Are Datasets for Nonlinear Mixed-Effects Models Large Enough For a Bootstrap to Provide Reliable Parameter Uncertainty Distributions?

Ronald Niebecker, Mats O. Karlsson

Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden

Background

- Nonparametric bootstrap is a frequently employed method to determine parameter uncertainty distributions.
- Compared to using bootstrap to calculate a single statistic from a vector of observations, application of bootstrap to assess uncertainty in nonlinear-mixed effects (NLME) models is more complex, due to
  - Simultaneous estimation of multiple parameters
  - Hierarchical models with ≥2 levels of random effects
  - Data-driven model development
  - Model misspecification
  - Heterogeneous designs including covariate distributions
- Concerning dataset size, the prerequisites for bootstrap to perform well have rarely been investigated.

Objectives

- Explore whether typical combinations of model complexity and dataset size allow for appropriate behaviour of bootstrap.
- Introduce bootstrap delta objective function value (dOFV) distributions as a method to diagnose whether a bootstrap will not provide appropriate parameter uncertainty distributions.

Methods

For each investigated combination of model and dataset, three dOFV distributions were generated and visually assessed (Fig. 1).

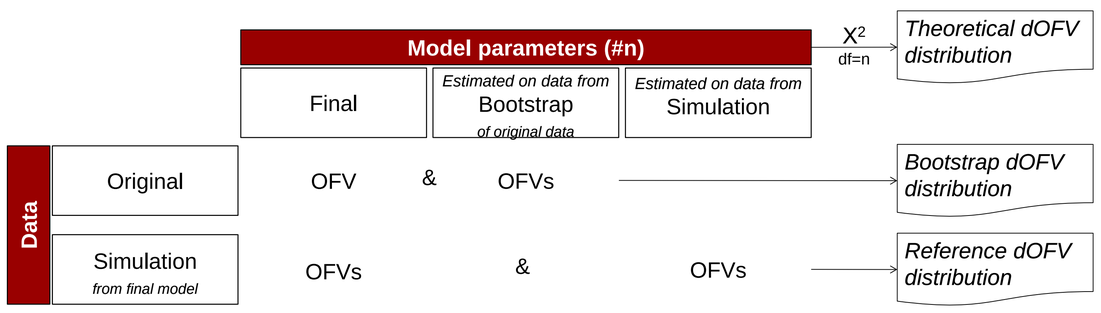

Figure 1. *Generation of dOFV distributions.*

Both real data [1–3] and simulation examples were investigated (Tab. 1).

Table 1: *Model characteristics of real data and simulation examples*

Example	Model	Parameters	Individuals	Observations/ Individual
Moxonidine [1]	Linear 1-compartment model with first-order absorption, lag time	12	74	14
Pefloxacin [2]	Linear 1-compartment model with IV bolus administration	10	74	4.6
Phenobarbital [3]	Linear 1-compartment model with IV bolus administration	7	59	2.6
Simulation I	Linear 1-compartment model with IV bolus administration	6	20–600	2–4
Simulation II	Emax model with baseline	7	20–500	4

Assessment of parameter uncertainty

- Confidence intervals (CIs) determined by bootstrap (1000 samples) and stochastic simulation and reestimation (SSE, 1000 samples) were compared.
- For simulation example II, coverage based on 100 bootstraps each for three different dataset sizes was computed.

The analysis was carried out in NONMEM 7.2 [4] aided by PsN [5].

References

[1] Karlsson et al., J Pharmacokinet Biopharm. 1998;26(2):207–46.

[2] Wählby et al., Br J Clin Pharmacol. 2004;58(4):367–77.

[3] Grasela et al., Dev Pharmacol Ther. 1985;8(6):374–83.

[4] Beal et al., NONMEM user's guides. Icon Development Solutions, Ellicott City, MD, USA; 1989–2009.

[5] Lindbom et al., Comput Methods Programs Biomed. 2005;79(3):241–57.

Results

Real data examples

- For the investigated examples, 27% to 51% of the bootstrap dOFV values exceeded the 95th percentile of the theoretical dOFV distributions (Fig. 2).
- Bootstrap based on simulated datasets of equal size confirmed these findings. For 8-times increased dataset sizes, bootstrap dOFV distributions converged to the theoretical and reference distributions, which were superimposed (Fig. 2).

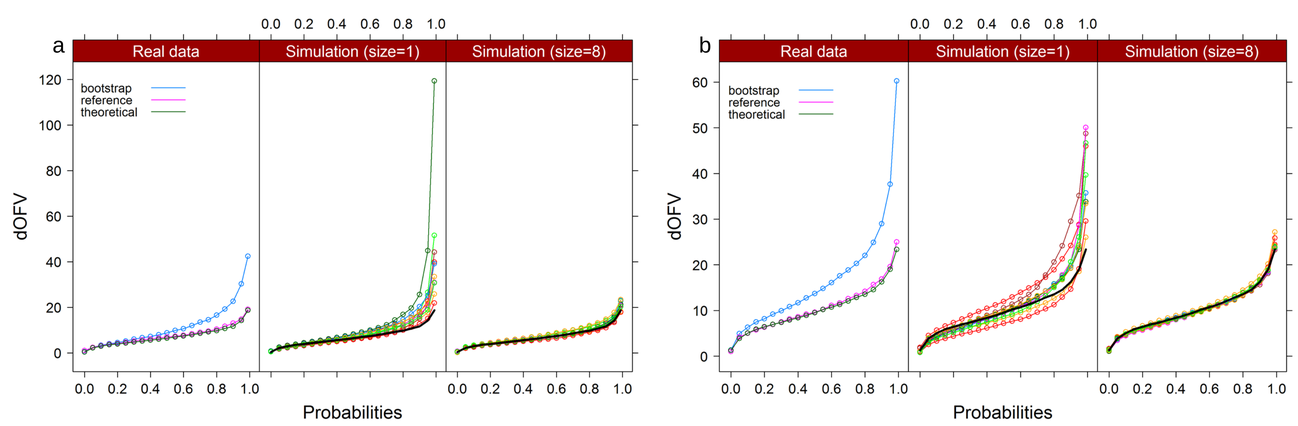

Figure 2a, b: *dOFV distributions for phenobarbital (a) and pefloxacin (b). Left panels contain dOFV distributions from bootstrap based on real data, plus reference and theoretical dOFV distributions. Middle and right panels: bootstrap dOFV distributions based on 10 bootstraps each, for same-sized and 8-times increased simulated datasets; theoretical dOFV distribution superimposed (black).*

Simulation examples

- Similar to the real data examples, bootstrap dOFV distributions converged to the theoretical dOFV distribution for increased datasets (Fig. 3).

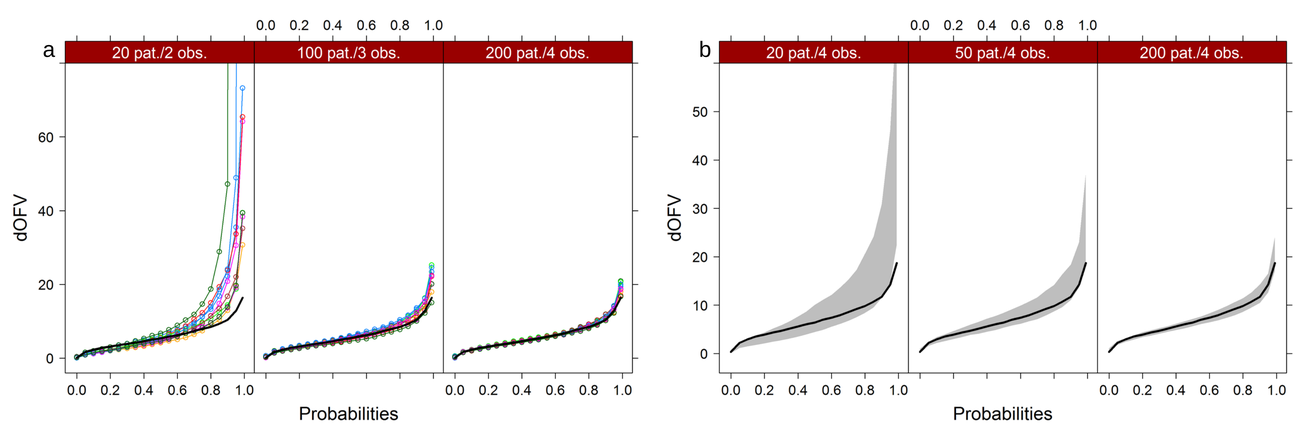

Figure 3a, b: *Bootstrap dOFV distributions for simulation example I (a) and II (b), based on 10 bootstraps (a)/100 bootstraps (b) each, theoretical dOFV distribution superimposed (black).*

- CIs from bootstrap more closely approximated CIs based on SSE especially for random-effects parameters; coverage improved for larger datasets (Fig. 4).

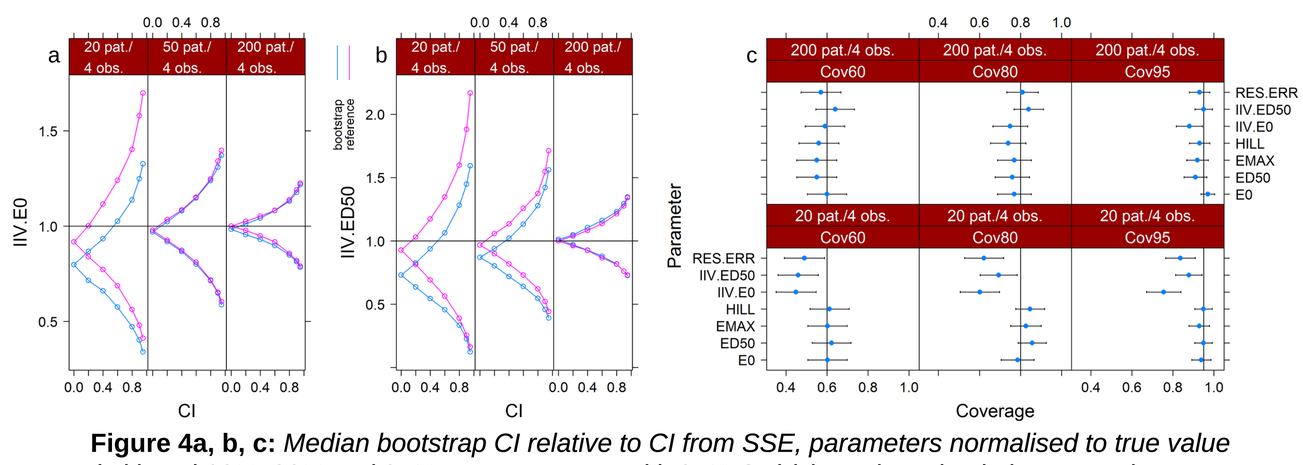

Figure 4a, b, c: *Median bootstrap CI relative to CI from SSE, parameters normalised to true value (a,b) and 60%, 80% and 95% coverage rates with 95% CI (c) based on simulation example II.*

Conclusions

- This analysis showed that with regard to providing uncertainty estimates, bootstrap may be unsuitable already for NLME analyses where datasets would commonly be considered "large enough".
- The bootstrap dOFV distribution provides an easy way to assess if bootstrap results in parameter vectors contradicted by the original data.

Acknowledgement: The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement n° 115156, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. The DDMoRe project is also financially supported by contribution from Academic and SME partners. This work does not necessarily represent the view of all DDMoRe partners.