Faster methods for case deletion diagnostics: dOFV and linearized dOFV

UPPSALA UNIVERSITET

Rikard Nordgren, Sebastian Ueckert, Svetlana Freiberga and Mats O. Karlsson
Pharmacometrics group, Department of Pharmaceutical Biosciences, Uppsala University

Background

Case deletion diagnostics (cdd) is a method for finding individuals that are especially influential on a model estimation. The standard way of performing a cdd for non-linear mixed effect models is to exclude one individual at a time and reestimate the model. This procedure has been automated in PsN [1]. Typically the cook score and the covariance ratio have been used as metrics for the individual influence on the parameter estimates [2]. Both these metrics require the covariance matrix of the parameter estimates which might be difficult to obtain. More importantly, the usefulness of the cdd is diminished by the fact that it is relatively slow as it requires one full estimation per individual.

The Cook score is defined as

$$\sqrt{(P_i - P_{orig})^T \operatorname{Cov}(P_{orig})^{-1}(P_i - P_{orig})}$$

where P_{orig} and P_i are the estimated parameter vectors for the original run and the run with individual i removed respectively.

Method

We propose a new metric for the cdd, delta-OFV (dOFV), to assess the influence of an individual on the quality of model fit to all other individuals. This metric does not rely on the covariance matrix. Furthermore, the switch to the dOFV metric also allows to use linearized models which can substantially reduce runtime.

The dOFV metric was calculated for each case deleted run separately as

$$dOFV = OFV_{all} - iOFV_k - OFV_k$$

where OFV_{all} is the OFV of the full run with all individuals included, $iOFV_k$ is the individual OFV of the k:th individual in the full run and OFV_k is the OFV of the run with the k:th individual removed.

The dOFV based cdd was performed both using the NLMEM as well as using a linearization approximation[3][4] of the model and compared to Cook score based cdd. The runtimes of these runs were also compared.

All comparisons of the different metrics and runtime were done for 19 different pharmacometric models using PsN 4.7.16 and NONMEM 7.4.2. Models showing a big (more than 3) difference in OFV with the linearized model where excluded from the analysis leaving 14 models.

Results

The graphic comparison between dOFV and Cook score can be seen in figure 1. The relation is slightly non-linear and the Pearson (linear) correlation coefficient ρ ranged between 0.40 and 0.97 for the tested models.

There were no results for Lopinavir due to numerical problems with obtaining the covariance matrix inverse needed for the Cook score.

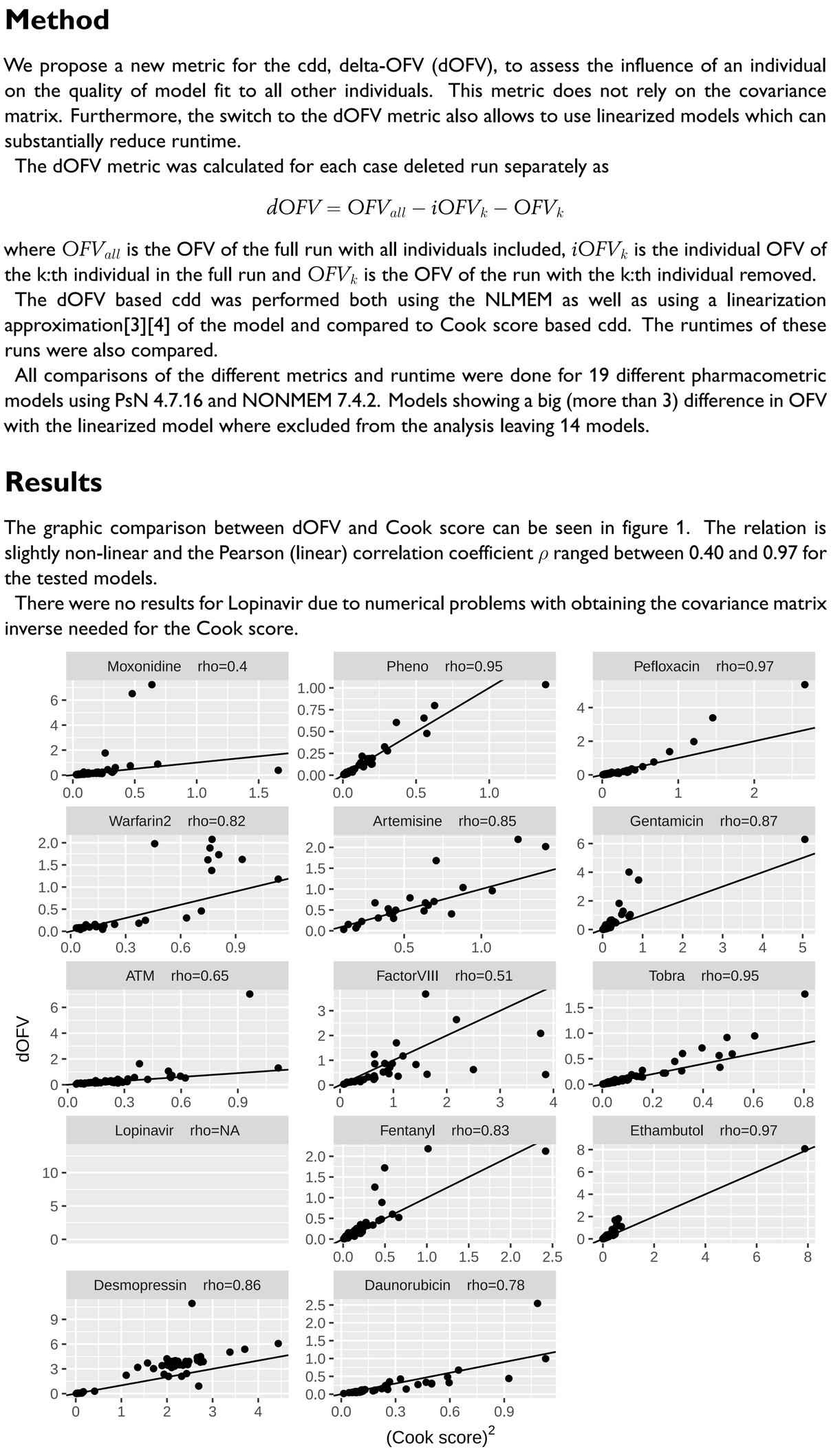

Figure 1: Correlation between Cook score and dOFV

The comparsion between dOFV and linearized dOFV can be seen in figure 2. The correlation coefficient is here above 0.9 for all models but Gentamicin (0.62) and Desmopressin (0.52).

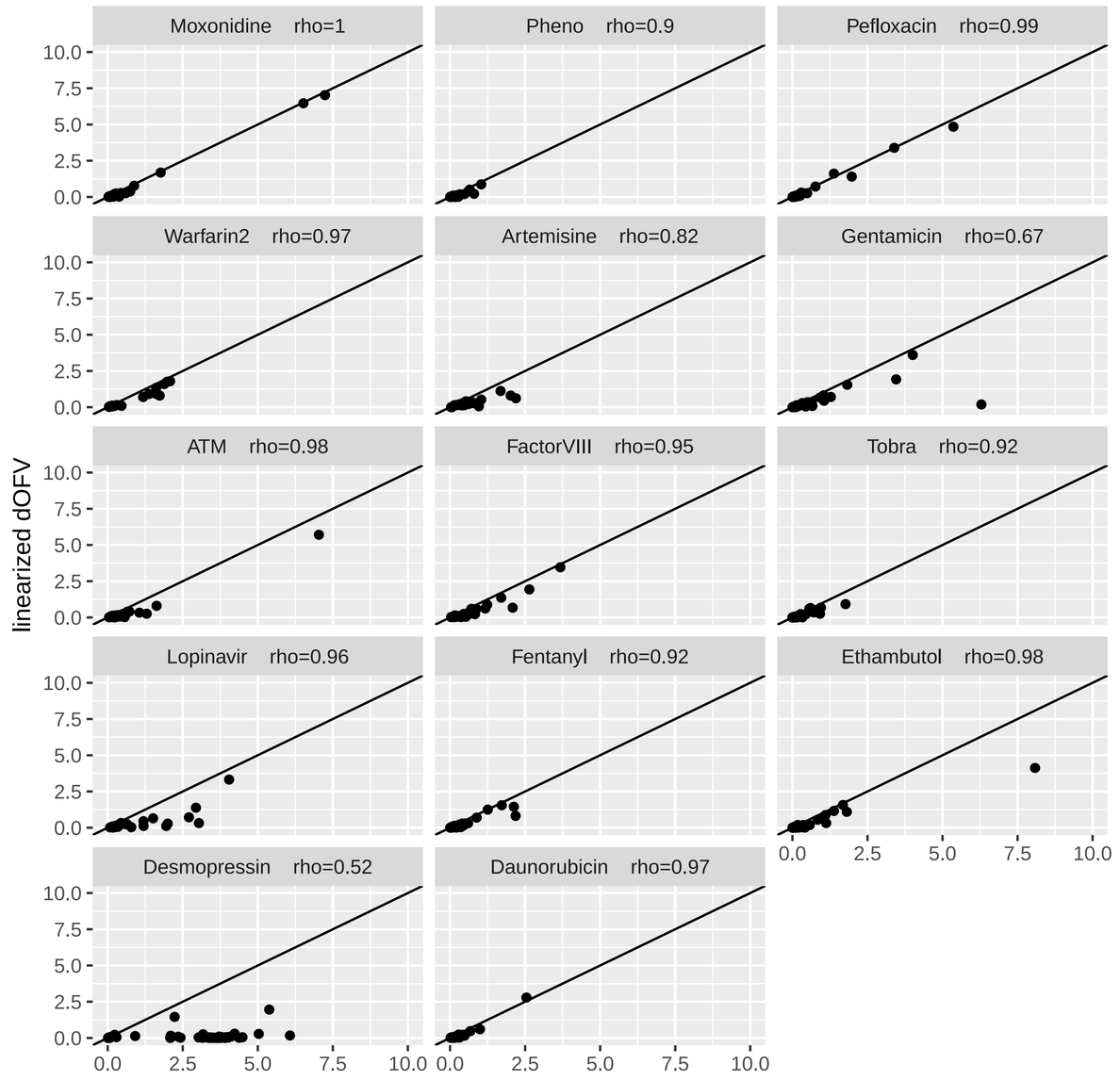

Figure 2: Correlation between dOFV and linearized dOFV

If we define an influential individual to be one with a higher dOFV than 3.84 the linearization does not display any false influentials and misses 4 out of 9 except for the desmopressin case where the original model had 16 influential individuals and the linearized only 1.

A runtime comparison (figure 3) between the three methods show a clear gain in using the proposed methods. The linearization has its greatest advantage for slow running models.

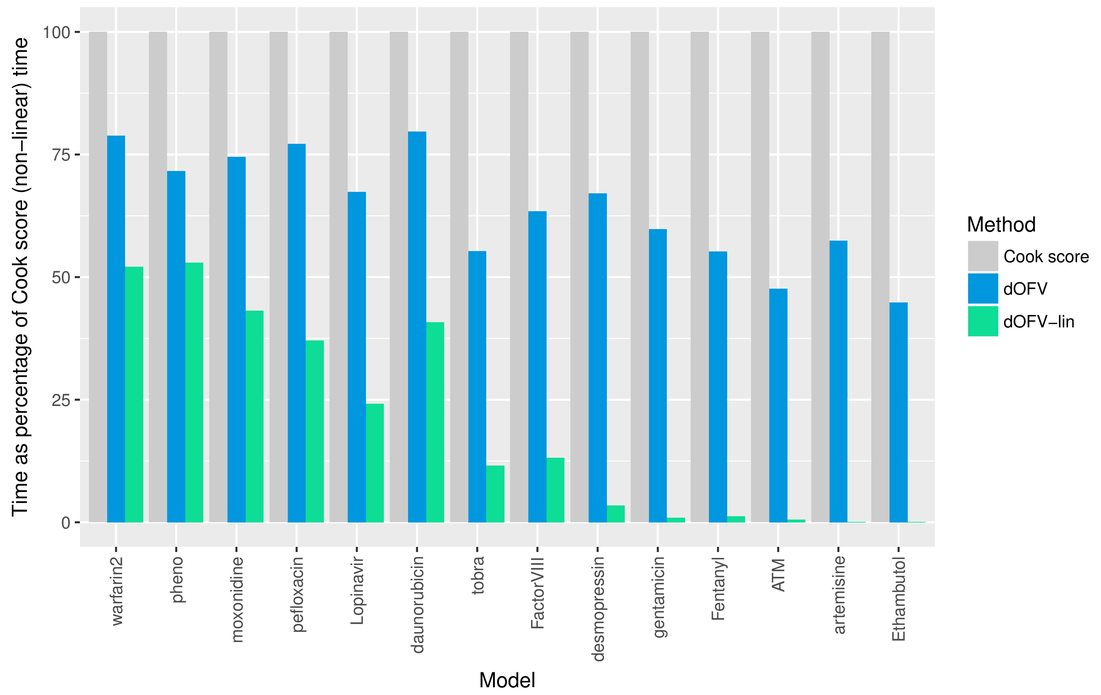

Figure 3: Comparison of percentage of run time between the three metrics ordered from shortest to longest actual runtime

Conclusion

Having a case deletion diagnostic method that does not rely on the covariance step would both increase the stability and reduce the runtime as well as making linearization possible. Our experiments indicate that the dOFV could be used as an alternative to the tradional cook score based metric. Linearization also largely kept the properties of the dOFV except for a few outlying models.

Further investigation could include an investigation of the influence of the number of individuals on the accuracy of the linearized cdd, a closer look on cutoffs or other definitions of influential individuals and ways of robustifying the methods.

[1] L. Lindbom, P. Pihlgren, and E. N. Jonsson. PsN-toolkit–a collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM. *Comput Methods Programs Biomed. 79(3):241-57*, 2005.

[2] Cook R. D. Detection of influential observation in linear regression. *Technometrics 19 (1977) 15-18*, 1977.

[3] A. Khandelwal, K. Harling, E. N. Jonsson, A. C. Hooker, and M. O. Karlsson. A fast method for testing covariates in population pk/pd models. *AAPS J. 2011 Sep;13(3):464-72*, 2011.

[4] E. M. Svensson and M. O. Karlsson. Use of a linearization approximation facilitating stochastic model building. *J PKPD. 2014, 41:153-158*, 2014.