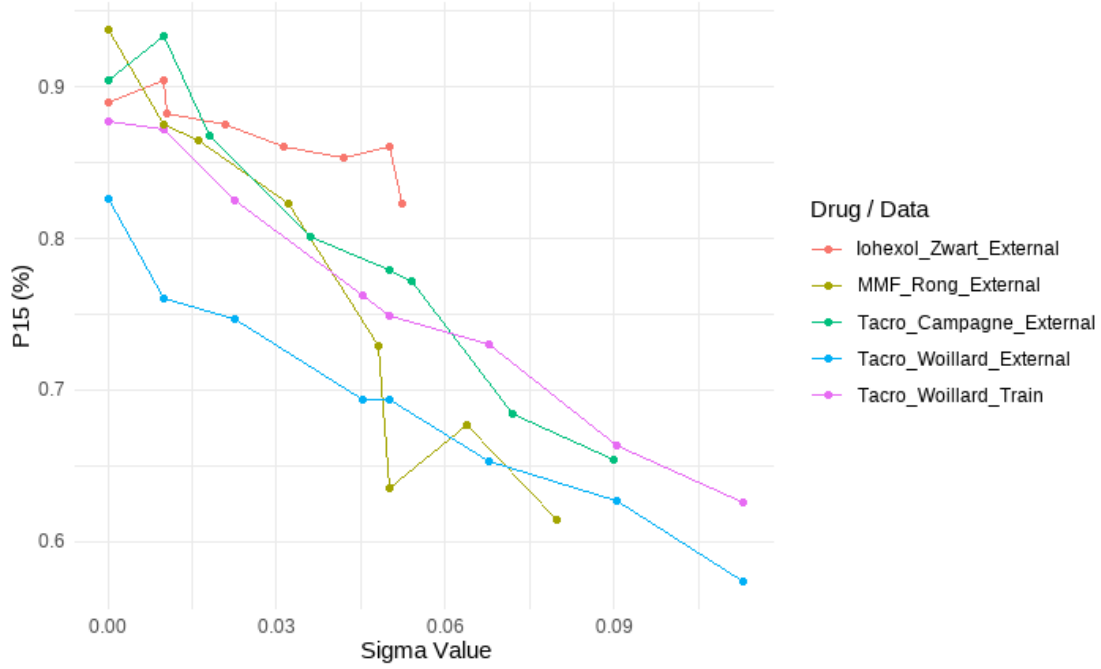
608 Figure 1. %RMSE trends of individual AUC prediction across residual error scenarios  
609 for each drug model for external datasets (a) and compared to train dataset for  
610 Tacrolimus (b)



611

612

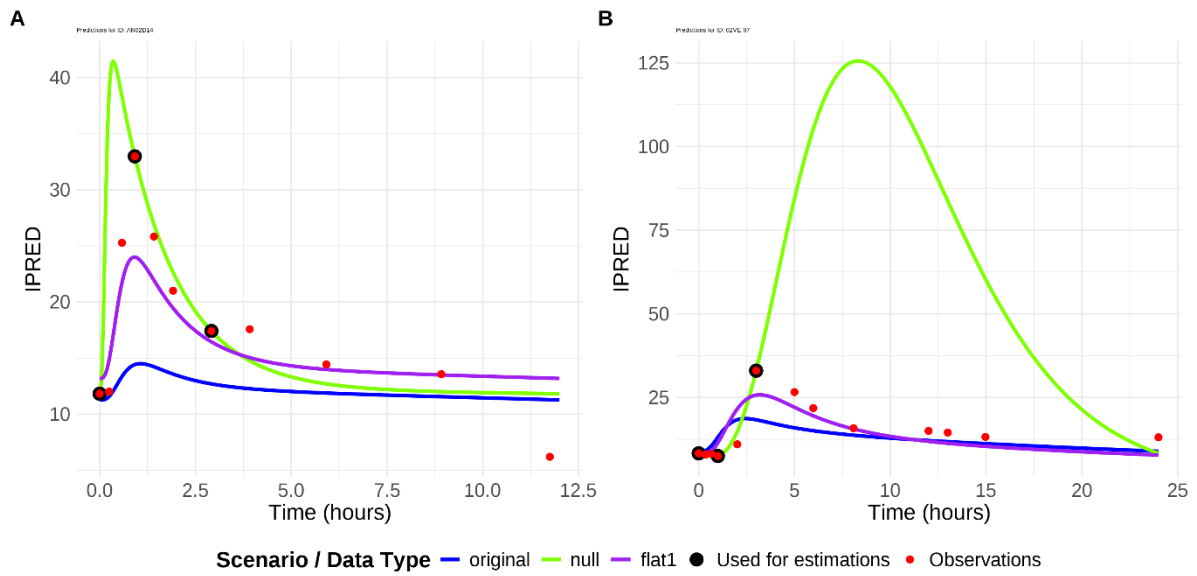
613 Figure 2 : Proportion of predicted AUCs within  $\pm 15\%$  of the reference (P15) as a  
614 function of  $\sigma$ , by drug and dataset.



615

616

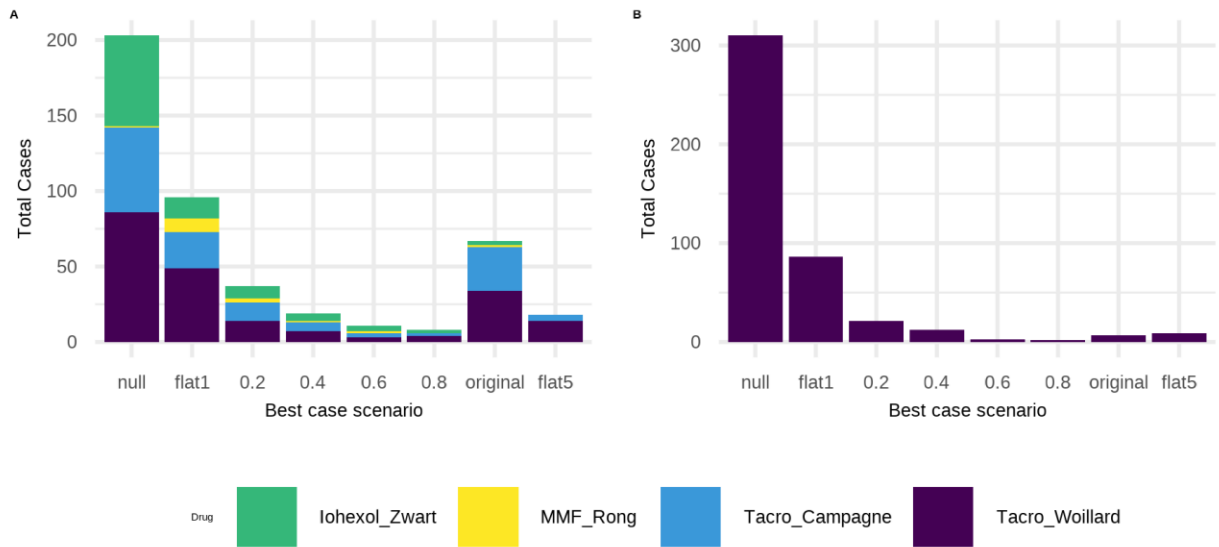
617 Figure 3. Representative individual good fits (a) or overfit (b) obtained using MAPBE  
618 for different scenarios



619

620

621 Figure 4: Number of patients for whom each residual error scenario produced the  
 622 most accurate AUC estimate in external datasets. (i.e., lowest absolute % error  
 623 compared to reference AUC). (A) real clinical datasets only (n = 457 counts; 321  
 624 unique profiles plus the same 136 PCCP profiles counted under both tacrolimus  
 625 models) and (B) simulated cohort (n = 450).



626

627

628 Figure 5: Proportion of patients for whom the best AUC estimate was obtained using  
 629 either the Null, Flat1, or original method, stratified by drug model



630