Interpretation of response categories in patient-reported rating scales: A test-retest study among people with neurological disorders

Peter Hagell The PRO-CARE Group, School of Health and Society, Kristianstad University, Kristianstad, Sweden, Department of Neurology, Lund University Hospital, Lund, Sweden. **Ida Knutsson, Karin Samuelson** Department of Health Sciences, Lund University, Lund, Sweden

Introduction

Unambiguous use and interpretation of rating scale data assumes that response categories are interpreted and work as intended.

This study investigated the stability of interpretations of commonly used patient-reported rating scale response categories among people with neurological disorders.

Methods

Twenty-one rating scale response categories representing ratings of frequency, intensity and level of agreement were presented in random order to 46 people with neurological disorders (26 men; mean age, 57; Parkinson's disease, 50%; multiple sclerosis, 41%).

Respondents indicated their interpretation of each response category on 100-mm visual analog scales (VAS; see Figure for example) anchored by "never" – "always" (frequency), "not at all" – "extremely" (intensity), and "totally disagree" – "completely agree" (agreement) on two occasions ≥2 weeks apart (median interval, 29 days). Patient-reported health and perceived disease severity was stable between the first (T1) and second (T2) occasion. Data were analyzed using intraclass correlation (ICC; should be >0.4-0.5), mean (95%CI) differences, and the standard error of measurement (SEM).

Results

Fourteen ICC values were <0.4 (overall mean ICC, 0.279; frequencies: mean ICC, 0.224; intensities: mean ICC, 0.265; levels of agreement: mean ICC, 0.362). There were no systematic differences in VAS values between time points according to 95%CIs. The mean difference across all 21 response categories was 0.43 mm (mean absolute difference, 3.36 mm). The overall mean SEM was 16.1 (range, 12.3-20.1) mm (frequencies, 17.8mm; intensities, 14.8mm; levels of agreement, 15.3mm). The average proportions of absolute difference scores exceeding the SEM were 32.2% (frequencies), 36.3% (intensity), and 29.8% (agreement).

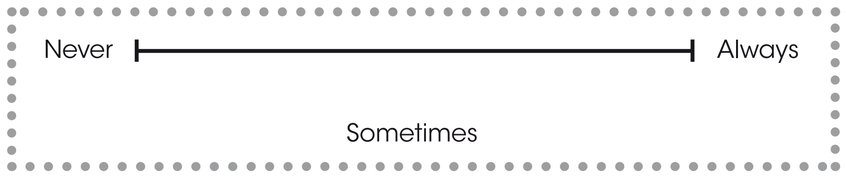

Figure. Sample VAS assessment sheet for recording of respondents' interpretations of rating scale response categories. Respondents were instructed to draw a line from the assessed response category ("sometimes" in the example) to the position on the VAS line where they perceived it to belong relative to the anchors.

Conclusions

Stabilities in interpretations of patient-reported rating scale response categories were generally low. Categories expressing levels of agreement showed best reproducibility, suggesting that these may be preferable, if appropriate with respect to the scale and its items.

These observations suggest caution when interpreting raw rating scale data and argue for the use of modern rating scale methodologies such as the Rasch measurement model, which allows for direct empirical testing of response category functioning.

Table Visual analog scale (VAS) values (possible range, 0-100 mm) for 21 rating scale response categories as determined at two time points (T1 and T2) by people with neurological disorders. [a]

	T1				T2						VAS differences (T1-T2)				
	n	Mean	SD	95% CI	n	Mean	SD	95% CI	ICC	95% CI (ICC)	n	Mean	SD	95% CI	SEM
Frequency:															
Seldom	46	25.2	21.1	18.9-31.4	45	22.6	13.4	18.6-26.7	0.123	-0.175, 0.400	45	3.1	23.4	-4.0, 10.1	19.8
Occasionally	45	36.7	24.3	29.4-44.0	45	32.7	21.6	26.2-39.2	0.512*	0.262, 0.699	44	4.8	22.6	-2.1, 11.7	17.0
A little of the time	45	41.4	24.0	34.2-48.7	46	34.0	21.6	27.6-40.4	0.315	0.038, 0.550	45	7.1	26.6	-0.9, 15.1	19.9
Some of the time	45	43.4	18.6	37.8-49.0	45	47.7	19.1	41.9-53.4	0.442*	0.173, 0.650	44	-3.6	19.8	-9.6, 2.5	13.9
Sometimes	46	50.7	17.3	45.6-55.9	45	46.8	18.0	41.4-52.2	-0.006	-0.296, 0.285	45	3.9	25.1	-3.7, 11.4	17.4
A good bit of the time	46	67.3	21.6	60.9-73.8	46	68.4	17.4	63.3-73.6	0.144	-0.156, 0,417	46	-1.1	25.7	-8.7, 6.5	20.0
Often	46	70.0	17.5	64.8-75.2	46	75.3	13.2	71.4-79.2	0.120	-0.159, 0.387	46	-5.3	20.5	-11.4, 0.8	16.4
Most of the time	46	76.0	19.3	70.2-81.8	46	79.2	11.5	75.8-82.6	0.142	-0.151, 0.412	46	-3.2	20.9	-9.4, 3.0	17.9
Intensity:															
Slightly	46	26.3	19.0	20.6-31.9	46	26.1	17.8	20.8-31.4	0.441*	0.172, 0.648	46	0.1	19.6	-5.7, 5.9	14.2
A little bit	46	26.6	19.5	20.8-32.4	46	30.0	18.4	24.5-35.5	0.509*	0.263, 0.694	46	-3.5	18.7	-9.0, 2.1	13.7
Somewhat	46	31.8	19.6	26.0-37.6	45	35.8	18.7	30.2-41.4	0.205	-0.086, 0.466	45	-4.4	24.1	-11.7, 2.8	17.5
Moderately	44	45.6	14.3	41.2-49.9	46	47.3	14.9	42.8-51.7	0.169	-0.132, 0.442	44	-2.4	18.7	-8.1, 3.3	13.0
Quite a bit	46	69.8	16.4	64.9-74.6	46	74.6	10.7	71.4-77.8	0.182	-0.096, 0.439	46	-4.8	17.6	-10.0, 0.5	14.8
A lot	46	74.1	16.4	69.2-79.0	46	74.4	11.9	70.9-78.0	0.085	-0.215, 0.367	46	-0.3	19.4	-6.1, 5.5	15.7
Agreement:															
Disagree	46	16.2	22.1	9.6-22.7	46	17.4	21.9	10.8-23.9	0.692*	0.505, 0.817	46	-1.2	17.4	-6.3, 4.0	12.3
Mostly false	46	27.0	19.7	21.1-32.8	46	22.1	14.5	17.8-26.4	0.261	-0.019, 0.506	46	4.9	20.9	-1.4, 11.1	16.9
Don't know	40	39.1	22.1	32.0-46.1	42	35.7	18.3	30.0-41.4	0.171	-0.148, 0.458	39	4.3	26.5	-4.3, 12.8	20.1
Do not agree or disagree	42	46.7	18.6	40.9-52.5	45	43.6	18.3	38.2-49.1	0.441*	0.163, 0.655	42	2.7	19.0	-3.2, 8.6	13.9
Mostly true	46	68.7	14.4	64.4-73.0	46	69.6	15.5	65.0-74.2	0.134	-0.165, 0.409	46	-0.9	19.7	-6.7, 5.0	13.4
Agree	46	81.7	20.9	75.5-87.9	46	77.2	24.2	70.0-84.4	0.391	0.120, 0.608	46	4.5	24.9	-3.0, 11.9	16.3
Strongly agree	46	89.3	18.6	83.8-94.8	46	84.9	24.9	77.5-92.3	0.444*	0.184, 0.647	46	4.4	23.1	-2.4, 11.3	13.9

[a] Categories are organized in ascending order (from lower to higher mean VAS values) according to the values from the first time point (T1).
* Response categories with ICC values above the recommended 0.4-0.5 criteria.
T1, time 1; T2, time 2; SD, standard deviation; CI, confidence interval; ICC, intraclass correlation; SEM, standard error of measurement.

The study was accomplished within the Basal Ganglia Disorders Linnaeus Consortium (BAGADILICO) at Lund University, and was supported by the Swedish Research Council, the Swedish Parkinson Academy, and the Faculty of Medicine at Lund University.

Contact: **Peter.Hagell@hkr.se**