

Impact of Rater Change on data variability in MDS-UPDRS part III total score in Parkinson's Disease trials

Objectives

The MDS-UPDRS part III is the most widely used tool to measure changes in motorfunction over time in clinical trials of Parkinson's Disease (PD). Previous studies have reported satisfactory inter-rater and intra-rater reliability of the motor assessment (part III) (1.2.3); however, the impact of multiple raters assessing subjects' motor symptoms at different time points during trials is unknown. This study evaluates the impact of rater change in visit-to-visit score variability on the MDS-UPDRS part III total score in PD clinical trials.

Methods

Data from six multinational clinical trials of early PD (H&Y I-III) were analyzed. MDS-UPDRS part III motor assessments conducted by site raters were arranged according to visit sequence and divided into two groups: Rater Change (different raters administered subject's consecutive visits) and No Rater Change (same rater administered consecutive visits). All participating raters were certified to administer and rate the MDS-UPDRS via MDS training. The two groups were matched on mean interval times between visits using random sampling. Visit-to-visit absolute score changes for Part III total score were calculated, and frequency distributions were evaluated.

Results

Welch corrected, two sample t-tests were conducted to compare absolute score changes between rater change groups. For MDS-UPDRS part III, the Rater Change group (Figure 1) showed a higher mean visit-to-visit score change (N = 2,523, mean = 6.7, SD = 5.53) compared to the No Rater Change (N = 7,553, mean = 4.1, SD = 3.74) group (Figure 2). This group difference reached statistical significance (t = 22.31, df = 3325.45, p < 0.0001). Rater Change and No Rater Change groups did not differ significantly on mean interval times (t = 1.80, df = 3891.41, p = .07). (Table 1)

Conclusions

Findings from the study indicate significantly higher score changes when there is a rater change between visits than not for MDS-UPDRS motor assessments (part III total score). These results indicate that ratings from multiple raters can potentially introduce undesirable noise in the data and highlight the importance of rater consistency in reducing variability in an effort to improve data quality in PD trials.

References

1. Post B, Merkus MP, de Bie RM, de Haan RJ, Speelman JD. Unified Parkinson's disease rating scale motor examination: are ratings of nurses, residents in neurology, and movement disorders specialists interchangeable? *Mov Disord.* 2005 Dec; 20(12):1577-84. doi: 10.1002/mds.20640.
2. Richards M, Marder K, Cote L, Mayeux R. Interrater reliability of the Unified Parkinson's Disease Rating Scale motor examination. *Mov Disord.* 1994 Jan;9(1):89-91. doi: 10.1002/mds.870090114.
3. Siderowf A, McDermott M, Kieburtz K, Blindauer K, Plumb S, Shoulson I; Parkinson Study Group. Test-retest reliability of the unified Parkinson's disease rating scale in patients with early Parkinson's disease: results from a multicenter clinical trial. *Mov Disord.* 2002 Jul;17(4):758-63. doi: 10.1002/mds.10011.

Rater Change

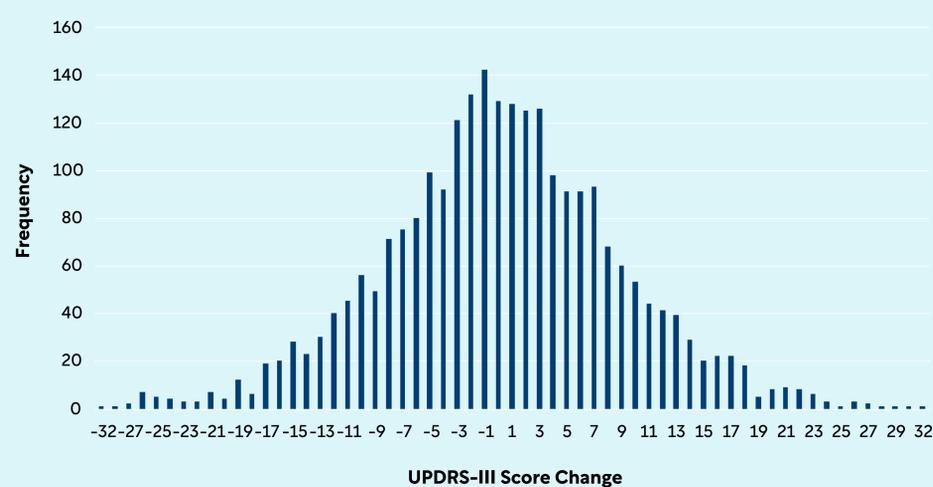


Figure 1

No Rater Change

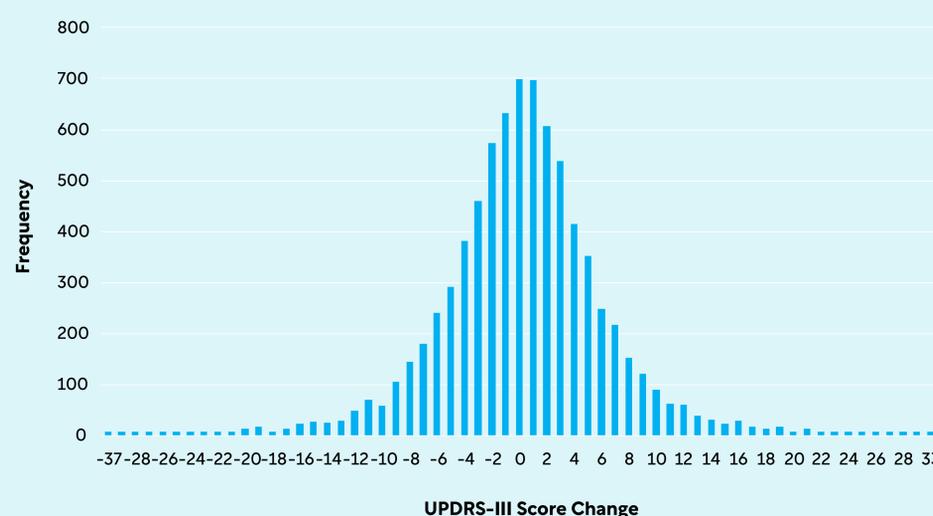


Figure 2

Group	n	Mean	SD	Min	Max	Mean visit to visit interval (weeks)
No Rater Change	7553	4.1	3.74	0	37	6.22
Rater Change	2523	6.7	5.53	0	32	6.33

Significant differences were found between the rater change and no change groups: t = 22.31, df = 3325.45, p < 0.0001.

Table 1