

UPPSALA

UNIVERSITET

Sample size calculations in multiple sclerosis using pharmacometrics methodology: comparison of a composite score continuous modeling and Item Response Theory approach

Ana Kalezic¹, Radojka Savic², Alain Munafo³, Elodie L. Plan¹, Mats O. Karlsson¹

¹ Dept of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden ² Dept of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, USA ³ Modelling and Simulation, Merck Serono S.A., Switzerland

Objectives

To demonstrate the application of pharmacometric methods for power/sample size calculation based on Expanded Disability Status Score (EDSS) [1], a widely used measure of disability in multiple sclerosis (MS), as efficacy endpoint.

Background

• Calculation of a sample size, one of key steps in the planning of a clinical trial, is aimed at appropriately powering the study on the basis of primary endpoint. The calculation typically assumes a given effect size and variance of the response variable.

Results

Power calculation

total

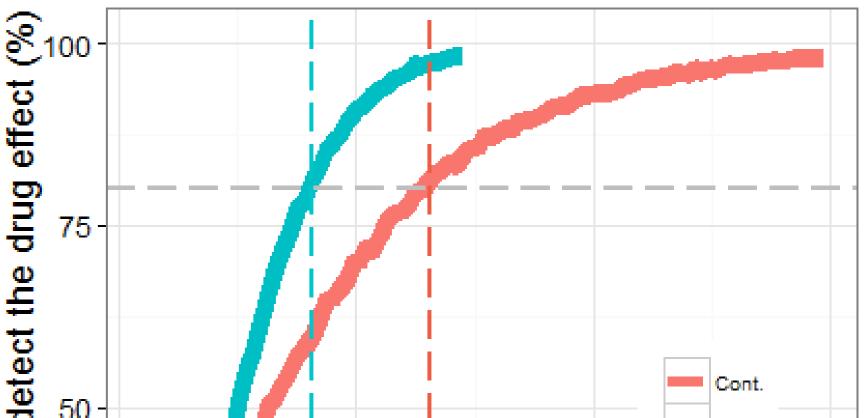
continuous model.

versus

• Application of the IRT modelling approach allows a reduction in sample size of 40% compared with a continuous composite model, to achieve 80% power to detect a dual (offset and disease modifying) drug effect (Fig.1)

number

of



- Clinical trials in multiple sclerosis therapeutic area are particularly long due to the variable and slowly progressive nature of the disease.
- Phase III trials are often conducted for over two years and frequently include more than a thousand patients. Therefore, increased efficiency would be valuable.

Methods

<u>Study design</u>

- EDSS is a 20 point scale (0=normal neurological exam and 10=death due to MS), based on the neurological assessment of 7 functional systems and of walking ability.
- Study design was a 96-week Phase III clinical study with relapsing-remitting MS, where patients received active treatment or placebo, and EDSS assessment was conducted before the start of the treatment (at baseline) and every 12 weeks during the duration of the study.

Clinical trial simulations

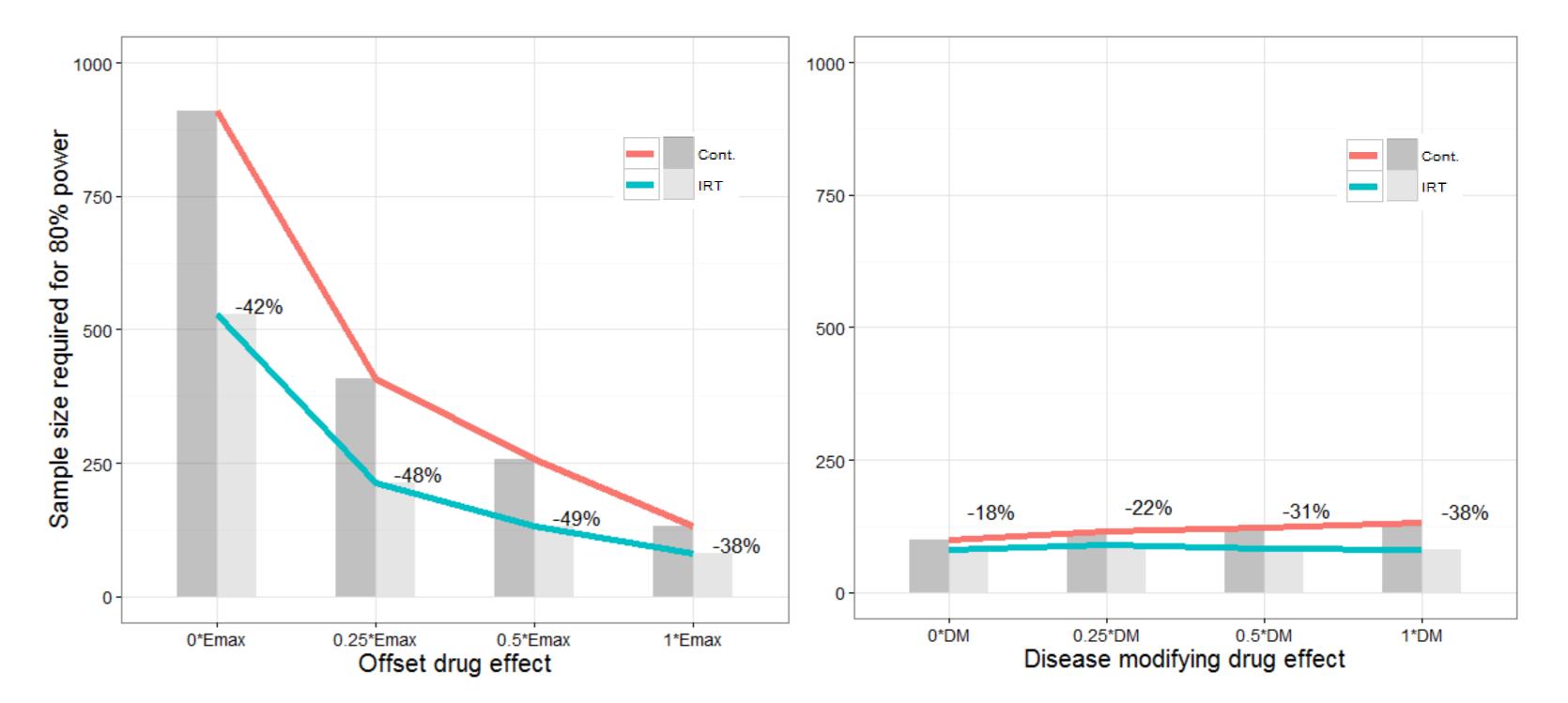
- Clinical trial simulations were used to compare the power to detect the drug effect for two Non-Linear Mixed Effect (NLME) models previously developed for EDSS: IRT and continuous model [2-3]
- The model used for simulations was the IRT model [2] using the parameter estimates previously obtained (Table 1)
- IRT model is a model assuming that the outcome of each item constituting EDSS depends on an unobserved variable

Parameter values (IIV)
1 FIX
0.094 (0.205)
0.733
0.186 (2.03)
409.05
0.216

σ ົບບ 2 Power ' Figure 1. Power to detect a drug effect subjects. Comparison of power curves for IRT and 25 200 100 300 Total sample size

Simulation of scenarios assuming various size of drug effect

- The demonstrated advantage of IRT model over continuous model was preserved accross different scenarios tested (Fig.2-3)
- The reduction in sample size needed ranged from 38% to 49% for various sizes of offset drug effect (Fig.2)
- 18% to 38% fewer patients were necessary with IRT model, in comparison with continuous model across various sizes of disease modifying drug effect investigated (Fig.3)
- In the absence of offset drug effect, sample size required for 80% power is considerably higher, which suggests disease modifying drug effect being more difficult to detect



"disability" (Eq. 1-3)

Corr Dis₀/SL 0.12 Table 1. Parameter values used for simulations

$$Dis(t) = Dis_0 + SL * Time^{pwr} * (1 - EffDm) - EffOff$$
(Eq.1)

$$Exps_{i} = \frac{CumDose*Cl_{cr median}}{Cl_{cr}} \quad (Eq.2) \qquad EffOff = \frac{Emax*Exps_{i}}{Exps_{50}+Exps_{i}} \quad (Eq.3)$$

Dis₀ :Baseline disability, SL: Slope, EffDm: Disease modifying drug effect, EffOff: Offset drug effect, Emax: Maximal obtainable effect, Exps₅₀: Exposure needed for half maximal effect based on cumulative dose administered, Exps_i: Exposure, CumDose: *Cumulative Dose administered, Cl_{cr}: Creatinine clearance, Cl_{cr median}: Median CL_{cr} of the observed population sample,*

- A large dataset (6000 individuals) simulated according to was predefined study design
- Algorithm, based on the same set of rules as used in clinical practice for EDSS assessment, was applied to map subscores, simulated with IRT model, to total EDSS score
- simulated data The were subsequently analyzed using two models, one assuming a drug model) and (full effect assuming no drug effect (reduced model). The same procedure was repeated for continuous composite

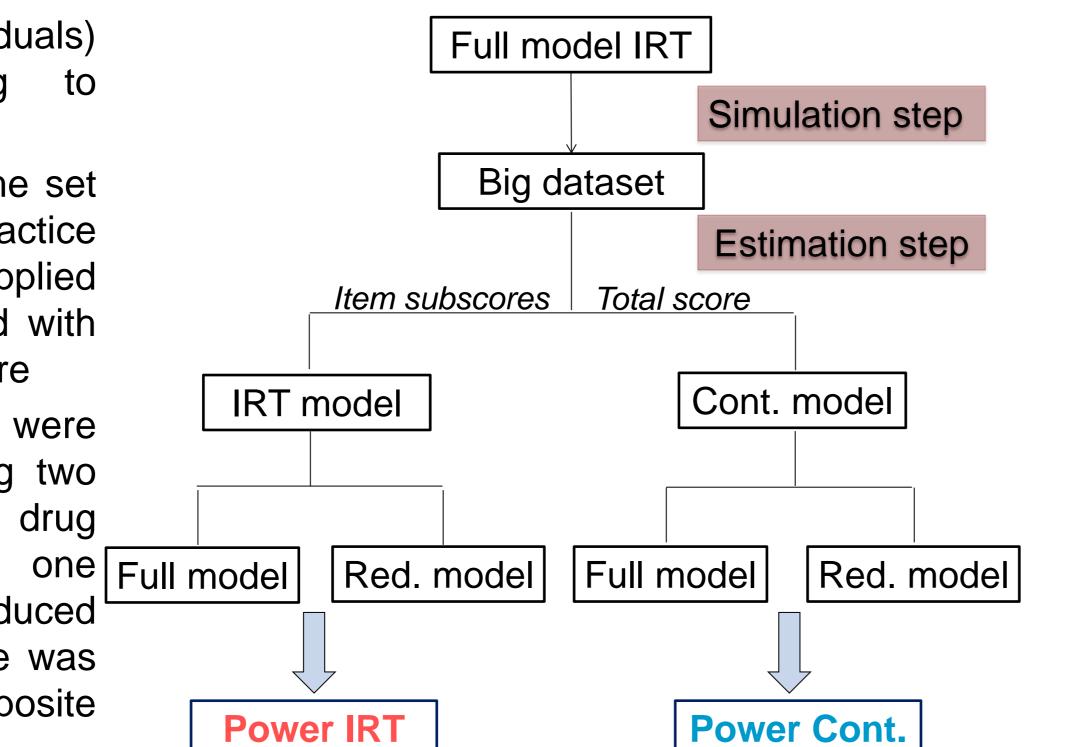


Figure 2. Sample size required for 80% power for a range of sizes of offset drug effect: Comparison of IRT and continuous composite model

Figure 3. Sample size required for 80% power for a range of sizes of disease modifying drug effect: Comparison of IRT and continuous composite model

Conclusions

- Taking a model-based approach offers an opportunity to investigate and maximize the efficiency in clinical trials.
- The application of IRT model demonstrated overall a need for lower sample size to detect the drug effect compared to continuous composite model regardless of the type and of the size of the drug effect

and IRT model:



IRT model

Same as simulation model

Continuous model

• Model treating total EDSS as a continuous, composite score [3] (Eq. 4)

 $EDSS(t) = \left| EDSS_0 + SL * \frac{(10 - EDSS_0) * Time}{2 * 365} * (1 - EffDm) \right| * (1 - EffOff) \quad (Eq.4)$

EDSS₀:Baseline EDSS, SL: Slope, EffDm: Disease modifying drug effect, EffOff: Offset drug effect (as defined in Eq,2-3) Power calculation

• The power calculations were performed using Monte Carlo Mapped Power (MCMP) method [4], implemented in PsN software.

Simulation of scenarios assuming various sizes of drug effect

- Different sizes of drug effect were simulated, by varying the size of one drug effect at time. Following scenarios were considered
- 0, ¹/₄, ¹/₂ and 1 * Emax parameter of Offset drug effect, maintaining *EffDm* unchanged • 0, $\frac{1}{4}$, $\frac{1}{2}$ and 1 * Disease modifying drug effect, maintaining *EffOff* unchanged

- A limitation of this approach is that one of the compared models (IRT model) was used for simulations
- This finding is in line with previous findings that IRT increases precision in predictions and power to detect drug effects and linkage to biomarkers [5, 6].

References

[1] Kurtzke JF, Neurology 33 (11):1444-52, 1983 [2] Kalezic A, PAGE 2013 (Abstract 2903) [3] Savic R, PAGE 2011 (Abstract 2282)

[4] Vong C, AAPS J 14(2):176-86, 2012 [5] Ueckert S, Pharm Res, 2014 [6] Balsis, Alzheimers Dement 8:288-294, 2012

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement n° 115156, resources of which are composed of financial contributions from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. The DDMoRe project is also supported by financial contribution from Academic and SME partners. This work does not necessarily represent the view of all DDMoRe partners.

