

Sequential parametric - nonparametric modeling: advantages, disadvantages and usefulness

Radojka M. Savic, Mats O. Karlsson

Division of Pharmacokinetics and Drug Therapy, Faculty of Pharmacy, Uppsala University, Uppsala, Sweden

Objectives: In recent years, nonparametric methods are increasingly used tools for random effects distribution estimation in non-linear mixed effects models. These methods are powerful to estimate underlying distribution shapes as they do not rely on distributional assumptions like the more widely used parametric methods (1-4). The use of nonparametric methods is associated with certain drawbacks, including increased computation time, lack of imprecision measurement, impossibility to estimate residual variability, difficulties to communicate final results with non-pharmacometrics audience etc, which limits their usage. To address some of the issues, nonparametric methods are utilizing outcome of preceding parametric analyses. One of the uses of the parametric method in a nonparametric analysis is in utilizing its variance-covariance matrix for defining the parameter space. With increasing number of random effects (n) and subjects (N), nonparametric estimation procedure rapidly becomes a complex problem as it requires estimation of $((n+1)*N)-1$ parameters. Rather than to search for such a large number of parameters in an infinite space, it is desirable to define the search space a priori. Boundaries can be set up arbitrary; however variance – covariance matrix from preceding parametric analysis could be utilized to get a good idea of parameter space. Nowadays, beside the information about variance-covariance matrix, information about residual error magnitude and individual parameter estimates are also utilized for different parts of nonparametric model building. In this work we summarize the advantages, disadvantages and usefulness of such a combined approach.

Methods:

a) Residual error magnitude In nonparametric analysis, residual variability has to be specified a priori. This term was, for a long time, identified with an assay error (Approach I), which is generally not an appropriate procedure as it ignores other components of residual variability, such as model misspecification, errors in study conduct, changes in individual biology over time etc. Thus, it is preferable to use the residual variability estimate from a preceding parametric analysis in subsequent nonparametric estimation as this is probably a better guess of residual error magnitude (Approach II). A newly available nonparametric method in software NONMEM allows for an estimation of nonparametric distribution of residual error magnitude which has the advantage to other available approaches of not fixing the error magnitude (Approach III). This approach may utilize a prior estimate of parametric residual variability distribution (4). A Monte Carlo simulation study was performed to compare different approaches for handling residual error magnitude. Data were simulated from one compartment IV bolus model and analyzed with the nonparametric method available in NONMEM VI using three above mentioned approaches. Parameter distributions were evaluated at different percentiles and relative estimation error was computed for each method.

b) Individual parameter estimates The points of support for the nonparametric parameter distribution estimation may be obtained from the prior POSTHOC (Empirical Bayes Estimates (EBE)) parameter estimates generated using data and a distribution defined by parameters from a preceding parametric analysis (default method). This approach has been implemented in the latest nonparametric method in NONMEM and it is probably the most radical change from other nonparametric methods. It has shown good estimation properties and it has been associated with several advantages (3). However, in certain circumstances, EBEs are not always providing a good range of points of support as they suffer from shrinkage phenomenon (5). When and if so, a recently developed “extended grid method” capable of providing nonparametric grid with better range of support points is suggested to be used (6). Again, this method utilizes preceding parametric analysis to generate the new set of points of support either via parametric simulations or via EBE estimation from parametric model using inflated variances. The performance of this method was evaluated and compared to the default method via simulations and subsequent analysis of pharmacokinetic data using an intravenous linear one-compartment model with various distribution shapes for the parameters CL and V and with various numbers of observations per subjects.

Results: Specifying residual variability based on an assay error was rather poor approach resulting with imprecise parameter estimates and inflated parameter distributions. A substantial improvement in parameter estimates was seen when Approach II and III were used, with Approach III showing slightly better results.

As expected, EBEs performed poorly as points of support when data were very sparse due to high shrinkage extent and imprecise residual error magnitude estimates. However, the parameter distribution estimated at the enriched grid showed a good agreement with the true parameter distributions examined via QQ plots, regardless the method used to enrich the nonparametric grid. This was also confirmed by inspection of cumulative probability density functions.

Conclusions: Both parts of the parametric analysis used in subsequent nonparametric analysis appeared to be useful, resulting in more precise parameter estimates and shorter CPU time for analysis. The latter which is solely associated with the usage of EBEs as support points, is of importance because long analysis time is one of the major limiting step for nonparametric methods usage. In certain circumstances, individual parameter estimates are poor starting points for the nonparametric estimation, but the enriched grid method successfully addresses this problem. A framework which allows for a subsequent parametric – nonparametric analysis with all above mentioned methods incorporated in it appears desirable for future efficient nonparametric analysis in non-linear mixed effects models.

References:

- [1] Mallet A., 1986. A maximum likelihood method for random coefficient regression models. *Biometrika* 73:3, 645-656
- [2] Schumitzky A., 1991. Nonparametric EM algorithms for estimating prior distributions. *Applied Mathematics and Computation*. 45:2, 143-157
- [3] Bustad A. *et al*, 2006. Parametric and nonparametric population methods: their comparative performance in analysing a clinical dataset and two Monte Carlo simulation studies. *Clin Pharmacokinet* 45 (4): 365-383
- [4] Savic RM. *et al*, 2006. Evaluation of the nonparametric method in NONMEM VI β . PAGE 15 (2006) Abstr 937 [www.page-meeting.org/?abstract=937]
- [5] Savic RM *et al*, 2006 (Un)informativeness of Empirical Bayes Estimate-Based Diagnostics, *The AAPS Journal*; Vol. 8, No. S2, Abstract T3360 (2006)
- [6] Savic RM *et al*, 2007 Evaluation of extended grid methods for estimation using nonparametric distribution. *The AAPS Journal* Vol. 9, No. S2, Abstract W4536 (2007)