Empirical Power Evaluations of an Item Response Model in Parkinson's Disease Patients

Ana Novakovic[1], Kris Jamsen[1], Camille Vong[2], Usman Arshad[2], Chao Chen[2]
[1] Parexel Clinical Pharmacology, Modeling and Simulation, [2] GSK Clinical Pharmacology Modeling and Simulation

Background

- The Unified Parkinson's Disease Rating Scale (UPDRS), a multi-item symptom evaluation tool that includes three sub scales, is the most widely used measure of disability in Parkinson's disease (PD) drug trials [1].
- Despite its validity, the assessment of all required items of the UPDRS can be burdensome on patients and their caregivers.
- Application of item response models (IRMs) can allow for sparser study designs when implementing the UPDRS, which can be beneficial to patients, caregivers and investigators.

Objective

- To evaluate the impact of study design on the statistical power to detect a drug effect within an IRM of the UPDRS in PD patients.

Methods

Development of the Model Used for the Evaluations

- The IRM was developed using data from a ropinirole trial in advanced PD patients [2, 3] and included 27 UPDRS items belonging to Part III: Motor Examination (3 sub-categories: non-sided, left-sided and right-sided; each consisting of 9 items).

Study designs

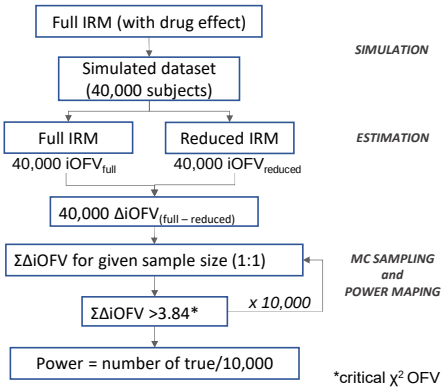
- The IRM was then used for the empirical power evaluations. The following designs were considered:

Scenario	No. of Items	Assessment times (weeks)
1	27	0, 4, 12, 24 for all patients
2	18*	
3	9**	
4	27	0, 4, 24 for 50% of patients; 0, 12, 24 for 50% of patients
5	18*	
6	9**	

*6 of 9 items selected randomly for each of the three subcategories
**3 of 9 items selected randomly for each of the three subcategories

Empirical Power Evaluations

- For each design, the power to detect the drug effect from the IRM was computed using a Monte Carlo Mapped Power (MCMP) procedure [4]:



- To account for *additional* inter-trial variability, stochastic simulation-estimations (SSEs) were explored at several sample sizes for Scenario 1 using the following approaches:
  - Classic application of SSE.
  - With parameter uncertainty based on the standard errors (SEs) from the original analysis (included via PRIOR subroutine).
  - Inflation of the shared placebo/drug effect between-subject variance (BSV; see Table 1) by (i) 25% and (ii) 50%.

Results

Model Used for the Evaluations

- 40,022 UPDRS Part III longitudinal records from 391 patients (190 placebo; 201 ropinirole; all treated over 24 weeks) were used.
- The structure of the underlying severity index was:

$$Latent\ variable_i(t) = BL_i + (PL + DE)_i * (1 - e^{-Onset\ Rate * t})$$

Where BL: baseline, PL: placebo effect, DE: drug effect

- Parameter estimates are shown in Table 1:

Table 1: Parameter estimates

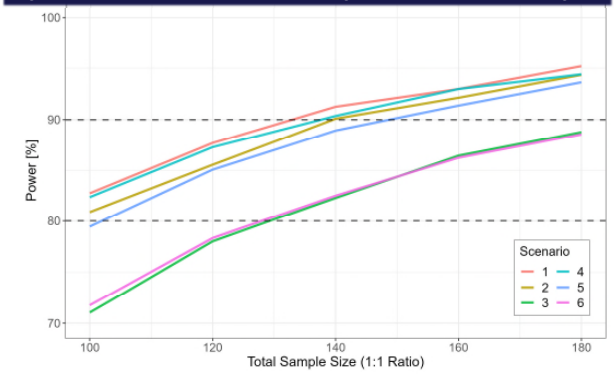
Parameter	Value (RSE%)	BSV (RSE%)
Baseline	0 FIX (-)	1 FIX* (-)
Placebo effect [week^{-1}]	-0.0467 (113)	0.438* (10)
Drug effect [week^{-1}]	-0.437 (17)	
Onset rate [week^{-1}]	0.153 (8)	-

*Correlation between the two BSVs was estimated at -0.259 (22%).

Empirical Power Evaluations

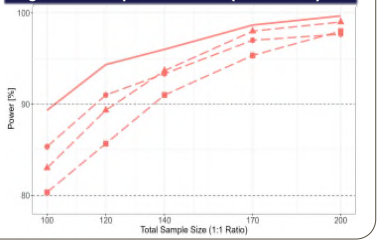
- For MCMP, a reduction in the number of assessed UPDRS items from 27 to 18 resulted in minimal sacrifice in power (Scenarios 1 vs. 2 and 4 vs. 5; Figure 1).
- A further reduction to 9 items (Scenarios 3 and 6) corresponded to a more notable drop in power.
- Specifying 3 visits per patient with stratification (Scenarios 4 - 6) yielded similar power to the corresponding designs with 4 visits.

Figure 1: Power via MCMP to detect the drug effect across different designs



- For SSE (based on 300 replicates), additional variability via parameter uncertainty or BSV inflation resulted in reductions in power for Scenario 1 (Figure 2).

Figure 2: SSE power curves (Scenario 1)



Solid line: classic; circles: with parameter uncertainty; triangles: BSV inflated by 25%; squares: BSV inflated by 50%.

Discussion Points

Summary of Current Findings

- The preliminary results suggest that sparser sampling of UPDRS items (≥18) reduces study power only slightly when using the IRM with a sufficient sample size.

Question for the Audience:

- Is including parameter uncertainty, inflation of BSV, or any other method appropriate and/or critical for accounting for *additional* between-trial variability that may occur, hence providing more conservative predictions of the outcome of a future trial?

References
[1] Ramarker et al., Mov Disord 2021;17:867-876
[2] Chen et al., CPT:PSP 2021 Apr;10(4):309-17
[3] Pahwa et al., Neurology 2007;68(14):1108-15
[4] Vong et al., *The AAPS journal* 2012;14:176-186

31st PAGE meeting, A Coruña, Spain, 28 - 30 June 2023