Jacobs, E[1]., DiGregorio, D[1]., Negash, S[1]., Blattner, R[1], and Randolph, C[1]
[1]WCG, Princeton, NJ, USA

The Use of Composite Z-Scores in Place of Normative-Based Scaling to Improve Signal Detection in Clinical Trials Involving Neurodegenerative Diseases

Background

Population-based norming for neuropsychological tests such as the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) is necessary for diagnostic purposes and useful for clinical interpretation of performance in research settings as well (Randolph, et al., 1998). The scaling procedures used to transform raw scores into index scores have the effect of constraining score ranges at the upper and lower limits of performance (Duff, et al., 2011). This may reduce the sensitivity of such tests to detect change at the outer limits of score ranges. This study explored the potential utility of a z-score analysis in capturing a greater range of performance in comparison to index scores in clinical trials of two different neurodegenerative disorders.

Methods

We examined pooled data from clinical trials of Alzheimer disease (AD) and progressive supranuclear palsy (PSP). For AD, there were 1591 subjects at baseline, 693 of which completed week 52. For PSP, there were 823 subjects at baseline and 455 at week 52. The standard index score data were compared to a composite z-score approach for both samples (**Figures 1a** and **1b**). Z-scores were calculated from the baseline mean and standard deviation for each of the 12 subtests, which were then combined to derive a composite z-score. Effect sizes were then calculated between baseline and week 52 and compared between two approaches for both samples.

Results

Tables 1 and **2** show descriptive summary for RBANS subtests for AD and PSP, respectively. The AD and PSP samples were both close to two standard deviations below the normal age-adjusted mean at baseline, as measured by the RBANS total scale index score. In both the AD and PSP samples, the z-score methodology resulted in more normal distribution (**Figures 2** and **3**). There were also larger effect sizes due to disease progression over time in the Z-score methodology than the normative-based index score approach for both AD (effect size: Z-Score = -0.59, normative-based = -0.51) and PSP (effect size: Z-Score = -0.18, normative-based = -0.12) samples (**Table 3**).

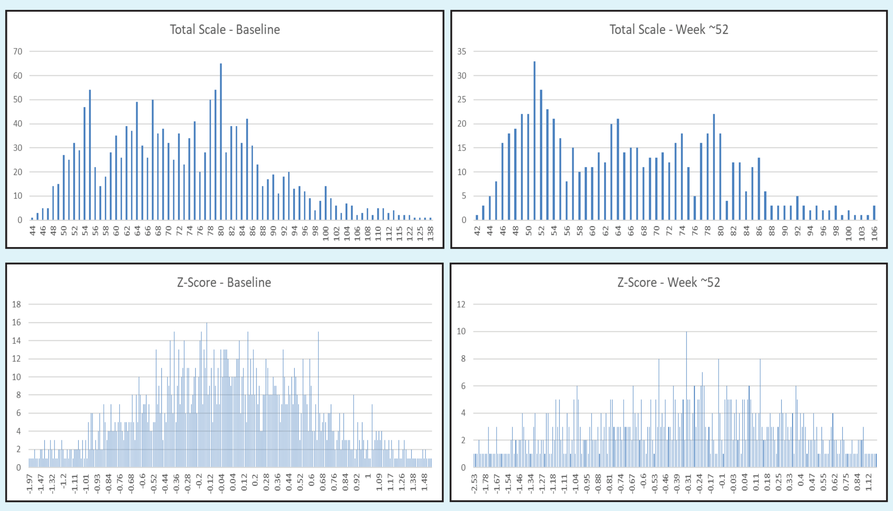

Figure 2. Frequency Distributions - AD

RBANS Subtest	Mean(BL)	Mean(LV)	SD(BL)	SD(LV)
1. List Learning	21.53	20.95	5.61	6.64
2. Story Memory	14.30	14.34	4.65	4.96
3. Figure Copy	11.08	9.12	5.16	5.53
4. Line Orientation	13.87	13.31	4.09	4.53
5. Picture naming	9.22	9.16	1.32	1.24
6. Semantic Fluency	11.36	11.18	4.69	4.44
7. Digit Span	9.18	8.95	2.63	2.69
8. Coding	16.17	13.24	10.20	9.85
9. List Recall	3.63	3.92	2.48	2.47
10. List Recognition	18.03	18.02	2.06	2.24
11. Story Recall	6.92	6.96	2.88	3.05
12. Figure Recall	8.77	7.72	4.89	5.18
Total Scale	73.96	72.39	13.66	13.69

Table 2. Descriptive Summary: PSP

RBANS Subtest	Mean(BL)	Mean(LV)	SD(BL)	SD(LV)
1. List Learning	18.93	16.70	5.59	5.94
2. Story Memory	11.08	7.43	5.09	4.65
3. Figure Copy	14.53	13.93	3.93	4.68
4. Line Orientation	15.22	14.25	3.88	4.55
5. Picture naming	9.10	8.44	1.44	2.05
6. Semantic Fluency	12.76	12.26	5.17	4.56
7. Digit Span	9.46	8.96	2.49	2.69
8. Coding	28.52	24.27	12.66	13.47
9. List Recall	1.73	0.70	2.15	1.38
10. List Recognition	16.23	14.28	2.82	3.09
11. Story Recall	3.70	1.78	3.37	2.14
12. Figure Recall	4.58	2.71	4.84	3.88
Total Scale	73.14	65.45	15.06	14.11

Table 1. RBANS Descriptive Summary: AD

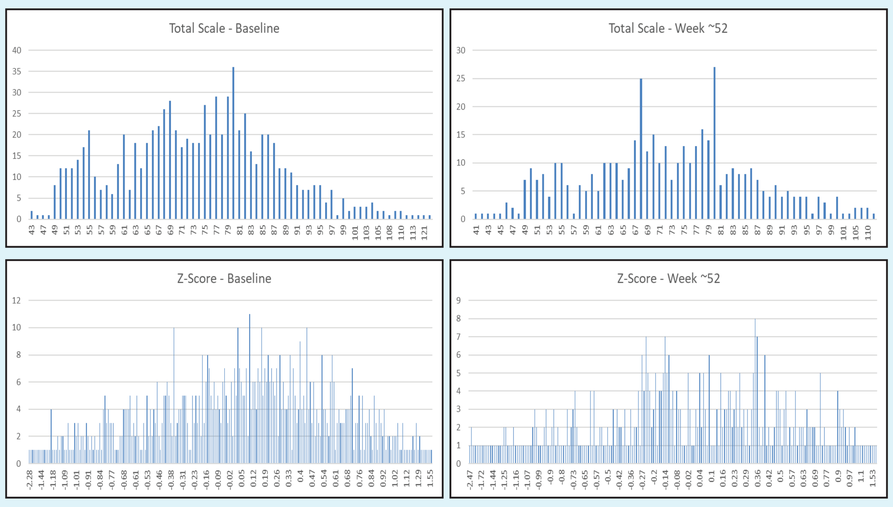

Figure 3. Frequency Distributions - PSP

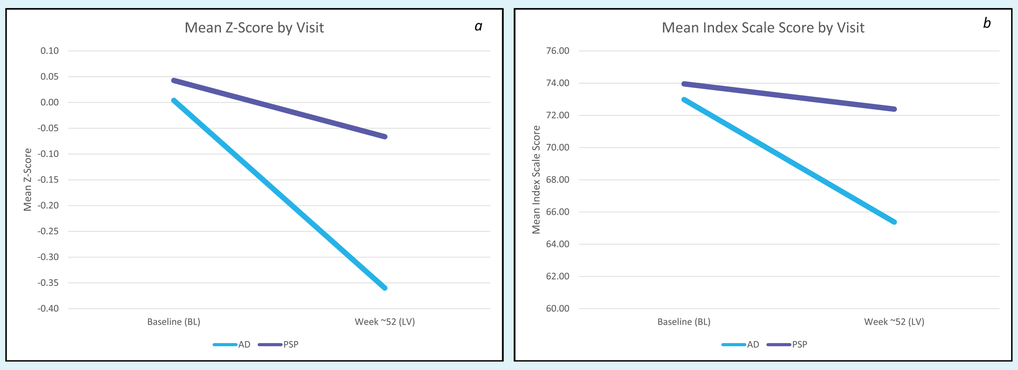

Figure 1.

Indication	Condition	Mean(BL)	SD(BL)	N(BL)	Min(BL)	Max(BL)	Mean(LV)	SD(LV)	N(LV)	Min(LV)	Max(LV)	Cohen
AD	Z-Score	0.00	0.60	1591	-1.97	1.80	-0.36	0.66	693	-2.53	1.37	**-0.59**
AD	Total Scale	72.98	15.12	1591	44.00	138.00	65.37	14.11	693	42.00	106.00	**-0.51**
PSP	Z-Score	0.04	0.59	823	-2.28	1.59	-0.07	0.67	455	-2.47	1.67	**-0.18**
PSP	Total Scale	73.96	13.66	823	43.00	129.00	72.39	13.69	455	41.00	116.00	**-0.12**

Table 3. Effect Sizes

Conclusions

The z-score methodology resulted in more normal distributions and was significantly more sensitive to change due to disease progression than the index score approach. We recommend the use of the z-score methodology for tracking change in study populations where performance is expected to fall at or below these levels relative to the normal population.

Key words:
Alzheimer's disease, neurodegenerative disease, clinical trials, endpoints

Disclosures:
Authors are employees of WCG.

References

Duff K, Spering CC, O'Bryant SE, Beglinger LJ, Moser DJ, Bayless JD, Culp KR, Mold JW, Adams RL, Scott JG. (2011). The RBANS Effort Index: base rates in geriatric samples. Appl Neuropsychol. 2011 Jan;18(1):11-7. doi: 10.1080/09084282.2010.523354.
PMID: 21390895; PMCID: PMC3074382.

Randolph C, Tierney MC, Mohr E, Chase TN. (1998). The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): preliminary clinical validity. J Clin Exp Neuropsychol 20: 310-319.