

Time to get intimate with your tools

Mats O Karlsson and Lars Lindbom

Dept of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden

The development of pharmacometric models depends on an understanding of the system to be described, the sources of variability and the error generation mechanisms. However, to arrive at adequate final models we also rely on estimation methods that can determine parameters of the model with sufficiently low bias and acceptable precision. Further, for the model building process we need to have procedures that efficiently and robustly can help bring a reasonable starting model to an adequate final model. In the model building and model evaluation, we need to have informative numerical and graphical diagnostics. For the purpose of this presentation, the tools we use include: estimation methods (FO, FOCE, Laplacian, NPAG etc), (semi)-automated procedures for building components of the model (e.g. covariate modeling strategies), numerical diagnostics (e.g. χ^2 OFV, SEs, Eigenvalues) and graphical diagnostics (using observations, predictions, residuals, simulations, covariates, parameters, etc). Most proposed, and used, tools have merit and are useful under many circumstances, but all have flaws and limitations. It therefore appears that we can't afford to abandon every tool that is not adequate in all situations.

A tool that have shown flaws under some circumstances, may in other lead to more efficient model building or prevent erroneous models to be accepted. On the other hand, to use a tool that truly is flawed for its purpose in a particular situation may result in inefficient model building and the acceptance of an inappropriate model. In between is a situation where modelers are aware of the potential shortcomings/strengths of a tool, but are not sure if those are applicable in the particular situation at hand. Although this is a relatively informed situation, it is still associated with inefficient decision making and potential communication difficulties between modeling stakeholders (drug developers/regulatory authorities, manuscript writers/reviewers).

Thus, we need more information about when a certain tool is appropriate and informative and when it is not. Despite the wide use of the tools in the pharmacometrics area, relatively few systematic studies have been performed in terms of evaluating these. Most studied of the modeling tools are probably the estimation methods, where the performance of several methods have been evaluated with respect to the bias and precision of the parameter estimates. However, such studies are still limited and if systematic, generally focused on simple models. For the few cases where more complex models have been studied, the investigation usually concerns single model/data sets and thus knowledge about performance is anecdotal. For procedures like semi-automated covariate model building, many approaches have been suggested, but thorough evaluations under different realistic conditions are virtually absent. Some evaluation of numerical diagnostics has been performed, but similar to studies of estimation methods, only a few simple scenarios have been studied.

Evaluations of graphical diagnostics involve particular difficulties. As these are intended for a qualified, but subjective, assessment of whether or not they are indicating an appropriate model or not, an objective evaluation in the same manner as for other tools is not possible. However, it is to some extent possible to assess their usefulness in diagnosing model misspecification/appropriateness by graphical questionnaires to modelers. The outcome of such a survey will be presented.

To promote efficient and informed model building with high probability of coherent judgment between different modeling stakeholders, several types of studies would be of value. Theoretical studies of behavior of different tools would be desirable, but most tools require systematic simulation or resampling studies under "realistic" and varied conditions to provide improved understanding of strengths and shortcomings of different procedures. Further, as all possible conditions hardly can be studied; availability of procedures to more easily assess a particular tool for a particular model/data set would be desirable. Additionally, development of new tools valuable across a wider range of conditions is warranted. Lastly, and the main purpose of this presentation, the complexity of choosing tools ought to be widely recognized.

