RESPONSE DOCUMENT

QUALIFICATION –
THIRD LIST OF ISSUES SIMCYP SIMULATOR

December 16, 2024

CONTACT: **KAREN ROWLAND YEO**
SENIOR VICE-PRESIDENT
CERTARA UK LTD
LEVEL 2-ACERO
1 CONCOURSE WAY
SHEFFIELD S1 2BJ
UK

EMAIL: KAREN.YEO@CERTARA.COM

Third List of Issues

PART A: Documentation for COU1-3

1. ***Please provide the Simcyp V19 compound files and the DDI and non-DDI qualification matrix in pdf format. It is envisaged that these files will be made available at the EMA website at the time of the potential qualification opinion.***

Response: Three pdf documents have been provided:

- The DDI qualification matrix
  - Composition, key characteristics of the compounds, clinical studies (design, dosage regimen and subject demographics) and results are described.
- Substrate compound file summaries for Simcyp Simulator V19
  - All compound file summaries for substrates have been compiled into a single pdf – a contents page indicates the starting page number for each individual compound file summary.
  - Previous comments made by the EMA have been taken on board and where possible, the individual compound file summaries have been updated.
- Inhibitor compound file summaries for Simcyp Simulator V19
  - All compound file summaries for inhibitors have been compiled into a single pdf – a contents page indicates the starting page number for each individual compound file summary.
  - Previous comments made by the EMA have been taken on board and where possible, the individual compound file summaries have been updated.

No information has been provided for the non-DDI qualification matrix as this relates to COU, which is no longer part of the qualification.

2. ***Please provide all the information in one folder, with separate sub-folders for each context of use. For each context of use, a table with systems model parameters should be provided and should include relevant and correct references to document their source. For optimized parameters (e.g. Fm, Ki), the correct study used to optimize the parameter should be provided. This information should be clearly understandable for each of the concerned enzymes and the number of studies in the validation set, their size and the doses levels should be presented for each of the concerned enzymes.***

Response: A single folder has been provided. Within this folder, there are 3 subfolders relating to each COU:

- COU 1: contains all data relevant to this COU (fmCYP)
- COU 2: contains all data relevant to this COU (competitive inhibition)
- COU3: contains all data relevant to this COU (MBI)

Within each folder, there is an excel file with several worksheets.

- On the 1st, there is a table showing all the clinical DDI studies relevant to the COU.
- On the 2nd, there is a figure showing predicted *versus* observed DDI ratios.
- On the 3rd, there is a table showing the source of the fm values (COU1), the Ki values (COU2) and inactivation parameters (COU3). It is clearly indicated when a parameter has been sourced from *in vitro* data and when a clinical study has been used for optimization. A link to the clinical study has also been provided.
- A comparison of the 1st and 3rd worksheets indicates that no clinical studies that were used for optimization were included in the DDI matrix – at least to the best of our knowledge!
- In the excel file for COU3, there is an additional worksheet with system parameter data – namely kdeg values. The actual values and the source references have been provided. We felt that provision of kdeg values was important as this parameter directly influences the magnitude of drug-drug interactions for COU3.

It should be noted that as the magnitude of DDI is mainly affected by fmCYP, dose, inhibitory potency (competitive and mechanism-based inhibition), which are all drug-related parameters, no other system data were provided. Hence, our focus was on clarification of the source/optimisation of the key drug-related data.

Within each folder, there is an additional sub-folder which contains the workspaces (population, compound files and study design) that are specific to the Simcyp Simulator V19R1. For each COU, only the workspaces representing the clinical studies simulated as part of the DDI qualification matrix have been included.

We have also provided an additional folder which has the DDI matrix described in its entirety. All of the requested information *(the number of studies in the validation set, their size and the doses levels should be presented for each of the concerned enzymes)* has been included in this

matrix and for each matrix relevant to each COU. A subfolder contains all publications for the clinical studies that were simulated.

PART B. Bias and Uncertainty

3. ***The SAWP would like the applicant to challenge the assumption of constant bias and between study variability (extra_var_of_ratio) across predicted GMRs. To this end, the applicant is asked to present and discuss alternative models that are not based on these assumptions. When discussing the models, the applicant is advised to take into account the SAWP's concerns around the presentation/assessment of the goodness-of-fit of the Bayesian meta-regression model as discussed above.***

Model-based uncertainty quantification (UQ) for COU1-3 based on the DDI QM

4. ***Please replicate the analyses, metrics and visuals developed by the SAWP.***
5. ***Please discuss the SAWP recommendations on the Bayesian meta-regression model and visuals as expanded in the scientific discussion.***
6. ***For CoU1, please discuss the possibility of assessing the predictive performance of Simcyp in predicting the effect of weak and moderate inhibitors, respectively. For this discussion, the Applicant should present posterior predictive checks (see Figure 1) only including substrate-inhibitor pairs including weak and moderate inhibitors that have not been used to optimise compound files.***

Response: Replies to the above comments/questions are captured globally in the description/discussion below.

Introduction

The model bias and imprecision model originally proposed by Certara (see conceptual representation in Figure 1) used a mix of stratification (by inhibition mechanism and by CYP) and hierarchical modelling (for between-subject variability); see Figure 2 left pane for a graph of its assumed statistical dependencies. This model, which we hereafter call "model A", is in a sense minimal. It has been recoded by the EMA using the *R* package RStan (script "m201.stan"), with minimal modifications.

Model A has been reviewed by the EMA who proposed an alternative formulation (model B, script "m200.stan", Figure 2 right pane) getting rid of preliminary stratification and introducing the type of inhibition as a covariate effect on GMR bias and between-study variability.

The EMA also asked us to examine the effect of using weak or moderate inhibitors on biases and between-study variability. To that effect, we developed model C based on model B. These three models are detailed and their results compared and discussed in the following.

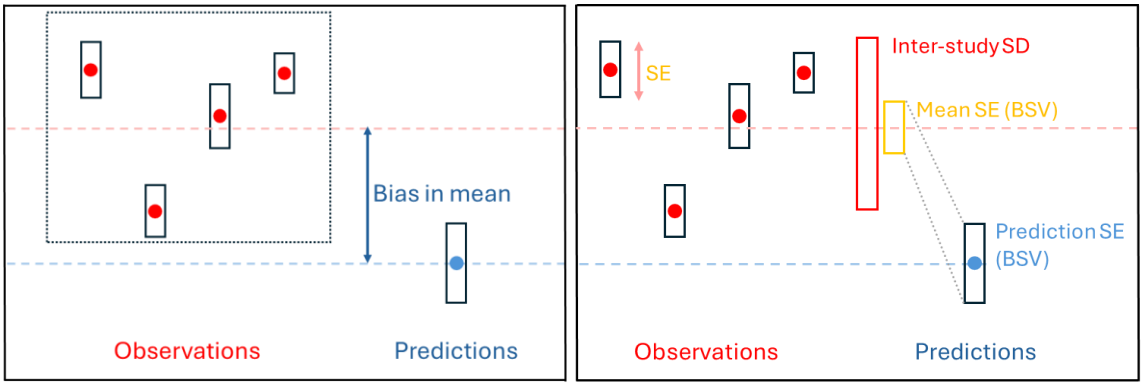

Figure 1: Schematic representation of the conceptual model. Left: Mean bias between observed and predicted geometric mean ratios. Right: The different variabilities and uncertainties studied; Observed and predicted between-subject variance (BSV) may differ by a factor to estimate (BSV bias); Inter-study variability can also be estimated, and it lumps in fact inter-study variability and potential prediction imprecision.

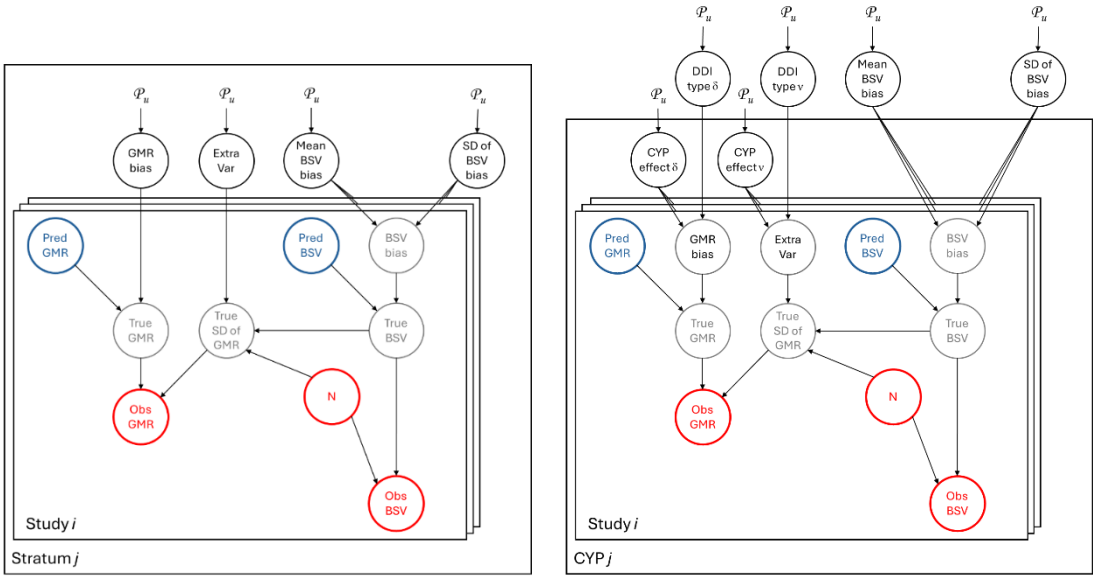

Figure 2: Directed acyclic graph representations of the model dependencies between variables and parameters. Left: model A; Right: model B. Literature data are in red; Simcyp® predictions in blue; Latent variables in grey; Estimands in black.

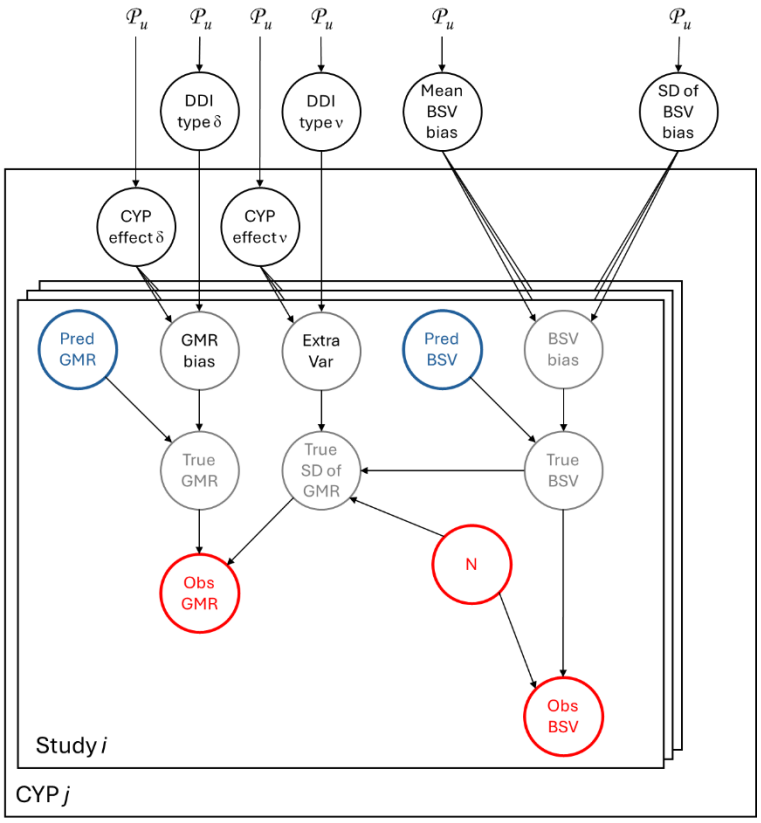

Figure 3: Directed acyclic graph representations of model's C dependencies between variables and parameters. Literature data are in red; Simcyp® predictions in blue; Latent variables in grey; Estimands in black.

Data

The same data as previously were prepared for analysis (see Certara/Simcyp. EMA Initial Qualification Procedure – Second List of Issues Simcyp Simulator – Response Document. June 26, 2024).

DB08: After preprocessing of the data as described previously (Certara/Simcyp. EMA Initial Qualification Procedure – Second List of Issues Simcyp Simulator – Response Document. June 26, 2024) 220 studies remained for which both the GMR and the geometric SD of individual ratios were available in clinical studies and used in the models. There were 246 clinical studies in the DDI qualification matrix.

Original statistical bias and imprecision model (model A)

The original model bridges Simcyp's predictions of GMR and observed GMR with a set of free parameters characterizing biases and uncertainties (see Certara/Simcyp. EMA Initial Qualification Procedure – Second List of Issues Simcyp Simulator – Response Document. June

26, 2024). It was applied to stratified data. It has been recoded in *Stan* and a few modifications were made which are indicated in red below:

The GMR observed in study i (out of $N_{studies}$) was assumed to be lognormally distributed around a true GMR ($logGMR_i$, in log-space), with true standard deviation $\sigma_{GMR,i}$ (in log-space):

$$GMR_{i,obs} \sim \mathcal{LN}\left(logGMR_i, \sigma_{GMR,i}\right) \quad (1)$$

The true GMR was assumed equal to the Simcyp-predicted value, $GMR_{i,pred}$, corrected for potential bias, β_{GMR}:

$$logGMR_i = \log\left(GMR_{i,pred} \times (1 + \beta_{GMR})\right) \quad (2)$$

This was rewritten by EMA as

$$logGMR_i = \log\left(GMR_{i,pred}\right) + \beta'_{GMR} \quad (3)$$

so that

$$\beta'_{GMR} = log(1 + \beta_{GMR}). \quad (4)$$

The true variance in log-space of $GMR_{i,obs}$, $\sigma^2_{GMR,i}$, was assumed to be equal to the sum of between-study variance, σ^2_{stu}, and a sampling variance equal to the true between-subject variance $\sigma^2_{sub,i}$, scaled by $N_{sub,i}$, the number of subject in study i:

$$\sigma^2_{GMR,i} = \sigma^2_{stu} + \frac{\sigma^2_{sub,i}}{N_{sub,i}} \quad (5)$$

Note that the between-study variance, σ^2_{stu}, may include a component due to Simcyp's imprecision in predictions, but this component is not separately estimable.

The true between-subject variance in log-space for study i was assumed to be equal to the Simcyp-predicted between-subject variance (in log-space), corrected for potential variance bias, $\beta_{BSV,i}$:

$$\sigma^2_{sub,i} = \sigma^2_{sub,i,pred} \times \beta_{BSV,i} \quad (6)$$

This was rewritten by the EMA as:

$$\sigma^2_{sub,i} = \sigma^2_{sub,i,pred} \times \exp\left(\beta'_{BSV,i}\right) \quad (7)$$

so that

$$\beta'_{BSV,i} = log\left(\beta_{BSV,i}\right). \quad (8)$$

A hierarchical specification for $\beta_{BSV,i}$, which allows for information sharing in BSV predictions bias estimates, was used:

$$\beta_{BSV,i} \sim \mathcal{N}\left(\mu_{\beta BSV}, {\sigma_{\beta BSV}}^2\right) \quad (9)$$

An analogous specification was used for $\beta'_{BSV,i}$ by the EMA:

$$\beta'_{BSV,i} \sim \mathcal{N}\left(\mu'_{\beta BSV}, {\sigma'_{\beta BSV}}^2\right) \quad (10)$$

Finally, the observed between-subject variance (in log-space) for study i is part of the data. According to the standard lognormal error model, it was assumed to be distributed according to the following gamma distribution, specified through shape and scale parameters k_i and θ_i respectively:

$$\sigma^2_{sub,i,obs} = \left[log\left(GSD_{i,obs}\right)\right]^2 \sim G\ (k_i, \theta_i) \quad (11)$$

$$k_i = N_{sub,i}/2 \quad (12)$$

$$\theta_i = \sigma^2_{sub,i}\ /k_i \quad (13)$$

The EMA used an alternative, but equivalent, parameterization of the gamma distribution, using rate instead of scale; rate is simply the inverse of scale.

The prior for the log GMR prediction bias, β_{GMR}, was set to a uniform distribution:

$$\beta_{GMR} \sim \mathcal{U}\,(-1, 10) \quad (14)$$

Logically, β_{GMR} cannot be lower than -1. We made sure that the upper bound was high enough to be never reached during MCMC sampling.

The EMA used a different prior for β'_{GMR}, but lying on the log-scale it does not have the same constraints:

$$\beta'_{GMR} \sim \mathcal{N}(0, 1) \quad (15)$$

The priors for the parameters of the variance of prediction bias, $\mu_{\beta BSV}$ and $\sigma_{\beta BSV}$ were respectively set to a uniform distribution (with upper bound high enough to be never reached during sampling) and to a normal distribution consistent with a degree of shrinkage:

$$\mu_{\beta BSV} \sim \mathcal{U}\,(0, 50) \quad (16)$$

$$\sigma_{\beta BSV} \sim \mathcal{N}(0{,}0.3)\ [0, 50] \quad (17)$$

The EMA used the following normal and truncated Cauchy priors for the corresponding $\mu'_{\beta BSV}$ and $\sigma'_{\beta BSV}$:

$$\mu'_{\beta BSV} \sim \mathcal{N}(0, 1) \quad (18)$$

$$\sigma'_{\beta BSV} \sim C\,(0{,}1)\ [0, \infty[\quad (19)$$

The between-study variance, σ^2_{stu}, was assigned a vague truncated half-normal prior on the log scale:

$$\sigma^2_{stu} \sim \mathcal{N}(0,1)\ [0,5] \quad (20)$$

The EMA used instead a Cauchy prior, truncated to positive values:

$$\sigma^2_{stu} \sim C\,(0,1)\ [0,\infty[\quad (21)$$

EMA-proposed statistical bias and imprecision model (model B)

The new model proposed by the EMA is based on the EMA variant of the previous one.

The GMR observed in study i (out of $N_{studies}$) is still assumed to be lognormally distributed around a true GMR ($logGMR_i$, in log-space), with true standard deviation $\sigma_{GMR,i}$ (in log-space):

$$GMR_{i,obs} \sim \mathcal{LN}\left(logGMR_i, \sigma_{GMR,i}\right) \quad (22)$$

Similar to EMA's implementation of the previous model, the true GMR is assumed equal to the Simcyp-predicted value, $GMR_{i,pred}$, but it is now corrected by a study-specific bias $\beta'_{GMR,i}$:

$$logGMR_i = \log\left(GMR_{i,pred}\right) + \beta'_{GMR,i} \quad (23)$$

The bias for study i depends on the CYP enzyme mediating the studied interaction and on the type of interaction occurring (competitive *vs*. mechanism-based):

$$\beta'_{GMR,i} = \delta_{GMR,CYP_i} + \mathbb{I}_{inh_i} \times \delta_{GMR,inh} \quad (24)$$

Parameter δ_{GMR,CYP_i} measures the bias affecting simulations of the CYP examined in study i; $\mathbb{I}_{inh_i}$ is an indicator function taking the value 1 if study i examined a mechanism-based inhibition and value 0 if it examined a competitive inhibition; $\delta_{GMR,inh}$ is a parameter measuring the effect on bias of studying a mechanism-based inhibition.

The true variance in log-space of $GMR_{i,obs}$, $\sigma^2_{GMR,i}$, is again assumed to be equal to the sum of between-study variance, $\sigma^2_{stu,i}$ (which is assumed to depends on the CYP enzyme mediating the studied interaction and on the type of interaction occurring), and a sampling variance equal to the true between-subject variance $\sigma^2_{sub,i}$, scaled by $N_{sub,i}$, the number of subject in study i:

$$\sigma^2_{GMR,i} = \sigma^2_{stu,i} + \frac{\sigma^2_{sub,i}}{N_{sub,i}} \quad (25)$$

Between-study variance, $\sigma^2_{stu,i}$ is computed as:

$$\sigma^2_{stu,i} = \nu_{stu,CYP_i} \times \exp\left(\mathbb{I}_{inh_i} \times \nu_{stu,inh}\right) \quad (26)$$

Parameter ν_{stu,CYP_i} is a CYP-specific between-study variance affecting simulations of the CYP examined in study i; $\mathbb{I}_{inh_i}$ is the same indicator function as above; $\nu_{GMR,inh}$ is a parameter measuring the effect on between-study variance of studying a mechanism-based inhibition.

As before, between-study variances, $\sigma^2_{stu,i}$, may include a component due to Simcyp's imprecision in predictions, but this component is not separately estimable.

The true between-subject variance in log-space for study i was assumed to be equal to the Simcyp-predicted between-subject variance (in log-space), corrected for potential variance bias, $\beta_{BSV,i}$, as EMA's implementation of the previous model:

$$\sigma^2_{sub,i} = \sigma^2_{sub,i,pred} \times \exp\left(\beta'_{BSV,i}\right) \tag{27}$$

Here also, a hierarchical specification for $\beta'_{BSV,i}$ allows for information sharing in BSV predictions bias estimates:

$$\beta'_{BSV,i} \sim \mathcal{N}\left(\mu'_{\beta BSV}, {\sigma'_{\beta BSV}}^2\right) \tag{28}$$

As in the previous model, with EMA reparameterization, the observed between-subject variance (in log-space) for study i is part of the data and assumed to be distributed according to the following gamma distribution, specified through shape and rate parameters k_i and τ_i respectively:

$$\sigma^2_{sub,i,obs} = \left[log\left(GSD_{i,obs}\right)\right]^2 \sim G\left(k_i, \tau_i\right) \tag{29}$$

$$k_i = N_{sub,i}/2 \tag{30}$$

$$\tau_i = k_i/\sigma^2_{sub,i} \tag{31}$$

The priors for the parameters of the log GMR prediction bias, δ_{GMR,CYP_i} and $\delta_{GMR,inh}$, were both normal:

$$\delta_{GMR,CYP_i} \sim \mathcal{N}(0,1) \tag{32}$$

$$\delta_{GMR,inh} \sim \mathcal{N}(0,1) \tag{33}$$

The priors for the parameters of the BSV prediction bias, $\mu'_{\beta BSV}$ and $\sigma'_{\beta BSV}$, were again normal and truncated Cauchy:

$$\mu'_{\beta BSV} \sim \mathcal{N}(0,1) \tag{34}$$

$$\sigma'_{\beta BSV} \sim C\left(0,1\right)\left[0,\infty\right[\tag{35}$$

The priors for the parameters of between-study variances, $\sigma^2_{stu,i}$, ν_{stu,CYP_i} and $\nu_{stu,inh}$, were respectively assigned a Cauchy prior, truncated to positive values and a standard normal:

$$\nu_{stu,CYP_i} \sim C(0,1)\ [0,\infty[\quad (36)$$

$$\nu_{stu,inh} \sim \mathcal{N}(0,1) \quad (37)$$

Bias and imprecision model including inhibition strength (model C)

To assess the performance of the Simcyp® Simulator V19R1 in predicting the effect of weak and moderate inhibitors, the above model, proposed by EMA, was modified to include the strength of the inhibitor used as a covariate of study bias (Eq. 24) and between-study variance (Eq. 26). Those two equations were modified as follows:

Bias for study *i* depends on the CYP enzyme mediating the studied interaction, on the type of interaction occurring (competitive *vs*. mechanism-based), and on the strength of the inhibitor used in study *i*:

$$\beta'_{GMR,i} = \delta_{GMR,CYP_i} + \mathbb{I}_{inh_i} \times \delta_{GMR,inh} + \mathbb{I}_{str_i} \times \delta_{GMR,str} \quad (38)$$

The new variable $\mathbb{I}_{str_i}$ is an indicator function taking the value 1 if study *i* used a moderate inhibitor and value 0 if it used a weak inhibitor; $\delta_{GMR,inh}$ is a parameter measuring the effect on bias of using a moderate inhibitor.

Between-study variance, $\sigma^2_{stu,i}$ also include a possible effect of inhibitor strength:

$$\sigma^2_{stu,i} = \nu_{stu,CYP_i} \times \exp\left(\mathbb{I}_{inh_i} \times \nu_{stu,inh} + \mathbb{I}_{str_i} \times \nu_{GMR,str}\right) \quad (39)$$

Parameter $\nu_{GMR,inh}$ measures the effect (on between-study variance) of using a moderate inhibitor.

The two new parameters are assigned normal priors:

$$\delta_{GMR,str} \sim \mathcal{N}(0,1) \quad (40)$$

$$\nu_{stu,str} \sim \mathcal{N}(0,1) \quad (41)$$

Methods and software

As previously, inference for model A was performed on all the data, but stratified according to mechanism of inhibition, and according to CYP enzyme involved. Inference for model B does not require stratification. Inference itself was performed using Hamiltonian Markov chain Monte Carlo (HMCMC) simulations with the *R* package *RStan* and all post-processing and plotting was done in *R* [1].

Results

Replication of the analyses, metrics and visuals developed by EMA

All the scripts provided by the EMA ran well and we can reproduce their results on our computing platform. The extension of model B to model C (with effect of the inhibitor strength) also runs well.

Convergence of MCMC and goodness of fit for models A, B, and C

HMCMC sampling converged well for all models. Gelman and Rubin $\hat{R}$ diagnostic values (which should be close to 1 at convergence) were at most 1.004 for all samples variables of all three models (model A was run in its modified version with *RStan*).

Figure 4 compares the distributions of published observed geometric ratios (top row) and between-subject variability in individual ratios (bottom row) to distributions of pseudo-data simulated using the parameters' joint posterior distribution of models A, B, and C. The leftmost panel of the Figure also shows the distribution of the published observations and raw Simcyp® Simulator predictions, for reference. The fits are quite good, and model B seems better than model A. The fit of model C also appears reasonable; the shape of its distributions differ from those of models A and B because the database for model C is a subset of that for the other two models (studies using strong inhibitors were not considered in model C). The bias corrections are obviously smaller on GMRs than on between-subject variability.

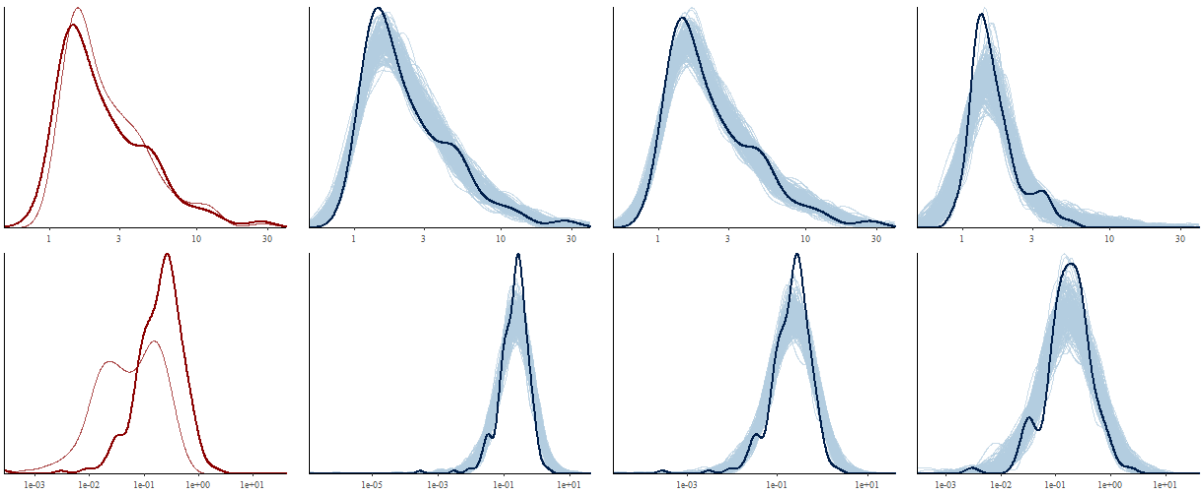

Figure 4: Comparisons of observed (dark lines) geometric ratios (top row) and between-subject variability in individual ratios (bottom row) with raw Simcyp predictions (left column) or posterior bias model predictions simulated using the parameters joint posterior distribution for models A, B, and C (three rightmost columns, respectively).

Comparison of parameter estimates for models A, B, and C

Since the models have converged and seem to fit the data, we can now examine the parameter estimates obtained (as samples from their joint posterior distribution). Table 1 gives the means and SDs of the parameters' marginal posterior distributions for the three models considered. Its companion, Table 2, gives the number of studies used per CYP and per model. For the main parameters of interest, Figure 5 shows a comparison of models A and B mean GMR bias, Figure 6 does the same for between-study variance estimates, Figure 7 and Figure 8 compare model B and model C. The estimates are very similar between model A and B. Model A parameters were transformed to be directly comparable with their model B equivalents (see Eqs. 4 and 8); in particular, GMR biases are on the log scale and correspond approximately to the percent change that should be applied to Simcyp predictions to recover the data (*e.g.*, $0.1 \approx 10\%$).

Mean GMR biases range from about -17% (for CYP2D6) to +2.5% according to model A and model B in the case of competitive inhibition. Those estimates are reasonable well-identified (Figure 5 and Figure 7). In the case of MBI, model B estimates a 4% up-shift in those values; the range would therefore be from -13% to 6.5% approximately, with less bias overall. Model C gives a mean GMR bias range from -29% to -1% in the case of competitive inhibition and weak inhibitors, with more uncertainty in the estimates because of the reduced data base (only three studies for CYP2C19, in particular). The estimate of inhibition type effect on GMR bias, for model C, is lower than that of model B (+1% in the case of MBI), and the effect of inhibitor strength is about +0.5% in the case of moderate inhibitors compared to weak inhibitors.

Between-study variances (on the log-scale) range from 0.014 to 0.16 for models A and B; this corresponds to CVs ranging from 12% to 41% on the natural scale. Remember that this is an upper limit of Simcyp imprecision. Model C identifies similar variances, except in the case of CYP2C19, but there are only 3 studies for that CYP and the estimate is very imprecise. For CYP2C8 and CYP2C9 in models A and B, and for those plus CYP1A2 for model C, it appears that the lower tails of the posterior distributions reach unrealistically low values (almost zero Simcyp imprecision *and* zero true between-study variability). This is probably due to identifiability problems but does not seem to affect the plausibility of the central estimates obtained. Models B and C identify different effects of the type of inhibition on between-subject variances: an increase by a factor 1.3 (model B) or 1.55 (model C) for each CYP in the case of mechanism-based inhibition compared to competitive inhibition. Model C identifies a further increase by a factor 1.7 in the case of moderate inhibitors compared to weak inhibitors.

The mean bias in between-subject variability is about a factor 2 (that is, the ratio between observed and predicted between-subject CV is $\sqrt{\exp(1.40)} = 2.0$) according to model B, with a large variability across studies. Models A and C bias estimates for between-subject variability are closer to factor 2.65, but this is expected because model A does not pool information about it across CYPs and model C has a reduced data base.

Overall, model B allows for better pooling of information (this is obvious given its structure, shown in Figure 2) and it yields estimates somewhat more homogeneous and precise. This is particularly true in the case of between-subject variability bias, which should be better estimated in model B. Model C (Figure 3) also pools information and estimates an inhibitor strength effect (if weak or moderate), but its identifiability and precision of estimates suffer from its much reduced dataset (see Table 2).

Table 1: Statistical summaries of the posterior distributions of the main parameters for models A, B, and C.

Parameter	Model A		Model B		Model C	
	Mean	SD	Mean	SD	Mean	SD
Mean GMR biases*						
CYP1A2	-0.0841	0.0347	-0.0709	0.0314	-0.0097	0.0485
CYP2C19	-0.0982	0.0939	-0.1019	0.1011	-0.0517	0.4725
CYP2C8	0.0252	0.0687	-0.0332	0.0628	-0.0993	0.0864
CYP2C9	-0.0374	0.0406	-0.0600	0.0393	-0.0440	0.0658
CYP2D6	-0.1548	0.0575	-0.1661	0.0539	-0.2884	0.0863
CYP3A4	0.0010	0.0306	-0.0083	0.0309	-0.0624	0.0533
Between-study variances*						
CYP1A2	0.0296	0.0119	0.0213	0.0102	0.0223	0.0161
CYP2C19	0.1574	0.0679	0.1471	0.0657	0.7750	1.0515
CYP2C8	0.0541	0.0380	0.0306	0.0245	0.0390	0.0450
CYP2C9	0.0173	0.0114	0.0140	0.0104	0.0086	0.0079
CYP2D6	0.0923	0.0328	0.0641	0.0266	0.0790	0.0469
CYP3A4	0.0550	0.0121	0.0327	0.0093	0.0374	0.0162
Effects of inhibition type						
on bias	-†	-†	0.0389	0.0448	0.0128	0.0765
on between-study variance	-†	-†	0.2798	0.2195	0.4403	0.3226
Effects of inhibitor strength						
on bias	-†	-†	-†	-†	0.0048	0.0639
on between-study variance	-†	-†	-†	-†	0.5173	0.3563
BSV‡ bias mean	2.0244	0.0702	1.3994	0.0950	1.9114	0.1372
BSV bias SD	1.9396	0.0638	1.2853	0.0720	1.3307	0.1112

* For model B those estimates are for the competitive inhibition (CI) case; For model C, they are for the case of CI and weak inhibitors.
‡ BSV: Between-subject variance.
† Not computed by the model.

Table 2: Number of studies considered for each CYP by the different models. Model C does not consider studies using strong inhibitors.

CYP	Number of studies	
	Models A and B	Model C
CYP1A2	39	16
CYP2C19	23	3
CYP2C8	17	9
CYP2C9	19	17
CYP2D6	40	18
CYP3A4	82	45
Total	**220**	**108**

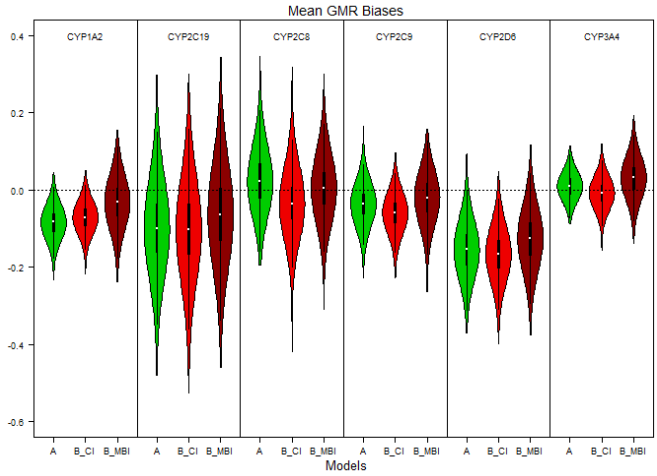

Figure 5: Comparison of the posterior distributions of mean GMR biases by CYP for models A and B (in that case, for competitive inhibition or mechanism-based inhibition).

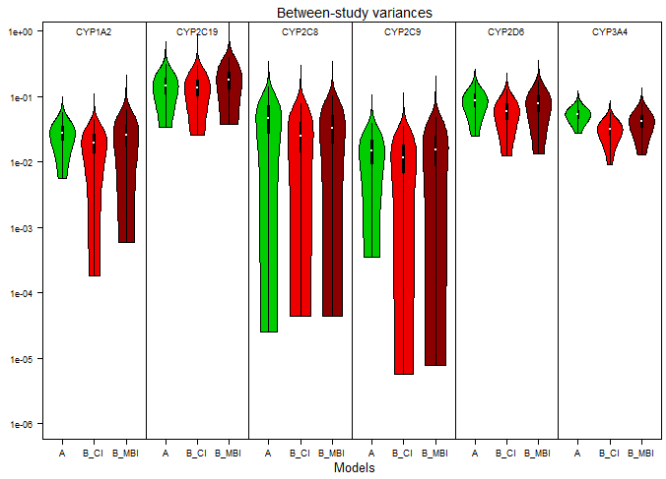

Figure 6: Comparison of the posterior distributions of between-study variances by CYP for models A and B (in that case, for competitive inhibition or mechanism-based inhibition).

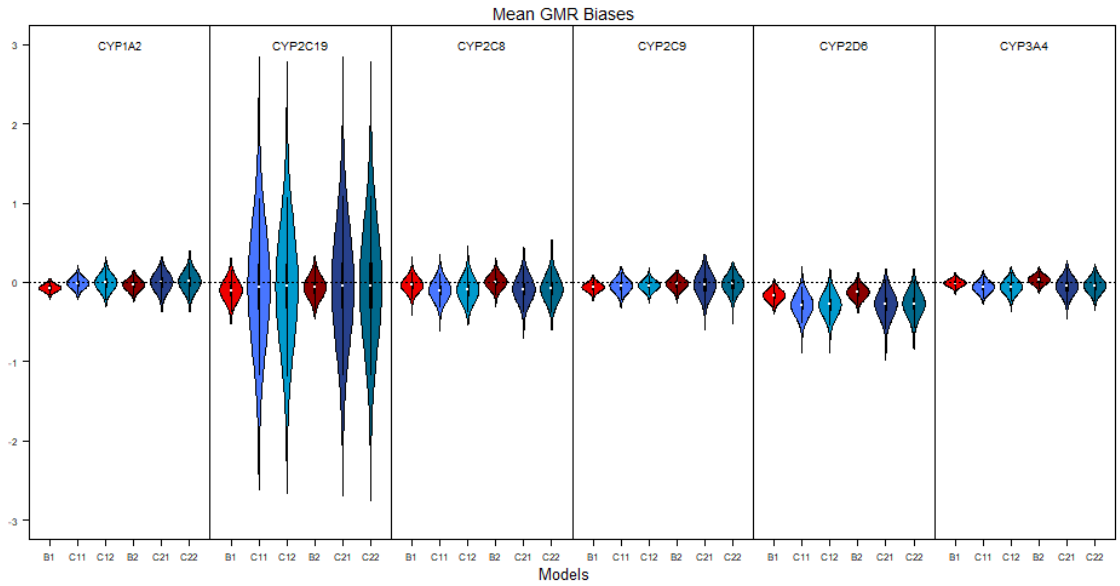

Figure 7: Comparison of the posterior distributions of mean GMR biases by CYP for models B and C. X-axis labels: B1 for model B, competitive inhibition case; C11 for model C, competitive inhibition and weak inhibitor case; C12 for model C, competitive inhibition and moderate inhibitor case; B2 for model B, MBI case; C21 for model C, MBI and weak inhibitor case; C22 for model C, MBI and moderate inhibitor case. The very large uncertainty in model C for CYP2C19 is due to the low number of studies available for that case (N=3)

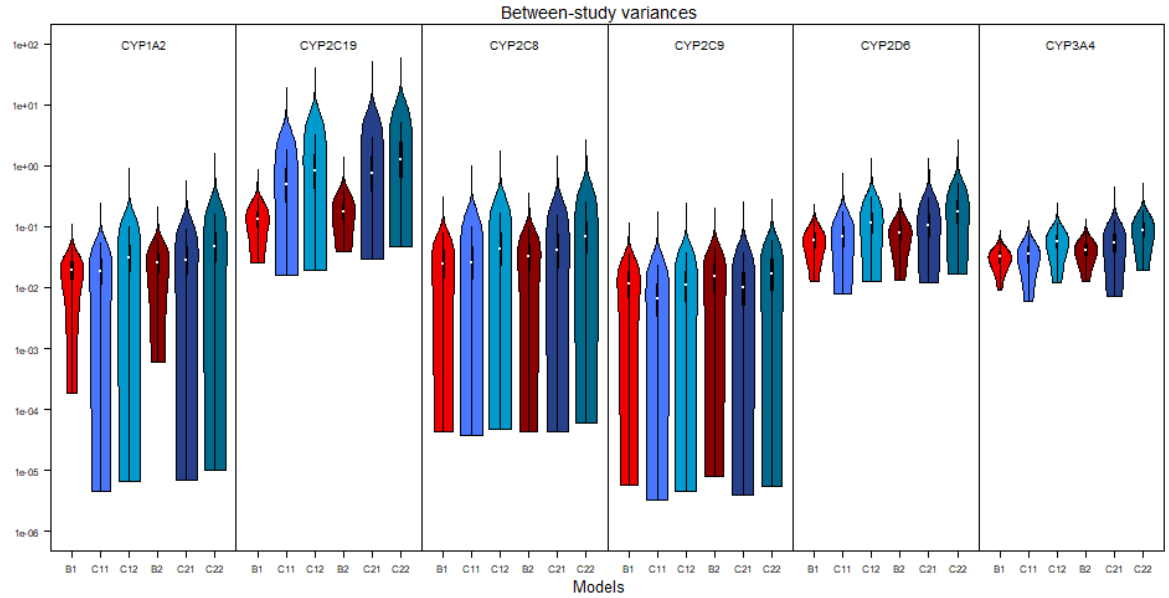

Figure 8: Comparison of the posterior distributions of between-study variances by CYP for models B and C. X-axis labels: B1 for model B, competitive inhibition case; C11 for model C, competitive inhibition and weak inhibitor case; C12 for model C, competitive inhibition and moderate inhibitor case; B2 for model B, MBI case; C21 for model C, MBI and weak inhibitor case; C22 for model C, MBI and moderate inhibitor case.

Posterior predictive plots for models C

The EMA proposed several predictive plots generated using the above model-based discrepancy analysis for application in drug development. Such plots are presented below for model C.

Credibility interval vs predicted GMR

The above meta-analysis models can be used to understand how uncertainty affects future DDI predictions for CYP inhibition in regulatory decision-making.

Figure 9 displays 90% credibility intervals for GMRs (*i.e.*, fold-changes) according to model C, by CYP and according to type of inhibition and strength of inhibitor studies. Posterior samples of GMR bias and between-study variability were used to sample a log-normal distribution, as per the scripts provided by EMA. The results are quite consistent across covariate effects: predictions of GMRs for CYP2C19 have the highest uncertainty, followed by CYP2D6.

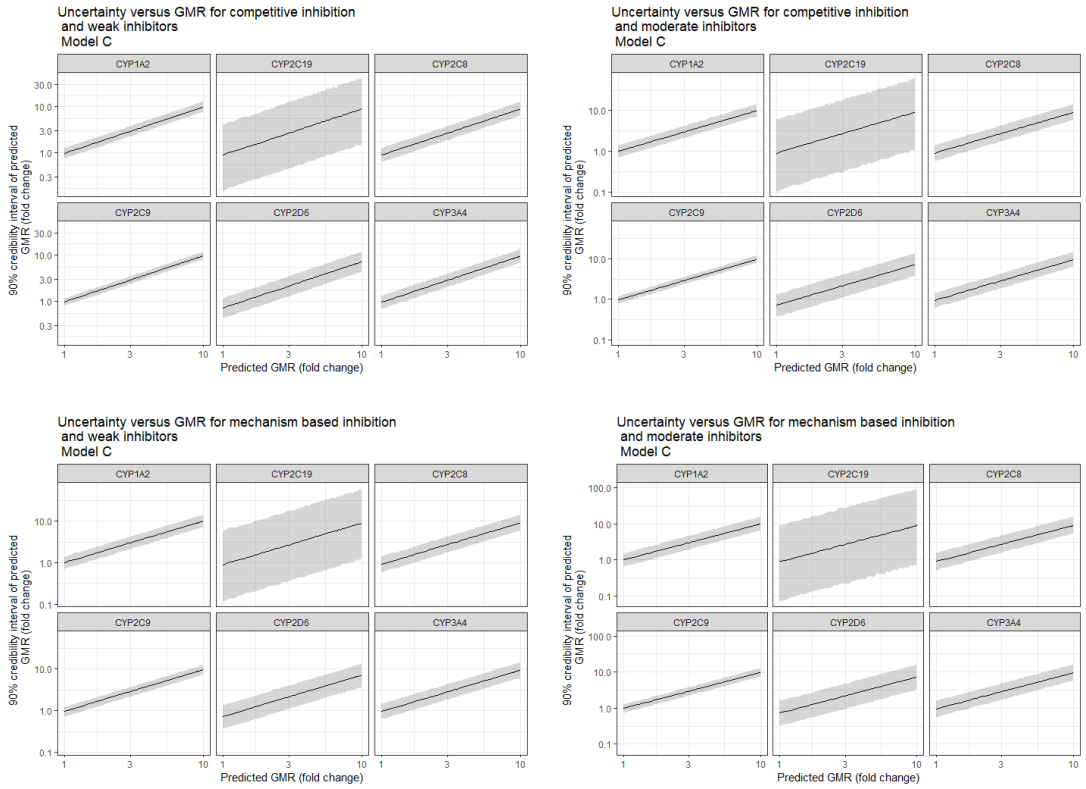

Figure 9: Model C-generated credibility intervals of predicted GMR values *vs*. those values for competitive inhibition (top row) or mechanism-based inhibition (bottom row), in the case of a weak inhibitor (left column) or moderate inhibitor (right column). The x-axis values represent hypothetical GMR point estimates predicted by model C. The grey shaded are 90% credibility intervals of the prediction.

Predicted GMR for hypothetical CYP substrates

The EMA suggested that the above displays could be extended to include information about the therapeutic range. Figure 10 shows the results obtained for a hypothetical scenario in which:

- A CYP substrate is being developed for which there is an adequate PBPK model.
- The therapeutic index is known for the drug in question and was hypothetically set to 0.5 to 2-fold compared to the expected geometric mean exposure at the therapeutic dose.
- A hypothetical DDI is predicted using the Simcyp® Simulator following concomitant administration with a weak or moderate CYP inhibitor. The type of inhibition is also considered. Model C posteriors were used, as above, to estimate GMR uncertainty.

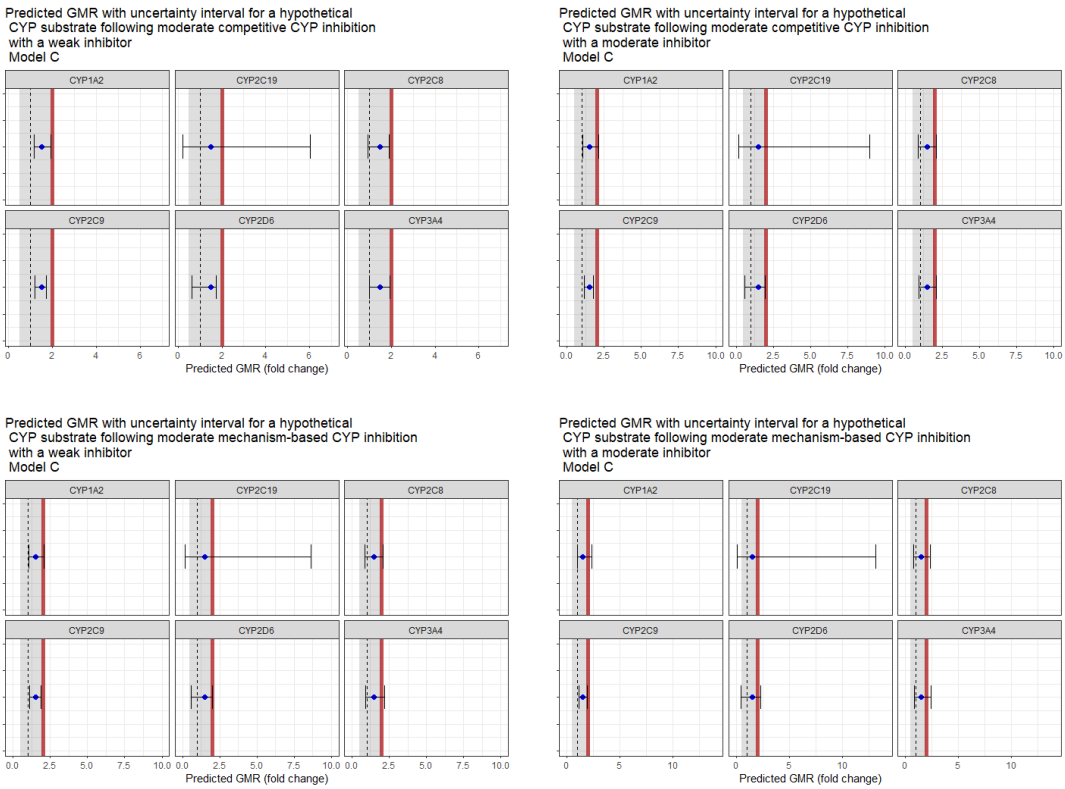

Figure 10: Model C predicted GMR following CYP inhibition for hypothetical CYP substrates in the case of competitive inhibition (top row) or mechanism-based inhibition (bottom row), in the case of a weak inhibitor (left column) or moderate inhibitor (right column). The grey shaded area represents the therapeutic window. The red vertical line indicates its upper limit. The dashed vertical line indicates a predicted GMR without CYP inhibition. The blue dot represents the point estimate of the GMR predicted by the Simcyp® platform. The error bar gives the 90% credibility interval associated with the predicted GMR.

Probability of exceeding a given therapeutic index versus predicted GMR

The above displays were further processed to give the probability (estimated according to model C) of exceeding the therapeutic window when several hypothetical therapeutic windows are considered.

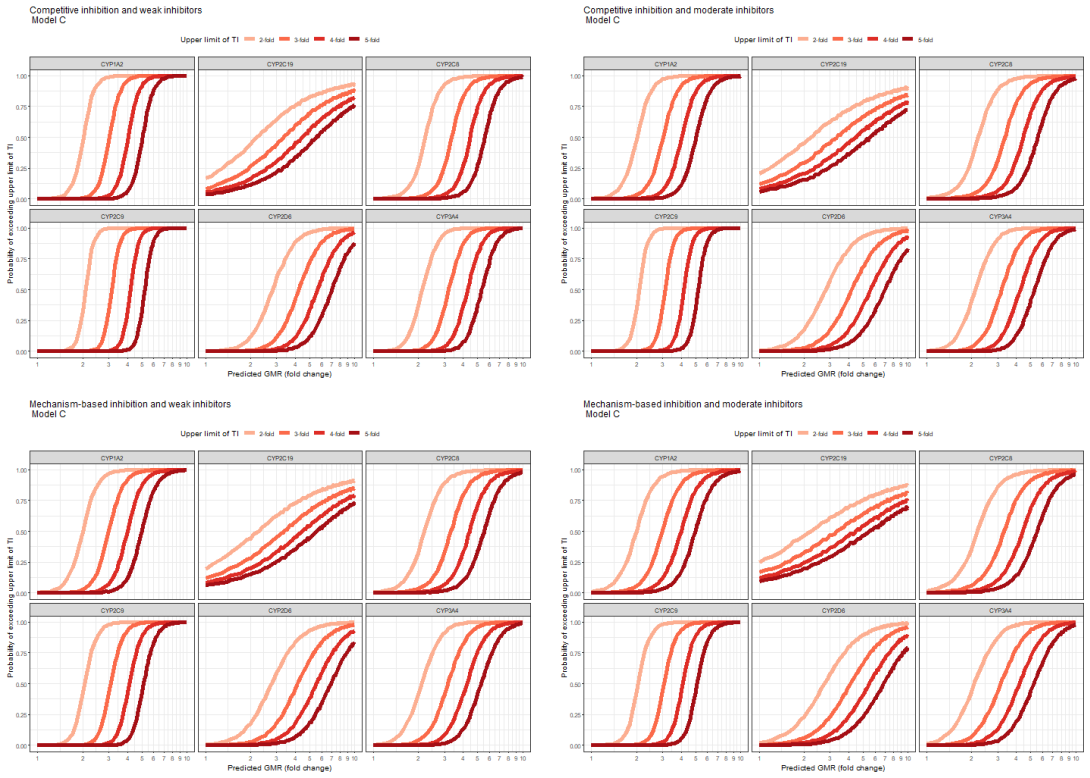

Figure 11: Probability, according to model C, of exceeding upper limit of the therapeutic index *vs*. predicted GMR for competitive inhibition (top row) or mechanism-based inhibition (bottom row), in the case of a weak inhibitor (left column) or moderate inhibitor (right column). The predicted GMRs on the x-axis were predicted using Simcyp®. Results for hypothetical therapeutic index upper limits of 2-, 3-, 4- and 5-fold are displayed.

Maximum predicted GMR for less than 5% risk of exceeding a given therapeutic index

The above analysis can be expanded to determine the predicted GMR that would lead to a 5% probability of exceeding the upper limit of a therapeutic window. The threshold of 5% is a preliminary proposal of EMA.

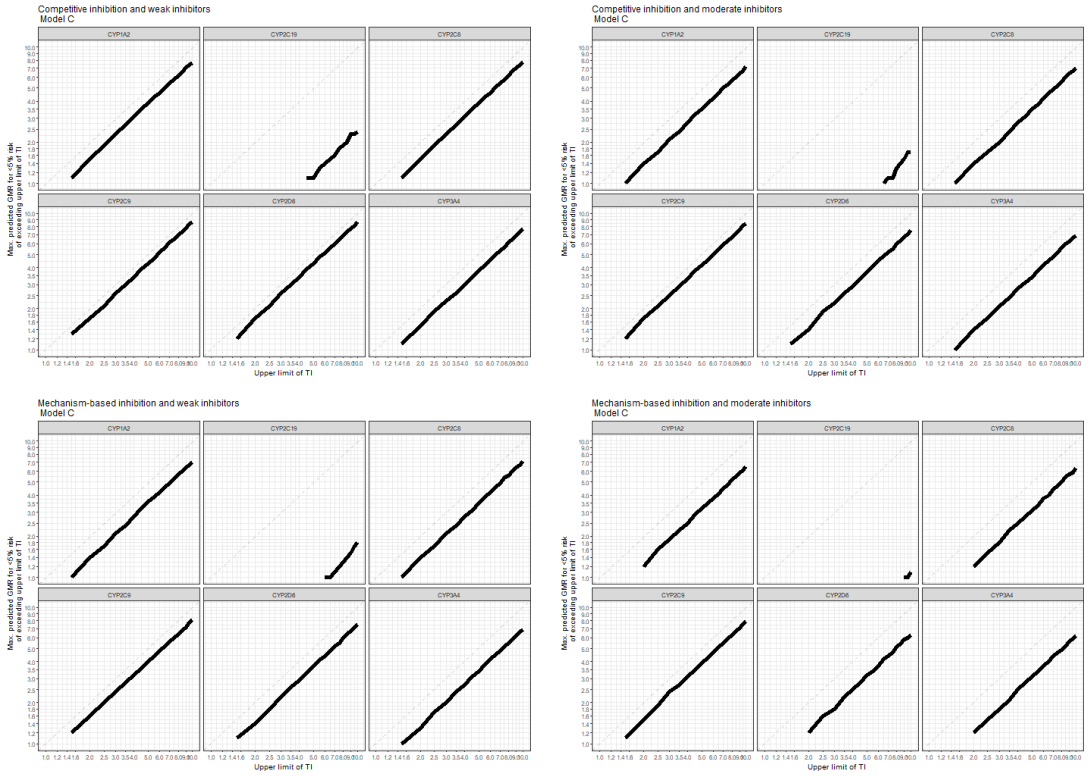

Figure 12: Maximum predicted GMR for <5% risk of exceeding the two-fold upper limit of a therapeutic index for competitive inhibition (top row) or mechanism-based inhibition (bottom row), in the case of a weak inhibitor (left column) or moderate inhibitor (right column). The x-axis shows the maximum GMR predicted by Simcyp® with <5% risk of exceeding the upper limit of therapeutic index given the uncertainty obtained from model C analysis.

Discussion

About the statistical models

Our initial model (model A) was a Bayesian meta-regression model with a nested random effect. It was fitted to data subsets obtained by stratification on individual CYPs. The EMA extended it to include inhibition type (competitive vs. mechanism-based) as a covariate effect and vectorized CYP-specific GMR bias and between-study variability (model B). Model B can therefore integrate all the data and is conceptually preferable to model A. Model B also uses at once all the information available to identify bias in between-subject variance and the posterior estimates obtained are stable without using an *ad hoc* prior, like we had to do for model A. As requested by the EMA we extended further their model to include the strength of the studied inhibitors (limited to classes "weak" and "moderate") as yet another covariate (model C). Model C does not consider strong inhibitors and uses a reduced data base. In model B and C effects of inhibition type and strength are the same for all CYPs. Overall, model B appears to be the most

general. Model C offers the possibility to assess performance when no fitting has occurred, but its performance is also degraded by a reduced database size. In any case, the three models give a coherent picture of the good performance of the Simcyp® Simulator.

About the results

Models' fit to the data

The EMA suggested to assess goodness of fit by comparing the marginal distribution of the data and the predictive distribution of Simcyp® PBPK model predictions, corrected for biases by the above meta-analysis models. A baseline comparison (without bias corrections) can also be made. Figure 4 shows that even without bias correction, DDI GMR predictions are quite close to published observations, but that between-subject variability is more clearly underestimated by predictions. The various meta-analysis models do, obviously, reduce those discrepancies but the plots do not show a clear advantage of a model B over model A. Model C seems less satisfying for GMR correction, but it is unclear whether this is due to worse fit rather than to the reduced database. We note also that the observation *vs*. prediction plots we were showing in our previous responses to EMA were un-marginalized equivalent to the plots in Figure 4.

In any case, since all the meta-analysis models studied are built to minimize the distance between observations and predictions, simple comparisons of them should not be expected to be very discriminant and do not consider issues of parameter numbers or database size. Model choice should also be guided by consideration of model structure, sensitivity to priors, as mentioned in the section above, but also consideration of the plausibility and coherence of the posterior distributions obtained (examined in the following sections).

Biases in GMR predictions

All models pretty much agree that GMR prediction bias is minimal (a few percent) for CYP2C8, CYP2C9, and CYP3A4; of the order of -10% (that is, over-prediction of GMRs by the Simcyp® Simulator) for CYP1A2 and CYP2C19; and of the order of -15% to -30% (the latter according to model C) for CYP2D6. The divergence of model C for CYP2D6 is probably due to the removal of the strong inhibitor data and issues related to lack of recording genotypes/phenotypes in the clinical studies. Models B and C indicate that bias is reduced by a few percents in the case of mechanism-based inhibition and that is coherent with the results of model A after stratification by type of inhibition (see Certara/Simcyp. EMA Initial Qualification Procedure – Second List of Issues Simcyp Simulator – Response Document. June 26, 2024). Model C further points to only a small effect of inhibitor strength on GMR bias.

Prediction imprecision

In any of the models considered, Simcyp®'s imprecision, linked to residual variance after GMR bias adjustment, cannot be separated from between-study variability. We would need a strong informative prior on true between-study variability, or specific data, to disentangle it from predictions' imprecision. For now, the total between-study variance after mean GMR bias correction can be used as an upper bound on Simcyp imprecision.

We found upper bounds on precision CVs ranging from ranging from 12% to 41%, depending on the CYP. So, overall, mechanism-based inhibition predictions seem less biased, but potentially less precise than predictions in the case of competitive inhibition. Model C identifies a further increase by a factor 1.7 in the case of moderate inhibitors compared to weak inhibitors.

Bias in between-subject variability predictions

The mean bias in between-subject variability is about a factor 2 according to model B. Models A and C estimates are higher but less reliable because of the reduced data bases that those models use.

Posterior predictions for use in drug development

The EMA has appropriately proposed a set of predictive posterior simulations that drug developers could use to gauge the confidence they could place in the Simcyp® Simulator V19 R1 predictions. Our own simulations, using model C, do confirm the usefulness of these proposals and their legitimacy for inclusion in a QO.

We think that the EMA's remark that the expected variability attributed to between-subject variability is not reflected in the proposals is well taken care of already, given that posterior samples of GMR bias ***and*** between-study variability were used to obtain those predictive estimates.

Conclusions

It can be concluded that:

- The model extensions and refinements proposed by the EMA work well and confirm that the Simcyp® Simulator DDI GMR predictions' bias is globally acceptable and well-understood.
- Simcyp imprecision in GMR predictions is *at most* 12% to 40%.

- The mean bias in between-subject variability is about a factor 2. This is an area where model-based assessment is strongly affected by data quality.

References

1. R Development Core Team. R: A Language and Environment for Statistical Computing, http://www.R-project.org. Vienna, Austria: R Foundation for Statistical Computing; 2013.

Appendix

Simulation script 1 (data preprocessing and model v7)

```
// R-code for fitting Bayesian meta-regression model
// SAWP model in Stan, modified to include inhibitor strength ("m202.stan")

// ------------------
data {
  int<lower=0> N;
  int<lower=0> K;
  int<lower=0> L;
  int NSub[N];
  int CYP[N];
  int TIN[N]; // indicator of inhibition type;     1 = CI;   2 = MBI
  int SIN[N]; // indicator of inhibitor strength; 1 = weak; 2 = moderate
  vector[N] Obs_BSV_of_ratio;
  vector[N] Obs_ratio;
  vector[N] BSV_simcyp_pred;
  vector[N] simcyp_ratio;
}

// ------------------
parameters {
  vector[K]          ratio_bias_CYP;
  real               ratio_bias_TIN;
  real               ratio_bias_SIN;
  real               mean_BSV_bias;
  real<lower=0>      log_sd_BSV_bias;
  vector<lower=0>[K] extra_var_CYP;
  real<lower=0>      extra_var_TIN;
  real<lower=0>      extra_var_SIN;
  vector[N]          BSV_bias_std;
}

// ------------------
transformed parameters {
  vector[N] mean_bias;
  vector[N] log_corrected_ratio;
  vector[N] BSV_bias;
  vector[N] true_BSV_of_ratio;
  vector[N] shape;
  vector[N] scale;
  vector[N] rate;
  vector[N] extra_var_of_ratio;
  vector[N] total_var_of_ratio;
  vector[N] total_SD_of_ratio;

  for (n in 1:N){
    // uses indicator variables.
    mean_bias[n] = ratio_bias_CYP[CYP[n]] + (TIN[n] - 1) * ratio_bias_TIN +
```

```
                                          (SIN[n] - 1) * ratio_bias_SIN;

    log_corrected_ratio[n] = log(simcyp_ratio[n]) + mean_bias[n];

    BSV_bias[n] = mean_BSV_bias + BSV_bias_std[n] * log_sd_BSV_bias;

    true_BSV_of_ratio[n] = exp(log(BSV_simcyp_pred[n]) + BSV_bias[n]);

    shape[n] = NSub[n]/2; // alpha in STAN
    scale[n] = true_BSV_of_ratio[n]/shape[n];
    rate[n]  = 1/scale[n]; // beta in STAN

    // uses indicator variables.
    extra_var_of_ratio[n] = extra_var_CYP[CYP[n]] *
                            exp((TIN[n] - 1) * extra_var_TIN +
                            (SIN[n] - 1) * extra_var_SIN);

    total_var_of_ratio[n] = true_BSV_of_ratio[n] / NSub[n] +
                            extra_var_of_ratio[n];

    total_SD_of_ratio[n] = sqrt(total_var_of_ratio[n]);
  }
}

// ------------------
model {
  // adaptive priors (describing the random effects distributions)
  BSV_bias_std ~ normal(0, 1);

  // hyper-priors (describing the priors for the parameters for the
  // random effects distributions)
  mean_BSV_bias   ~ normal(0, 1); // log scale; we use U(0,50)
  log_sd_BSV_bias ~ cauchy(0, 1); // log scale; we used half truncated normal

  // regular priors
  ratio_bias_CYP ~ normal(0, 1);
  ratio_bias_TIN ~ normal(0, 1);
  ratio_bias_SIN ~ normal(0, 1);

  extra_var_CYP  ~ cauchy(0, 1);
  extra_var_TIN  ~ normal(0, 1);
  extra_var_SIN  ~ normal(0, 1);

  // likelihood
  Obs_BSV_of_ratio ~ gamma(shape, rate);
  Obs_ratio        ~ lognormal(log_corrected_ratio, total_SD_of_ratio);
}

// ------------------
generated quantities {

  array[N] real predBSV = gamma_rng(shape, rate);
  array[N] real predR   = lognormal_rng(log_corrected_ratio,
                                        total_SD_of_ratio);

//  array[N] real predBSV2 = gamma_rng(shape*1000, 1000/scale);
//  array[N] real predR2   = lognormal_rng(log_corrected_ratio,
//                                         sqrt(extra_var_of_ratio));
//  vector[K] sim_log_corrected_ratio = log(sim_ratio) + mean_bias;
//  array[K] real simR   = lognormal_rng(sim_log_corrected_ratio,
//                                       sqrt(extra_var_of_ratio));

  vector[N] log_lik;
  for (n in 1:N) {
    log_lik[n] = gamma_lpdf(Obs_BSV_of_ratio[n] | shape[n], rate[n]) +
                 lognormal_lpdf(Obs_ratio[n] | log_corrected_ratio[n],
                                               total_SD_of_ratio[n]);
  }
```

```
}
// End.
```