

# Propagation of population PK and PD information using a Bayesian approach: dealing with non-exchangeability

Aris Dokoumetzidis and Leon Aarons  
School of Pharmacy and Pharmaceutical Sciences, University of Manchester, UK.

## Aim

The aim of this work is to implement a conservative prior that safeguards against population non-exchangeability of prior and data likelihood, in the framework of population PK / PD analysis, incorporating multi-level hierarchical modelling.

## Methods

Various datasets were simulated using an Emax dose-response model.

$$E = E_0 + \frac{E_{max} \cdot D}{ED50 + D}$$

where  $D$  is the drug dose. These datasets were denoted as prior and test datasets according to their role in the exercise. Two exercises were considered:

### Exercise (i): parametric priors

To investigate the use of parametric priors with the multilevel model, 5 prior studies and a test study were simulated incorporating three sources of variability, intra-individual, on the responses, inter-individual on the parameters and inter-study variability on the mean values of the parameters. The aim was to investigate whether the fitting of the test study with parametric priors that came out of the prior studies gave the same results as the fitting with non-informative priors of all 6 studies (prior and test) combined. The procedure followed was the same as the one followed in (1) and is outlined schematically in Fig.1, namely the these steps were followed:

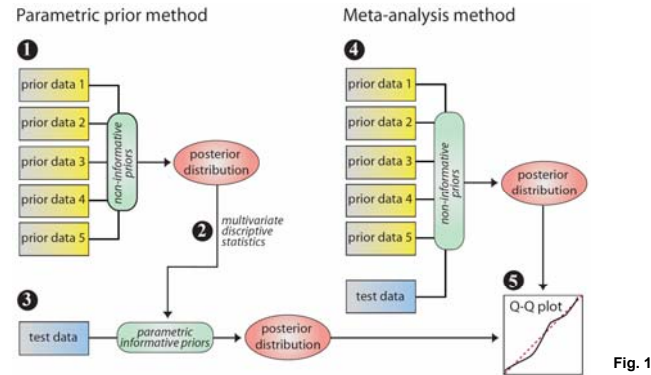


Fig. 1

1 A 4-stage hierarchical model was fitted to the 5 prior studies, with WinBUGS, where on top of the two standard stages of the mixed effects model that describe the intra- and inter-individual variability of each study, the population means of each study were considered to follow a Student-t distribution with the inter-study means and the inter-study variance-covariance matrix as hyperparameters. These hyperparameters, together with those for the intra-individual variability were considered to follow the distributions of the 4th hierarchical stage, the priors. The 4-stage hierarchical model was formulated as follows (2):

structural model  $\log y_{ij} = \log f(\theta_{ij}, x_{ij}) + \epsilon_{ij}$     intra-individual  $\epsilon_{ij} \sim N(0, \tau_i^{-1})$     inter-individual  $\theta_{ij} \sim N(\mu, \Sigma)$     inter-study  $\theta_i \sim St_i(\mu, \Omega)$     priors  $\tau_i \sim \Gamma(a_i, b_i)$   
 $\mu \sim N(\mathbf{m}, \mathbf{P}^{-1})$      $\mathbf{P} = \text{var}^{-1}(\mathbf{h}_{\text{post}})$      $\mathbf{h}_{\text{post}} = E(\Sigma_{\text{post}}^{-1})$   
 $\Sigma^{-1} \sim W_i(S_i, \nu_i)$      $\nu_i = \frac{2 \cdot E^2(\Sigma_{\text{post}}^{-1})}{\text{var}(\Sigma_{\text{post}}^{-1})}$      $S_i = \frac{E(\Sigma_{\text{post}}^{-1})}{\nu_i}$   
 $\Omega^{-1} \sim W_i(S_2, \nu_2)$      $\nu_2 = \frac{2 \cdot E^2(\Omega_{\text{post}}^{-1})}{\text{var}(\Omega_{\text{post}}^{-1})}$      $S_2 = \frac{E(\Omega_{\text{post}}^{-1})}{\nu_2}$

2 From the posteriors of step 1, parametric distributions were constructed for the mean inter-study parameters, the inter-study variability, and the inter-individual variability. The latter was considered common for all studies. The parameters of each distribution were:

$\mu \sim N(\mathbf{m}, \mathbf{P}^{-1})$      $\mathbf{m} = E(\mathbf{h}_{\text{post}})$      $\mathbf{P} = \text{var}^{-1}(\mathbf{h}_{\text{post}})$   
 $\Sigma^{-1} \sim W_i(S_i, \nu_i)$      $\nu_i = \frac{2 \cdot E^2(\Sigma_{\text{post}}^{-1})}{\text{var}(\Sigma_{\text{post}}^{-1})}$      $S_i = \frac{E(\Sigma_{\text{post}}^{-1})}{\nu_i}$   
 $\Omega^{-1} \sim W_i(S_2, \nu_2)$      $\nu_2 = \frac{2 \cdot E^2(\Omega_{\text{post}}^{-1})}{\text{var}(\Omega_{\text{post}}^{-1})}$      $S_2 = \frac{E(\Omega_{\text{post}}^{-1})}{\nu_2}$

3 Using these priors, and non-informative priors for the residual error,  $\tau_i$ , namely  $a_i = b_i = 0.001$  for the parameters of the  $\Gamma$  distribution, the 4-stage model was fitted to the 6th study data, estimating the study-specific population parameters, and also updating the inter-study population parameters.

4 Meta-analysis of all 6 studies combined, was carried out using the 4-stage model with non-informative priors, as in step 1.

5 The posterior distributions of steps 3 and 4 were compared directly. This was done by plotting quantile-quantile plots (Q-Q plots) of the posterior WinBUGS chains that came out of the informative prior fittings of the test study (step 3) and the meta-analysis fitting of all 6 studies combined (step 4).

### Exercise (ii): Performance of the multilevel model.

To assess the average performance of the multilevel hierarchical model, and compare it with the standard mixed effects model, 100 sets of 6 studies, with the same characteristics as in exercise (i), were simulated. The 5 first studies of each set were considered to be the prior while the 6th study was considered to be the test study. Fittings with WinBUGS were performed for the 6 simulated studies of each of the 100 sets, in three different ways:

- Using a 4-stage hierarchical model exactly as the meta-analysis step (step 4) of exercise (i). We focused our attention on the study-specific parameter of the test study,  $\theta_6$ .
- Using a standard mixed-effects, 3-stage model, where prior and test studies were pooled and the inter-study variability was ignored. We focused our attention on the population parameter  $\mu$ .
- The results were also compared with the fittings of the test study by itself, using the 3-stage model, with non-informative priors (no-prior).

### Exercise (iii): extreme cases

Also, 2 extreme cases were considered using on the exercise (ii) datasets. These were:

- To assess the performance of the 4-stage model when very few prior studies are available, the extreme case of fitting the 4-stage model to a single study was attempted with non-informative priors. This is the case where the fitting is dominated by the data and the prior is not very informative or non-informative at all. This result was compared with the 3-stage fitting of the same dataset, with a non-informative prior.
- A common danger when fitting with priors is the opposite case than above, when the prior is much more informative than the data and may overwhelm them, such that the posterior are the same with the prior distributions. To assess the performance of the 4-stage model on that issue, the extreme case of fitting the 4-stage model to a dataset with an infinitely informative prior, was considered. That is, in the 4th hierarchical stage, the values for the inter-study variability were fixed (to their theoretical values) and did not have any uncertainty at all. The results were compared with the 3-stage model fitting with non-informative priors.

## Results and discussion

### Parametric priors

Q-Q plots were used to compare the posterior distributions of the parametric prior and the meta-analysis fittings which are presented in Fig. 2. In most of the Q-Q plots, the curves lie on the identity line indicating that the two distributions in comparison are identical. The only exceptions are in the inter-individual variability parameters,  $\Omega$ . So, the conclusion is that parametric priors may be used to describe the information of the datasets that they came from, and therefore can be used alternatively to the combined fitting of prior and test datasets, what we refer to as meta-analysis. This approach has several advantages. The computational times are much smaller, as part of the information is included in the parametric prior which does not add to the computational effort. Also the parametric prior method works without having the original prior data available (1).

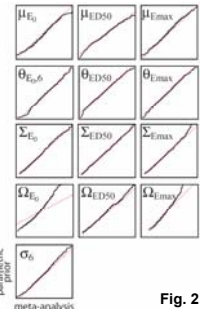


Fig. 2

### Performance

In exercise (ii), we assessed the average performance of the 4-stage model. For each test study the biases of the estimated parameter values, as a percentage, relative to the real parameter values of that study, were reported, for each of the three fitting methods. Also the uncertainty, i.e. the SD of the MCMC chain for each parameter for the 4-stage and 3-stage fitting methods, as percentage relative to the no-prior method, were reported as well. So, for the 100 sets, the average biases and relative uncertainties were estimated for each fitting method and the results appear in Table I. The average biases for all methods are small. This is because the inter-study variability was considered symmetrical and although biases do exist, on average they cancel out. However, the bias SDs (absolute bias) for the 4-stage method are small which, suggests that these estimates are unbiased, and this is reasonable given the fact that they are study-specific estimates. Further a drop in the uncertainty with the 4-stage method compared to no-prior method, is present. This means that the prior is actually working, increasing the precision of the estimates. On the other hand, the 3-stage method exhibits elevated levels of bias. This is because the estimated parameters refer to the overall population of all the studies, test and prior pooled. Also it is clear, that the decrease of uncertainty, is higher than the 4-stage method and this is because a stronger assumption for the prior is adopted, namely that all the studies have the same population values, instead of the more relaxed assumption of the 4-stage method, of the study population parameters being distributed and not identical. Overall, the 4-stage model seems to be a good compromise between bias and added precision, and therefore is a formal way of formulating a conservative prior which is in-between the unbiased, but less precise fitting with no prior, and the more precise, but biased, 3-stage method.

Table I

	no-prior		4-stage		3-stage		uncertainty relative to no-prior	uncertainty relative to no-prior
	average bias	bias SD	average bias	bias SD	average bias	bias SD		
$E_0$	0.20%	3.40%	0.20%	3.30%	89.10%	0.90%	14.50%	54.00%
$ED50$	0.20%	3.90%	0.20%	2.80%	74.50%	1.10%	5.80%	39.50%
$E_{max}$	1.10%	5.90%	0.60%	4.40%	73.80%	0.40%	10.10%	40.50%

### Extreme cases

(a) The 4-stage model was fitted to a single study, which corresponds to the case of a non-informative prior or equivalently when the data overwhelm the prior. The results (Table II) indicated that the presence of an extra (abundant) hierarchical level did not have any negative influence, and instead the inter-study variability posterior distributions turned out to be non-informative, while the rest of the posterior distributions were identical to the ones obtained with the 3-stage model.

(b) The prior was much more informative than the data and in fact the prior values were fixed. However, results (Table III) indicated that because they do not interact with the study specific parameters directly but only through the inter-study variability, the study-specific parameters were not biased. This means that the prior did not overwhelm the data and in fact gave similar results to the fitting with no priors, but with slightly less uncertainty. The inter-study priors interact with the study-specific parameters through the inter-individual variability, which in fact determines how influential the priors are. If the inter-study variability is very large, then the priors do not influence the study-specific parameters, and if it is too small then the model collapses to the 3-stage model.

Table II

node	4-stage		3-stage	
	mean	SD	mean	SD
$\mu_{E0}$	1.547	5.093	-	-
$\mu_{ED50}$	5.658	5.125	-	-
$\mu_{Emax}$	2.34	4.346	-	-
$\Sigma_{E0}$	70.23	2140	-	-
$\Sigma_{ED50}$	61.68	2020	-	-
$\Sigma_{Emax}$	55.43	2164	-	-
$\theta_{E0,6}$	1.541	0.04982	1.541	0.04989
$\theta_{ED50,6}$	5.644	0.1765	5.641	0.1766
$\theta_{Emax,6}$	2.311	0.09142	2.31	0.09202
$\Omega_{E0}$	0.04185	0.0168	0.04169	0.01701
$\Omega_{ED50}$	0.1621	0.1424	0.1681	0.1445
$\Omega_{Emax}$	0.06147	0.04118	0.06156	0.03959
$\sigma_6$	0.09228	0.00945	0.09229	0.0094

Table III

node	Simulated value	3-stage		4-stage	
		mean	SD	mean	SD
$\theta_{E0,6}$	1.522	1.541	0.04978	1.548	0.04845
$\theta_{ED50,6}$	5.975	5.648	0.1765	5.795	0.1349
$\theta_{Emax,6}$	2.435	2.313	0.09238	2.355	0.07608
$\Omega_{E0}$	0.04	0.04151	0.01677	0.04184	0.01682
$\Omega_{ED50}$	0.04	0.1514	0.1286	0.1756	0.1443
$\Omega_{Emax}$	0.04	0.05809	0.03579	0.05811	0.03805
$\sigma_6$	0.1	0.09267	0.00945	0.09244	0.00948

### Bayesian individualization

All the above exercises share a common feature. Due to the fact that studies which do not come from exchangeable populations were analyzed together, an extra level in the hierarchical model was added, and the priors were applied to this hierarchical stage. However the parameters of interest are one level below the prior; they are the study specific parameters of the test study and not the inter-study parameters. This in fact is equivalent to Bayesian individualization, where estimates for a population of individuals are available and are used as priors in order to estimate a specific individual's parameters. So, given a prior that came out of a population of studies, we estimate the study specific, i.e. individual parameters of the test study. This approach implements a conservative prior, appropriate when prior and test populations are not exchangeable, much like the way an individual is not exactly representative of the population in mixed effects modelling.

## Conclusion

In the present work we addressed the issue of population non-exchangeability of prior and test studies in the case where the bias between them is random. An extra hierarchical stage was introduced which accounted for the inter-study variability of the prior and test studies. With this multilevel hierarchical model, parametric priors were used successfully, giving the same posterior distributions, as meta-analysis of the test study and the studies that the parametric prior came from. By applying the prior one hierarchical level above the level of the parameters of interest, a more conservative prior was implemented, compared to applying the prior directly on the parameters of interest, because the prior influences them only indirectly. Further, the multilevel model has a few desirable properties, offering a safeguard against bias from the prior and also avoiding the danger of the data being overwhelmed by a strong prior.

The multilevel approach, is equivalent to Bayesian individualization, in the sense that given a population of studies as a prior, the parameters of another specific study are estimated, much like the parameters of a specific subject are estimated using a prior of population estimates.

## References

- Aris Dokoumetzidis and L. Aarons. Propagation of population pharmacokinetic information using a Bayesian approach: comparison with meta-analysis. *J. Pharmacokin. Pharmacodyn.* (accepted 2005).
- J. Wakefield and N. Rahman. The combination of population pharmacokinetics studies. *Biometrics* 56:263-270 (2000).