



Model Averaging Across Parametric and Non-Parametric Methods As an Aid for Clinical Decision Making

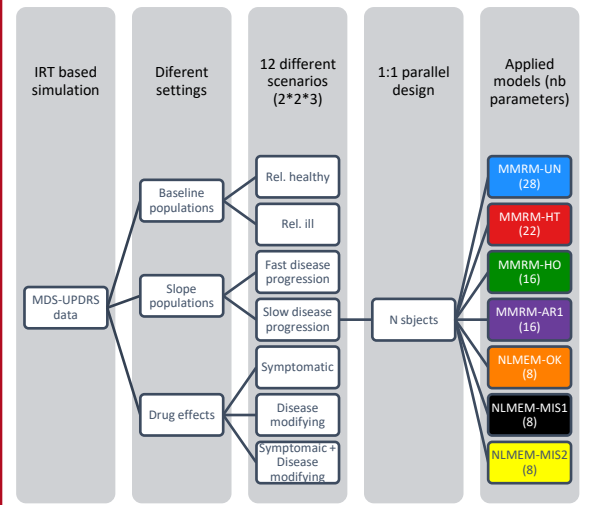
Mohamed Gewily, Gustaf J. Wellhagen, Mats O. Karlsson
Department of Pharmacy, Uppsala University, Uppsala, Sweden

Objectives

Parametric and non-parametric models are often used to describe longitudinal data. On one side, the parametric Non-linear Mixed Effects Model (NLMEM) are built progressively upwards, and usually require external validation and have a considerable number of assumptions. On the other hand, the non-parametric Mixed Models for Repeated Measures (MMRM) rely on modelling the means and standard deviations of observations at each sampling time to make individual predictions, requiring less assumptions.

In this study, we investigated how model averaging across NLMEM and MMRM could benefit predictions of treatment effects in terms of reducing bias and improving accuracy

Methods



A previously published Item Response Theory (IRT) model (1) was used to simulate MDS-UPDRS motor total score data for 42 months, across 4 visits with 100 simulations for 12 different scenarios.

Seven models were applied : MMRM with unconstrained residual error correlation structure (MMRM-UN), with heterogenous Toeplitz (MMRM-HT), with homogeneous Toeplitz (MMRM-HO), with AR1 (MMRM-AR1), a correctly specified continuous variable model in terms of drug effect (NLMEM-OK), and finally two mis-specified continuous variable models in terms of drug effect and vice versa (NLMEM-MIS1&2), while no drug effect parameters were estimated in MMRM by definition.

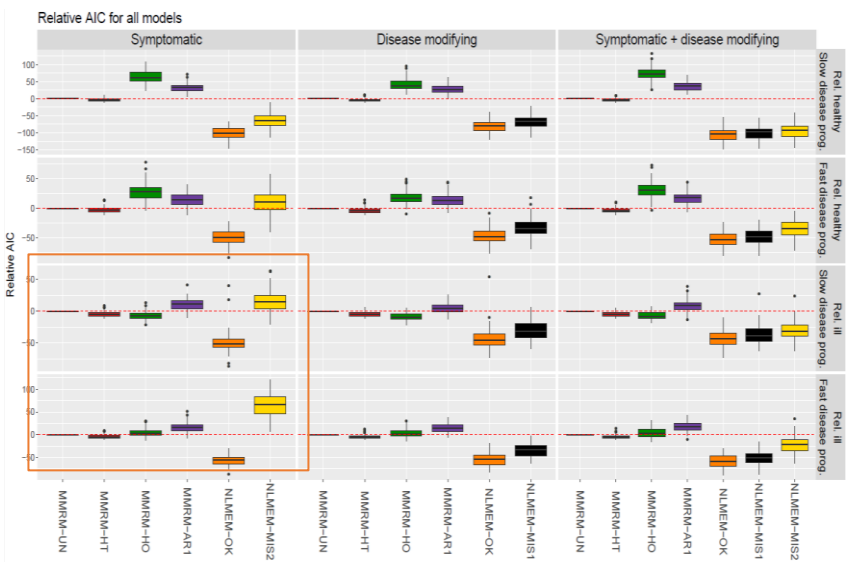
To account for the different number of parameters, the model fits were compared using the Akaike Information Criterion (AIC). For all fitted models, a predicted treatment effect was calculated through a simulation which predicted the difference in the treatment effect at the last time point between the treatment arm and the placebo arm.

Model averaging was carried out using weighted AIC (2).

Conclusions

- Model averaging including both NLMEM and MMRM for the analysis of Randomized Clinical Trials (RCTs) has the possibility to provide: (i) high precision in estimated drug effects through NLMEM when such model(s) is an adequate representation of data, (ii) robustness through MMRM(s) when NLMEM is not adequate
- The MMRM used for this averaging, while standard, were not the best fit for the data. Further improvements of this model averaging strategy may come from: (i) exploration of other types of MMRM, (ii) inclusion of NLMEM that have a higher flexibility than the very parsimonious models used here, (iii) joint inclusion of MMRM and NLMEM, models with a very wide difference in number of parameters.

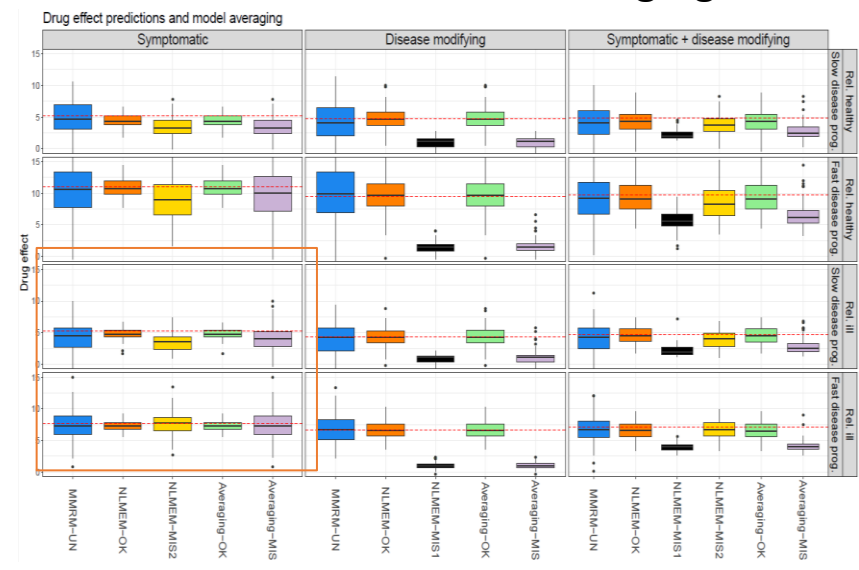
Model fit results



NLMEM had a better data fit compared to MMRM across all 12 scenarios. The AIC penalty favored NLMEM. Nonetheless, NLMEM advantage was also present in terms of the Objective Function Value (OFV). It is also clear that in the case of MMRM, the AIC penalty negates the advantage of the increasingly flexible error correlation structures, resulting in nearly similar AIC values for all tested MMRM.

The misspecified NLMEM models had a better data fit relative to MMRM except for 2 scenarios where the baseline population was relatively ill and the drug effect was symptomatic (highlighted).

Model averaging results



When NLMEM-OK was part of the model averaging scope, it dominated the weighted averaging. When excluded from the model averaging scope, it was either NLMEM-MIS, MMRM-UN that dominated the weighted average

The weight domination of NLMEM-OK proves that when NLMEM is a good representation of the data, including MMRM in the averaging will not negatively influence the drug effect predictions. In contrast, when NLMEM fails to describe the data adequately, including MMRM in the model averaging could result in better predictions in terms of bias and accuracy as in the highlighted scenarios.

References

- Buatois S, Retout S, Frey N, Ueckert S. Item Response Theory as an Efficient Tool to Describe a Heterogeneous Clinical Rating Scale in De Novo Idiopathic Parkinson's Disease Patients. Pharm Res. oct 2017;34(10):2109-18
- Aoki Y, Röshammar D, Hamrén B, Hooker AC. Model selection and averaging of nonlinear mixed-effect models for robust phase III dose selection. J Pharmacokinet Pharmacodyn. déc 2017;44(6):581-97. .