The bootstrap of Stepwise Covariate Modeling using linear approximations

Ron J. Keizer, Akash Khandelwal, Andrew C. Hooker, Mats O. Karlsson
Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden

Background and Objectives

Stepwise covariate modeling (**scm**) is a tool for automatized building of a covariate model on top of a base structural model [1]. Using a linear approximation to the base model in the scm provides similar results compared to an scm based on the original model, while greatly reducing computation times [2].

Performing a **bootstrap** analysis of the **scm** can provide valuable information, e.g. about the type I error of covariate inclusion, correlations between covariate inclusions, and identification of influential individuals. However, a regular bootstrap-scm may require considerable computation time. A bootstrap of a **linearized** scm can be performed much faster.

Objective: to compare results obtained from the bootstrap-scm using linearized and non-linearized models
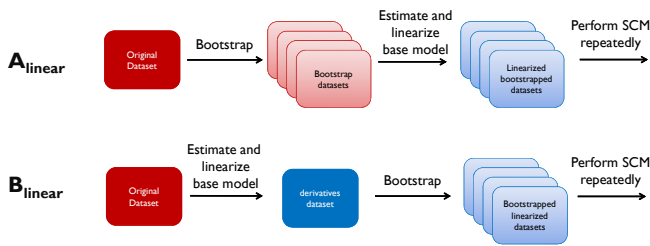
Methods

Data
- Two real PK datasets (moxonidine, pefloxacin). PK model developed earlier.
- Dummy covariates introduced into the datasets based on the randomized original covariates (evaluation of type I error for covariate inclusion)

Bootstrap
Two methods of performing a **linearized bootstrap-scm** were implemented:

For comparison, the bootstrap-scm was also implemented using the non-linearized model and based on the original dataset (**A_non-linear**).

Optimism analysis
Final models obtained in the bootstrap repetitions were refitted on the original dataset and fit compared: to the original model: $dOFV = OFV_{bs,i} - OFV_{orig\_scm}$.

Results

Speed: The linearized bootstrap based on the dataset with derivatives (B) was fastest (about 1 day) but only slightly faster than method A. The non-linearized bootstrap-scm took almost a week to complete. With more complex models, runtime for $A_{non-linear}$ will be much longer, while runtime for methods A_{linear} and B_{linear} will stay similar.

Covariate inclusion %: small differences between methods, especially for intermediately strong covariates.
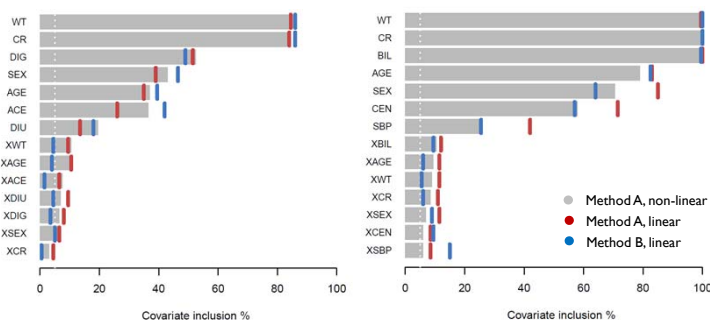

Figure 1. Covariate inclusion % for moxonidine (left) and pefloxacin (right). "X" covariates are randomized (dummy-) covariates.

Conclusions

- Linearization of the model allows the implementation of a bootstrap-scm within a reasonable time-span.
- Results are comparable to a non-linearized bootstrap-scm.
- Several diagnostic plots are available for the bootstrap-scm to aid covariate model construction.

Covariate model sizes: distributions were highly similar between the two linearized methods, and were also similar to the non-linearized bootstrap-scm.
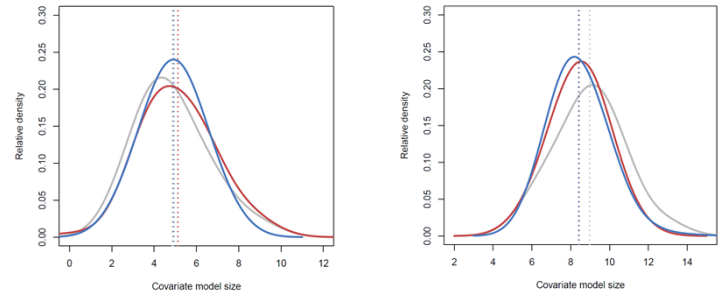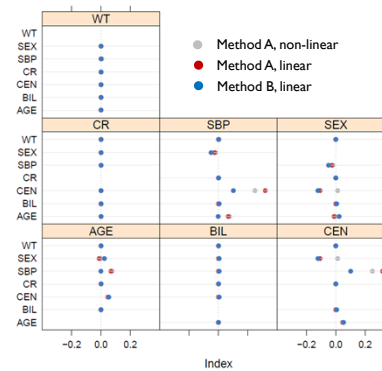

Figure 2. Covariate model size moxonidine (left) and pefloxacin (right).


Correlations: This diagnostic plot can help determining if covariate inclusion is correlated between covariates. Results differ somewhat between methods. In general, results from method A_{linear} were closer to $A_{nonlinear}$ than results from B_{linear} were.

Figure 3. Correlation inclusion rate (pefloxacin).
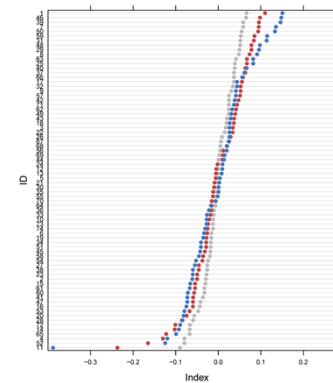
Influential individuals: For intermediately strong covariates (here shown for SEX), both linearized bootstrap methods seem a bit more sensitive to influential individuals than the non-linearized method, especially method B_{linear}.

Figure 4. Covariate inclusion rates individuals (pefloxacin)

- Method A, non-linear
- Method A, linear
- Method B, linear
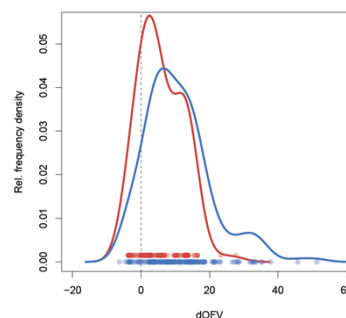
Optimism: Interestingly, about 10% of the final full covariate models obtained in the bootstrap procedures showed a lower OFV than the final model in the original scm, when the final (non-linearized) covariate model was refitted on the original dataset.

Figure 5. Optimism analysis
- Moxonidine
- Pefloxacin

References

[1] Karlsson & Jonsson. PAGE 7 (1998) Abstr 678 [www.page-meeting.org/?abstract=678]
[2] Khandelwal et al. PAGE 19 (2010) Abstr 1925 [www.page-meeting.org/?abstract=1925]