1. **IPRED versus DV**

I have identified why our durations in severe neutropenia appear to be different. It is because I summarised simulations using IPRED, while you have been summarising with DV.

In the Shiny app, I’ve used PRED for the line to represent the “average/typical” value and used IPREDs for the 95% prediction intervals. This is because any decisions made for an individual should ideally be based on the underlying process and in the absence of the observed variable’s (i.e., ANC) random noise/measurement error or random intra-individual variability.

The ANC fluctuations you see when you simulate DV are due to the residual error model and simulated ERR terms that are added to IPRED to make DV. Therefore, any fluctuations you see are random and more due to measurement error rather than individual’s underlying ANC profile. For example, there is no physiological or pharmacological process (i.e. in the code) that explains why individual ANC profiles would bounce up and down near the nadir. This is shown by the smooth curve in this time-period when you simulate IPRED or population predicted (PRED) from your model. An unwanted consequence of using DV is that the number of ANC samples taken will affect the calculated occurrence of Grade 4 (e.g. if we took 100 samples near the nadir, the chances are that one of them will be below 0.5 due to RUV noise alone is pretty high). I talked to Richard and David about this, and they would recommend using simulations based on IPRED. There are scenarios where it may be better to simulate DVs to capture the range of profiles likely to be measured include; clinical trial simulation – which this is not – or for a VPC.

1. **PRED versus mean**

I have also identified another reason why our summary plots are different. I have been using PRED for the typical value, while you have been using the mean and CI of simulated DV’s. It is very common for these to give different results, as simulated DV’s are rarely normally distributed.

The PRED ANC profile for the model already provides a nice summary as to the average or “typical” profile for an individual with *x* covariates. It may not be average as in a “mean” or “median” sense – but it is the “most likely average” as estimated by maximum likelihood estimation. As this is the “average” for *x* covariate profile, you do not need to calculate the mean or median of the simulations as simulated inter-individual variability is essentially centred about the PRED. (However, it may be appropriate to calculate the median when the simulation study contains mixed covariate groups as each of these different covariate groups will have a different PRED).

No matter how many simulations are performed, PRED will always be the same as it is not subject to random effects, but 95% prediction intervals based on IPRED will slightly differ as the random number generator may spit out different ETA values every time when given a different seed. The mean, however, is much more dependent on the number of simulations – you might find that if you performed a few studies with varying numbers of simulations and seeds, you will get different values for the mean, but the PRED would be the same every time.

Finally, we commonly we use prediction intervals based on percentiles e.g. (5th, 50th and 95th percentiles) to summarise simulated distributions as this makes no assumption about the underlying distribution. The confidence interval method based on mean and SD assumes that the data is normally distributed and that the mean is the centre of the distribution which is often not the case.

I have made an R script (attached, Melphalan\_Simulation\_Model\_RUV\_Demonstration.R) that illustrates the phenomena.

I talked to Richard and David about this too, and they would recommend using PRED for summarising the “average.”

1. **Methods for calculating duration in severe neutropenia**

We have a difference in this aspect of our models to. As discussed above, the method that I used for calculating the duration in severe neutropenia in the Shiny app is based on the PRED, and the 95% prediction intervals are from the IPRED.

I think you have used a method that assumes that in the 24-hour interval where an individual enters or exits Grade 4 neutropenia, the slope of change in ANC is linear? This method can only provide an approximation and would be quite exhaustive if written into R. There is actually a neat trick in NONMEM that allows you to do this. I’ve attached a NONMEM control stream of your model (PKPD\_Final\_SIM\_JW.ctl), with an extra compartment added that collects the duration each individual spends in Grade 4 neutropenia. You’ll find that it only requires an extra 3 lines of code and uses differential equations.

This only applies to IPRED, so is obviously dependent on accepting the arguments presented in 1 & 2.

100sims_PRED_versus_Mean.pdf

**Number of simulated individuals = 100**

**1000sims_PRED_versus_Mean.pdf**

**Number of simulated individuals = 1000**