Sample size for detection of drug effect using item level and total score models for Unified Parkinson's Disease Rating Scale data

Siv Jönsson[1], Shuying Yang[2], Chao Chen[2] , Elodie L. Plan[1], Mats O. Karlsson[1]

[1]Department of Pharmaceutical Biosciences, Uppsala University, Sweden; [2]GlaxoSmithKline, London, UK

Contact: siv.jonsson@farmbio.uu.se, www.farmbio.uu.se/research/researchgroups/pharmacometrics

Objective

To estimate the sample size required to reach 80% power for detection of a drug effect using an item response model (IRM) and a total score model (TSM), describing longitudinal 44-item Unified Parkinson's Disease Rating Scale (UPDRS) data in advanced Parkinson's disease (PD) patients.

Methods

Data and Models

- ✓ Longitudinal (24 weeks) UPDRS in advanced PD patients [1]
- ✓ Comparison of ropinirole to placebo as adjunct therapy to L-dopa
- ✓ Individually titrated doses between 6 and 24 mg/day
- ✓ Item-level and total score UPDRS data (Table 1)

Table 1. Number of observations [number of patients].

	ITEM	TOTAL
Placebo	31,212 [190]	663 [189]
Ropinirole	33,951 [201]	727 [200]

- ✓ Previous IRM [2] re-estimated for patients in present population of PD patients (Eqs. 1-3)
  - Response for item j (*Yj*), function of unobserved *disability* for subject i (D_i), as a random effect
  - Three unobserved (latent) variables describing Patient reported, Non-sided and Sided responses [3]
  - Extent ($Ext_{ITEM,j}$) and onset ($On_{ITEM,j}$) of symptom relief over time (*t*) in D_i
  - Exposure independent symptomatic drug effect [$SY_{ITEM,j}$] in D_i
  - UPDRS total score and D_i related through item characteristic curves
    - probability for a response (*0, 1, k, .. K*) for each item
    - item specific parameters a_j (slope/discrimination) and b_j (difficulty/location)

$$D_i(t) = D_{i,t=0} + \left(Ext_{ITEM,j} + SY_{ITEM,j}\right) \cdot \left(1 - e^{-On_{ITEM,j} \cdot t}\right) \quad \text{Eq. 1}$$

$$P\left(Y_{ij} \geq k\right) = \frac{e^{a_j(D_i - b_{jk})}}{1 + e^{a_j(D_i - b_{jk})}};\ P\left(Y_{ij} = k\right) = P\left(Y_{ij} \geq k\right) - P\left(Y_{ij} \geq k+1\right) \quad \text{Eqs. 2-3}$$

- ✓ TSM settled for advanced PD patients with complete item records (Eq. 4)
  - exponential placebo time course (Ext_{TS} and On_{TS})
  - symptomatic drug effect [SY_{TS}]

$$TS_i(t) = TS_{i,t=0} + (Ext_{TS} + SY_{TS}) \cdot (1 - e^{-On_{TS} \cdot t}) \quad \text{Eq. 4}$$

Power calculations

- ✓ Monte-Carlo Mapped Power (MCMP) method [4]
  - 10,000 MC samples stratified by treatment
  - Observed individual difference in OFV (iΔOFV) between full (with drug effect) and reduced (without drug effect) models, for IRM and TSM
  - p=0.05 (1 df, ΔOFV 3.84)
- ✓ Support for ΔOFV cut-off used in MCMP
  - Randomisation test [5] for TSM
  - 1000 data sets sampled: placebo, or, placebo and ropinirole data
  - Treatment (placebo:ropinirole, 1:1) randomly assigned for each data set
  - Full and reduced models estimated for each data set
  - Empirical ΔOFV cut-off (p=0.05, 1 df) obtained from ΔOFV distribution

References

1. Pahwa R et al.; EASE-PD Adjunct Study Investigators. Ropinirole 24-hour prolonged release: randomized, controlled study in advanced Parkinson's disease. Neurology, 2007; 68(14):1108-15
2. Jönsson S et al. Placebo and drug response assessment on Unified Parkinson's Disease Rating Scale using longitudinal item response modelling. PAGE 26 (2017) Abstr 7236 [www.page-meeting.org/?abstract=7236]
3. Gottipati G et al. Modeling a composite score in Parkinson's Disease using item response theory. AAPS J,. 2017;19(3):837–45
4. Vong C et al. Rapid sample size calculations for a defined likelihood ratio test-based power in mixed-effects models. AAPS J, 2012; 14(2):176-86
5. Wählby U et al. Assessment of actual significance levels for covariate effects in NONMEM. J Pharmacokinet Pharmacodyn, 2001;28(3):231-52
6. Ueckert S et al.; Alzheimer's Disease Neuroimaging Initiative. Improved utilization of ADAS-cog assessment data through item response theory based pharmacometric modeling. Pharm Res, 2014; 31(8):2152-65
7. Kalezic A et al. Sample size calculations in multiple sclerosis using pharmacometrics methodology: comparison of a composite score continuous modeling and item response theory approach. PAGE 23 (2014) Abstr 3262 [www.page-meeting.org/?abstract=3262]
8. Schindler E et al. Comparison of item response theory and classical test theory for power/sample size for questionnaire data with various degrees of variability in items' discrimination parameters. PAGE 24 (2015) Abstr 3468 [www.page-meeting.org/?abstract=3468]
9. Buatois S et al. Item response theory as an efficient tool to describe a heterogeneous clinical rating scale in de novo idiopathic Parkinson's disease patients. Pharm Res, 2017;34(10):2109-2118

Conclusions and discussion

- ✓ **IRM analysis superior over TSM analysis from study size perspective**
- ✓ **Sample size reduction ~50% for drug effect detection (80% power) using the current data set**
- ✓ **Confirmed, or even exceeded, previous studies where IRM analysis reported sample size reductions varying from 18% to 49% [6-9]**
- ✓ **Use of observed iΔOFVs in MCMP beneficial, but estimated sample size expected to be less precise given low number of iΔOFVs [4]**

Results

Models

- ✓ The IRM and TSM predicted the total score reasonably well (Figure 1).
- ✓ Statistically significant drug effect (p<0.001) with ΔOFVs of -210 and -37 for IRM and TSM, respectively.
- ✓ Approximate model predicted change in total UPDRS at week 24 in placebo group: -2 [-4; -1] (95% CI) and -4 [-6,-2] for IRM and TSM.
- ✓ Approximate model predicted change in total UPDRS at week 24 in ropinirole group: -9.0 [-11,-8.0] and -9.3 [-12,-7.1] for IRM and TSM.

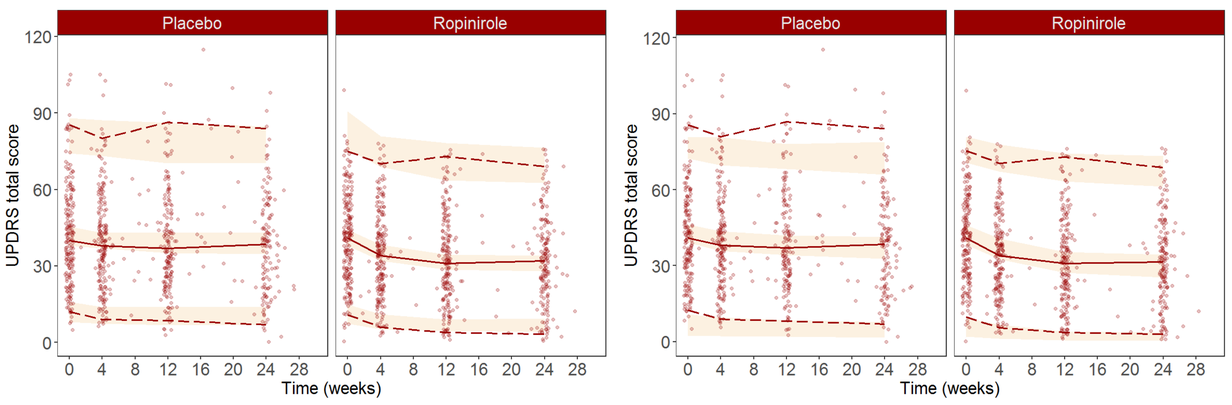

Figure 1. Visual predictive checks for total UPDRS scores based on IRM (left) and TSM (right), respectively.

Power calculations

Monte-Carlo Mapped Power (MCMP) method [4]

- ✓ At 3.84 cut-off, sample size required for 80% power in detecting a drug effect was 54% lower using IRM compared with TSM.

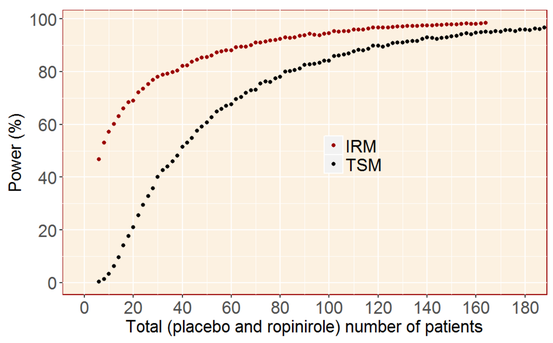

Figure 2. Power versus total sample size (placebo and ropinirole) for IRM and TSM analysis, respectively.

- ✓ The reduction in required sample size tended to be larger when applying a higher cut-off value; sample size reduction of 69% at ΔOFV of 10.8 (Table 2).

Table 2. Estimated total number (placebo and ropinirole) of patients to reach 80% power, analysing with IRT and TSM, respectively.

Critical value (ΔOFV, χ^2)	Sample size (N) IRM	TSM	Ratio N_{TSM}/N_{IRM}
-3	32	72	2.3
-3.84	38	82	2.2
-6.635	44	124	2.8
-10.828	54	176	3.3

Support for ΔOFV cut-off used in MCMP (randomisation test)

- ✓ Type I error rates sampling from only placebo or combined placebo and ropinirole treated patients were similar (Table 3, Figure 3).
- ✓ Sampling from all patients indicated that using a ΔOFV cut-off of 3.84 would be appropriate

Table 3. Actual ΔOFV when sampling from placebo or from placebo and ropiniriole treated patients (sampling 3 times, n=189).

Placebo treated patients			Placebo + ropinirole treated patients		
Actual ΔOFV at 5th percentile					
-4.98	-4.47	-3.97	-3.32	-3.59	-3.36
Actual percentile at ΔOFV 3.84 (χ^2, 1df)					
0.074	0.063	0.051	0.035	0.039	0.039

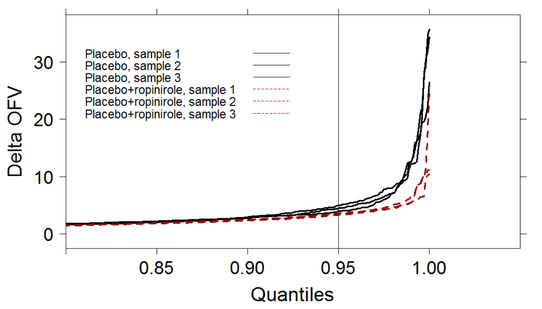

Figure 3. Quantiles of ΔOFV, from randomisations test using placebo or placebo+ropinirole treated patients

Disclosure

SJ, MOK and ELP are employed at Uppsala University (UU). SY and CC are employees of GlaxoSmithKline (GSK), London, UK. UU has received funding from GSK.