

Background

There is an acute need for a quantitative summary outcome measure to assess global neurological disability from the perspective of MS patients. Existing measures are burdensome, clearly ordinal, and/or overemphasize physical disability over mental disability.

Objectives

To estimate (nearly) linear patient-reported scores of global neurological disability in MS by applying Item Response Theory (IRT) methods to North American Research Committee on Multiple Sclerosis (NARCOMS) data and; 2) to assess relations between IRT-derived disability scores and patient characteristics

Measures

Disability measures consisted of 11 single-item (6-to-7-point) **Performance Scales - PS** (mobility, dexterity, spasticity, tremor/coordination, sensory, bowel-bladder, cognition, fatigue, pain, depression, vision) and the **Patient-Determined Disease Steps (PDDS)**, which approximates steps of the Expanded Disability Status Scale (EDSS). Other patient characteristics included gender, age at assessment, race/ethnicity, disease modifying therapy, disease duration, year of assessment, and employment status.

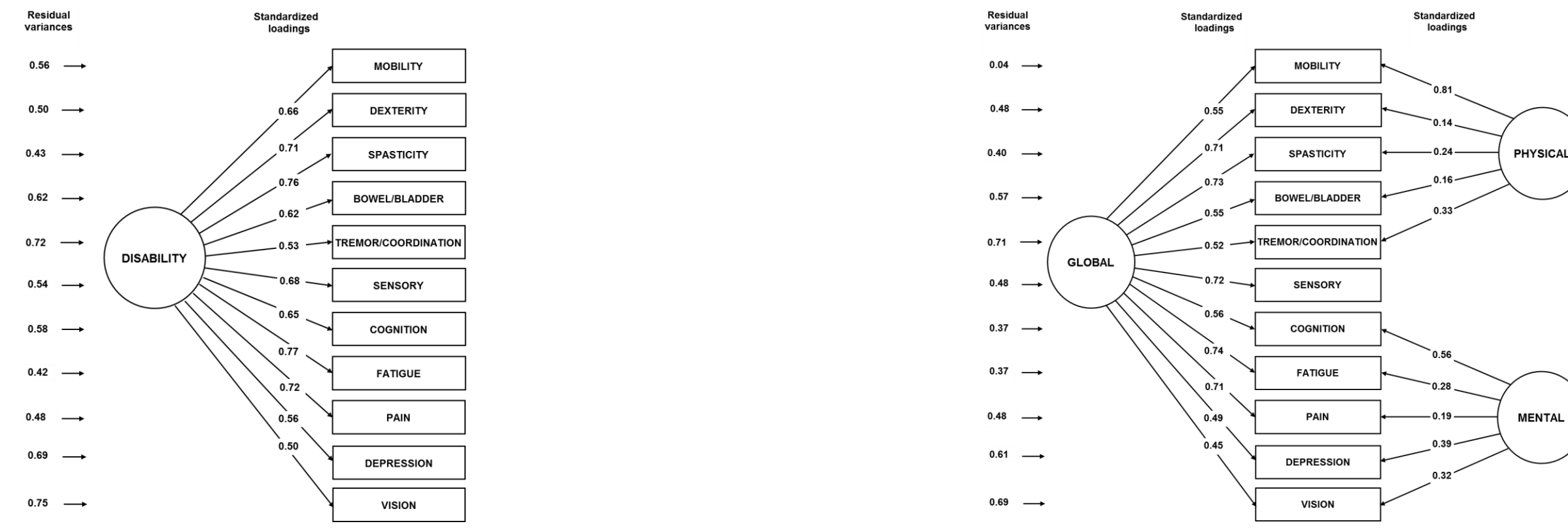
Sample and methods

Study data: Intake surveys collected in 1998-2011. Analyses were restricted to 7,851 patients with complete PS data. The sample was randomly split into development (n=3,926) and validation (n=3,925) sets.

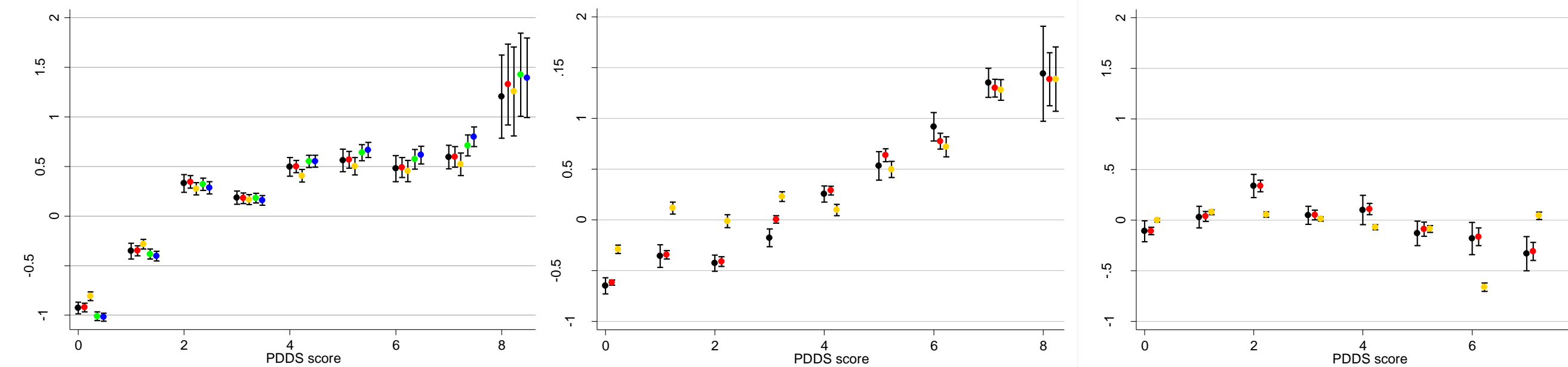
Methods: Item factor analysis and IRT analysis were used to evaluate the measurement structure of the 11 PS and to generate IRT scale scores of disability expressed on the standard deviation (z-score) metric. We applied three methods of IRT scale score estimation: plausible score (gold standard, laborious), Bayes expected a posteriori (EAP, most precise but biased), and summed score EAP (SSEAP, simplest, most convenient in practice but most biased). We employed linear regression to identify patient characteristics associated with each disability score and logistic regression to test the hypothesis of independent association of the scores with unemployment.

Results

Average inter-item tetrachoric correlation was 0.41, Cronbach's alpha was 0.88, and median item-total correlations was 0.70 (range, 0.57-0.77).

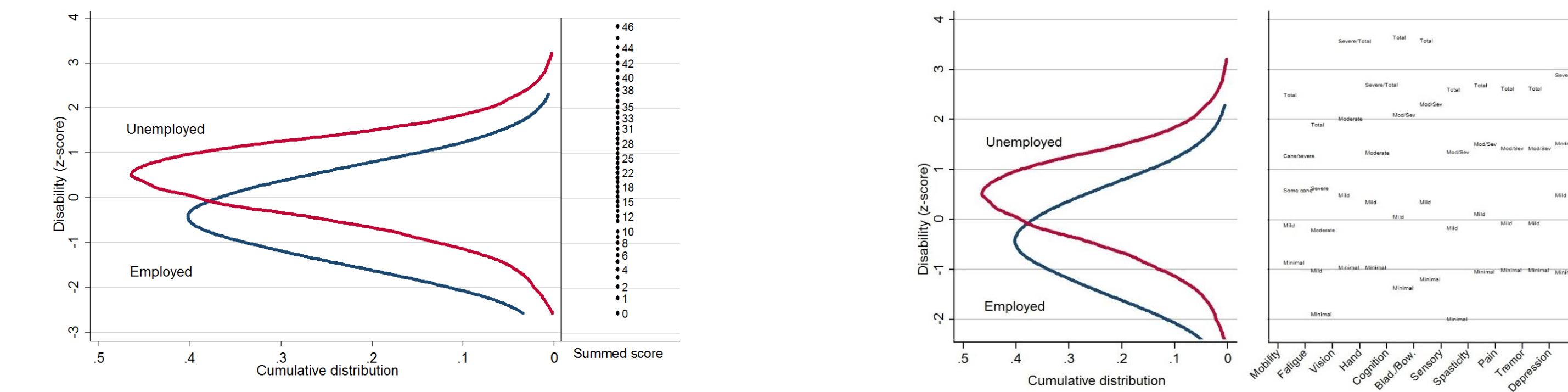


Factor and IRT analyses did not clearly indicate whether the measurement structure of disability was best represented by a one-dimensional model (left figure above) or by a bifactor model including a general factor of global disability and two uncorrelated auxiliary factors accounting for unexplained residual variance of “physical” and “mental” disability (right figure). Several results suggested that the factor of global disability was “strong” and the auxiliary factors “weak”. For instance, 87.8% of reliable variance in PS raw summed scores was explained by the dimension of global disability and only 12.2% by the auxiliary factors ($\Omega_H=0.82$ and $\Omega_{Total}=0.91$).



• Plausible score of global disability (bifactor); • EAP score of global disability (bifactor); • EAP score of global disability (unifactor); • SSEAP score of disability (unifactor).

Scores of both global and physical disability strongly discriminated among PDDS levels. The relation between mental disability and PDDS was more tenuous, possibly reflecting the lack of sensitivity of the PDDS to mental disability or low participation in NARCOMS of severely impaired patients. The method of IRT score estimation (plausible, EAP, SSEAP) had relatively little impact on scores of global disability, but more on the other two scores.



The map to the left establishes a correspondence between PS raw summed scores and SSEAP score estimates of global disability (•). The raw summed scores are on an ordinal scale. The map to the left locates the categories of each PS on the IRT score scale. Notice that “minimal” fatigue disability contributes less to IRT scores than “minimal” mobility disability. In contrast to SSEAP estimates, EAP estimates (•) account for the fact that different patterns of response to the PS can yield the same PS raw summed score.

Table 1. Simple linear regression analysis of association between patient-assessed disability and patient characteristics.

Characteristic	N (%)	Global disability			Residual Physical		Residual Mental	
		Bifactor		Unidimensional	Bifactor	Bifactor	Bifactor	Bifactor
		EAP	SSEAP	EAP	EAP	EAP	EAP	
Gender	3102 (79.4)	0	0	0	0	0	0	
Female	853 (20.6)	0.06 (0.05, 0.11)	-0.08 (-0.13, 0.01)	0.04 (0.02, 0.06)	-0.01 (-0.02, 0.01)	-0.01 (-0.02, 0.01)	-0.01 (-0.02, 0.01)	
Male	2249 (56.8)							
Race/Ethnicity	3481 (80.2)	0	0	0	0	0	0	
White	132 (3.1)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	
African American	120 (3.1)	-0.08 (-0.10, 0.01)	-0.08 (-0.10, 0.01)	-0.08 (-0.10, 0.01)	-0.08 (-0.10, 0.01)	-0.08 (-0.10, 0.01)	-0.08 (-0.10, 0.01)	
Hispanic/Latino	76 (2.0)	-0.08 (-0.10, 0.01)	-0.08 (-0.10, 0.01)	-0.08 (-0.10, 0.01)	-0.08 (-0.10, 0.01)	-0.08 (-0.10, 0.01)	-0.08 (-0.10, 0.01)	
Native American	49 (1.3)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	
Other	49 (1.3)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	
Disease modifying therapy	2814 (66.6)	0	0	0	0	0	0	
Yes	1312 (33.4)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	
No	1502 (38.0)							
Disease duration (y.)	1713 (43.6)	0	0	0	0	0	0	
< 10	1109 (28.3)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	
11 - 20	1104 (28.1)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	
> 21	500 (12.6)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	
Year of assessment	1640 (41.8)	0	0	0	0	0	0	
1998-2005	957 (24.4)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	
2006-2008	1329 (33.8)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	
2009-2011	354 (9.0)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	
Age at assessment (y.)	749 (19.1)	0	0	0	0	0	0	
36 - 55	1760 (44.8)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	
> 56	1417 (36.1)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	0.08 (0.05, 0.11)	

Table 2. Adjusted relative risk of unemployment by disability type and level.

Characteristic	Bifactor model		Unidimensional model	
	EAP	SSEAP	EAP	SSEAP
Global disability (z-score)	1.0	1.0	1.0	1.0
< -1	1.50 (1.28, 1.77)	1.83 (1.55, 2.17)	1.73 (1.48, 2.04)	2.14 (1.83, 2.51)
-1 to 0	2.74 (2.33, 3.22)	3.66 (3.12, 4.30)	3.79 (3.24, 4.43)	
0 to 1				
> 1	1.17 (0.9, 1.38)	1.56 (1.32, 1.84)	2.02 (1.70, 2.40)	
Residual physical disability (z-score)	1.0	1.0	1.0	1.0
< -1	1.11 (0.8, 1.26)	1.35 (1.10, 1.52)	1.08 (1.02, 1.15)	1.08 (1.01, 1.14)
-1 to 1	1.53 (1.32, 1.78)			
> 1				
Residual mental disability (z-score)	1.0	1.0	1.0	1.0
< -1	1.11 (0.8, 1.26)	1.35 (1.10, 1.52)	1.08 (1.02, 1.15)	1.08 (1.01, 1.14)
-1 to 1	1.53 (1.32, 1.78)			
> 1				
Gender (male)	1.0	1.0	1.0	1.0
Female	1.04 (0.8, 1.10)	1.08 (1.02, 1.15)	1.08 (1.01, 1.14)	
Race/Ethnicity	1.0	1.0	1.0	1.0
White	1.0	1.0	1.0	1.0
African American	1.14 (0.9, 1.31)	1.16 (1.00, 1.33)	1.14 (0.9, 1.31)	1.12 (0.9, 1.31)
Hispanic/Latino	1.14 (0.9, 1.31)	1.11 (0.95, 1.31)	1.12 (0.9, 1.31)	
Native American	1.14 (0.9, 1.31)	1.11 (0.95, 1.31)	1.12 (0.9, 1.31)	
Other	1.14 (0.9, 1.31)	1.11 (0.95, 1.31)	1.12 (0.9, 1.31)	
Disease modifying therapy	1.0	1.0	1.0	1.0
Yes	1.05 (0.9, 1.11)	1.08 (1.02, 1.14)	1.08 (1.01, 1.13)	
No	1.0	1.0	1.0	1.0
Disease duration (y.)	1.0	1.0	1.0	1.0
< 10	1.0	1.0	1.0	1.0
11 - 20	1.06 (0.8, 1.14)	1.10 (1.02, 1.19)	1.09 (1.01, 1.18)	1.21 (1.13, 1.30)
> 21	1.21 (1.13, 1.30)	1.27 (1.16, 1.37)	1.25 (1.16, 1.34)	
Year of assessment	1.0	1.0	1.0	1.0
1998-2005	1.01 (0.9, 1.08)	1.01 (0.9, 1.08)	1.02 (0.9, 1.09)	1.02 (0.9, 1.08)
2006-2008	1.01 (0.9, 1.08)	1.01 (0.9, 1.08)	1.02 (0.9, 1.09)	1.02 (0.9, 1.08)
2009-2011	1.04 (0.8, 1.11)	1.01 (0.9, 1.07)	1.01 (0.9, 1.08)	
Age at assessment (y.)	1.0	1.0	1.0	1.0
36 - 55	1.01 (0.9, 1.11)	1.03 (0.9, 1.14)	1.02 (0.9, 1.13)	1.22 (1.10, 1.35)
> 56	1.22 (1.10, 1.35)	1.31 (1.16, 1.45)	1.28 (1.16, 1.42)	

Conclusions

There are several IRT options to generate PS-based scores of global disability. All of them yield scores that have better metric properties than the raw summed score composite. The simplest methods of score derivation might prove useful in clinical practice and the most complex in research. The three dimensions of disability defined by the bifactor model have independent explanatory power, but the auxiliary dimensions are weak (poorly identified).