



Background and Objectives

Improving the calculations of the weighted residuals has proven to be of high importance; especially if the model is highly nonlinear in the random effects [1,2]. Various new methods for calculating the weighted residuals have been implemented in NONMEM 7 (NM7) [3]. The aim of this project is to investigate these new methods.

Methods

A highly non-linear sigmoidal Emax-model ($\gamma=4.5$) previously used to demonstrate the importance of using CWRES versus WRES (NWRES in NONMEM 7) was used to investigate the different residuals [1]. Emax and EC50 both had between subject variability (BSV) corresponding to ~71% CV. The study design was rich; 200 individuals with 25 observations each.

Five different scenarios were investigated with the true and a misspecified model ($\gamma=1$): 1) Additive residual unexplained variability (RUV), 2) proportional RUV, 3) exponential RUV, 4) exponential BSV on the proportional RUV, 5) Between occasion variability on Emax with an additive RUV.

The residuals investigated were (fig. 1): NWRES (First order (FO) residuals without interaction), WRESI (FO residuals with interaction), CWRES (FO conditional residuals with some interaction), CWRESI (FO conditional residuals with full interaction), ECWRES (Monte Carlo calculated weighted residuals with some interaction), EWRES (Monte Carlo calculated weighted residuals with full interaction) and the NPDE (Normalised Prediction Distribution Errors). The simulation based residuals were calculated with the default number of samples (300) but in some cases a more intense sampling was also investigated (3000 samples). Interaction was always used in the estimation line, however MAXEVAL=0 or EONLY=1 was used to disable any population parameter estimation. All the residuals were calculated for 100 simulated data sets with the true and misspecified model and hypothesis tests for mean 0 (Wilcoxon signed-rank test), variance 1 (Chi-square test) and normality (Kolmogorov-Smirnov test) were calculated (fig. 3).

Different weighted residuals in NM7

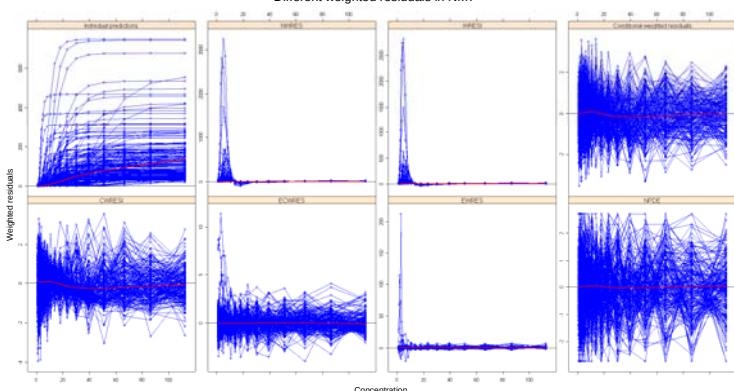


Fig 1. The individual predictions and the different weighted residuals available in NM7 from one data set with the proportional residual error model. FOCEI was used to initiate the residual calculations and 300 samples were used to calculate the simulation based residuals.

Results

The CWRES and NPDE outperform the other residual diagnostics (figs. 1 and 3) in identifying the true model. The NPDE occasionally underperform in the variance test (figs. 2, 3) compared to the CWRES. Furthermore the CWRES was, in general, better than the NPDE when the NM7 default number of samples was used. When more samples were used, either NPDE or CWRES could be better in different situations and different tests (fig. 3).

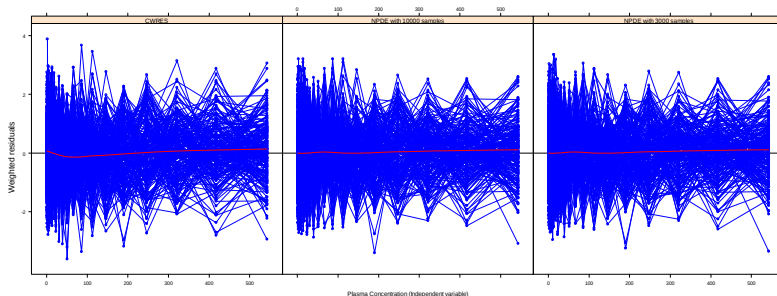


Fig 2. The CWRES and the NPDE calculated with 3000 samples and 10 000 samples showing the weighted residuals versus concentration for one of the data set with model 1). The empirical variance differ between the residuals, i.e. CWRES = 1.012, NPDE_3000 = 1.103, NPDE_10000 = 1.097 which indicates that the NPDEs are sensitive to the chi-square test and the number of samples. However the visual differences are very small.

The other simulation based residuals (ECWRES, EWRES) didn't perform as well as the CWRES and NPDE. CWRESI performed worse than CWRES or NPDE even when there was interaction in the model. As expected the NWRES and WRESI were not performing well in any of the investigated cases. In nearly all cases the misspecified model was correctly identified. Using FOCEI in the estimation line instead of IMP did not change the results to a large extent (except for the CWRES and model 4). Here the percentage of accepted H0 changed to (K-S=35%, Wilcoxon=94%, Chi-square=0%) for the true model.

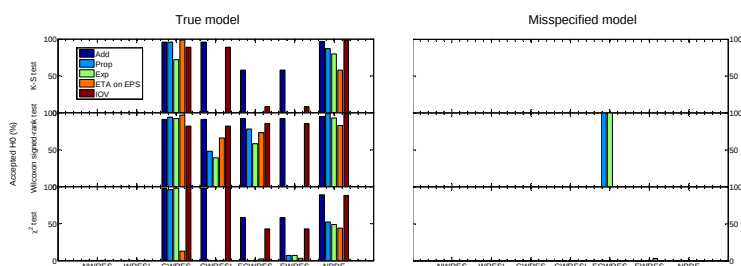


Fig 3. Accepted null hypothesis in percent for the Kolmogorov-Smirnov test, the Wilcoxon signed rank test and the Chi-square test. All tests used a significance level of 5% and tested 100 simulated data sets. The left panel used the same model for simulation of data as for calculation of the residuals while the right panel used a misspecified model when calculating the residuals. Importance sampling with 3000 samples (EONLY) was used as an initial starting point for the residual calculation. The simulation based residuals used 3000 samples for integration.

The CWRES is not exactly the same as in Hooker et al [1], but is identical to the methods implemented in PsN and Xpose. The Hooker implementation evaluates the residual model around individual values=0 while the PsN/Xpose/NM7 implementation around the individual values = individual mode or mean, accounting for some interaction which makes the CWRES method more useful for many types of residual models.

Conclusion

- ✓ NPDE or CWRES give overall the best diagnostic
- ✓ NPDE or CWRES will perform best in different situations
- ✓ Both methods work well with interaction models

References

- [1] Hooker AC, Staats CE, Karlsson MO. Conditional weighted residuals (CWRES): a model diagnostic for the FOCE method. *Pharmaceutical research* 2007; 24: 2187-97.
- [2] Comets E, Brendel K, Mentre F. Computing normalized prediction distribution errors to evaluate nonlinear mixed effects models: the npde add-on package for R. *Computer Methods and Programs in Biometrics* 2008; 90:154-166.
- [3] NONMEM users guide. Introduction to NONMEM 7. Icon Development Solutions (2009)