The likelihood ratio test appears robust to most residual error model misspecifications

H.E. Silber, M.C. Kjellsson, M.O. Karlsson
Department of Pharmacokinetics and Drug Therapy, Division of Pharmaceutical Biosciences, Uppsala University, Sweden

Background and objective

The likelihood ratio test (LR-test) has previously been shown to be sensitive to residual error model misspecifications in that ignoring an existing η-ε interaction when using the first-order conditional method (FOCE) in NONMEM resulted in higher significance levels for the type I error than the nominal level [1].

The objective of this study was to assess the LR-test sensitivity to residual error misspecifications through simulations.

Methods

Data sets containing 250 individuals with six or twelve observations per individual were simulated multiple times (n=1000). The structural model that was used for simulation was a 1-compartment model with a constant infusion at steady state. The inter-individual variability on clearance was set to 30 percent. The different residual error models that were implemented were: (1) autocorrelation (AR), (2) inter-individual variability in the residual error magnitude (ETAonEPS), (3) replication error (L2), (4) time-varying residual error magnitude (TIME), (5) heavy-tailed residual error distribution (HEAVY), (6) inter-occasion variability (IOV) and (7) an η-ε interaction (INTER). The residual error was 30 percent except in the case where the value of ε was varied. Details about the parameter values used for the different scenarios are given in table 1.

The simulated data were analyzed using the correct residual error model as well as a reduced model with an additive residual error (on log-transformed data). Further, the data were analyzed with or without a covariate relationship on clearance. The type I error rate of inclusion of a non-informative covariate was calculated as the number of runs where the drop in the objective function value (OFV) was larger than 3.84 (5-percent level) when the covariate relationship was included in the model. This was done using both the correct and the reduced error model (figure 1). To assess the difference in goodness-of-fit between the correct and the reduced residual error model the difference in OFV was also calculated (figure 1). The study was performed using the FOCE method in NONMEM.

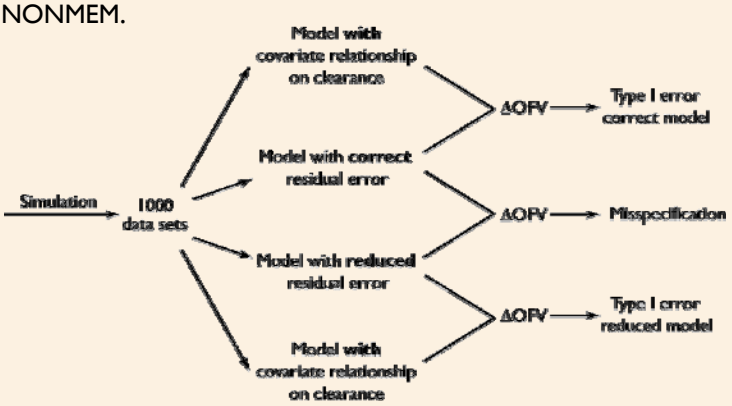

Figure 1. Data were analyzed to assess the type I error rate for the correct and the reduced error models and to estimate the difference in goodness-of-fit (model misspecification) for the different residual error models.

Results

When the reduced error model was used the model misspecifications were pronounced which was indicated by the OFV being on average 206-2269 higher than with the corresponding correct error models. However, the significance levels for the LR-test with the reduced models were still appropriate and similar to those when the correct models were used. The type I error were in all cases between 4.1 and 5.8 percent (figure 2). The only exception was, as expected based on findings in [1], the case where the reduced model ignored an existing η-ε interaction. The type I error rate was then 31.1 percent.

Model	Parameters	
AR	t_{corr} = 4	
L2	correlation = 28%	
ETAonEPS	ω_ε = 30%	
IOV	π = 30%	3 different occations
TIME	ε_{early} = 10% ε_{late} = 50%	2 different magnitudes
HEAVY	ε_{normal} = 30% $\varepsilon_{heavy} = \varepsilon_{normal} \cdot 5$	10% of observations

Table 1. Parameter values for some of the residual error models that were evaluated.

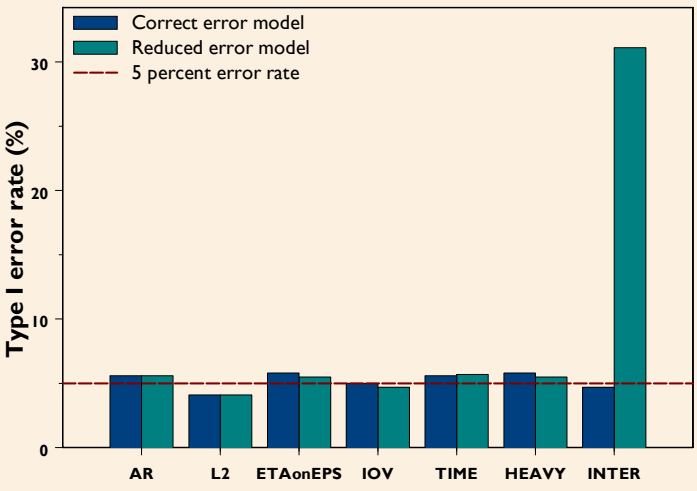

Figure 2. The type I error (%) calculated for the different scenarios using the correct error model (blue) and the reduced error model (green). The red broken line shows the 5-percent error rate.

Conclusion

The LR-test appears robust towards all tested residual error misspecifications but ignoring the η-ε interaction.

Reference:

1. Wahlby et al. *Assessment of type I error rates for the statistical sub-model in NONMEM.* J Pharmacokinet Pharmacodyn, 2002. 29:251-69.