

Mika Karsikas

AGE-RELATED DIFFERENCES IN PSYCHOMETRIC PROPERTIES OF WORLD HEALTH ORGANIZATION
DISABILITY ASSESSMENT SCHEDULE 2.0 (WHODAS 2.0)

Syventävien opintojen kirjallinen työ

Syyslukukausi 2024

Mika Karsikas

AGE-RELATED DIFFERENCES IN PSYCHOMETRIC PROPERTIES OF WORLD HEALTH ORGANIZATION
DISABILITY ASSESSMENT SCHEDULE 2.0 (WHODAS 2.0)

Fysiatria

Syyslukukausi 2024

Vastuuhenkilö: Juhani Juhola

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

Syventävien opintojen kirjallinen työ, 14 s.

Fysiatría

Syyskuu 2024

The aim of this study was to detect a possible age-related differential item functioning (DIF) of the WHODAS 2.0 amongst people with musculoskeletal disorders.

This was a cross-sectional study of 1,739 consecutive patients with chronic musculoskeletal pain at a university clinic. Difficulty and discrimination parameters were calculated for each of 12 items of the WHODAS 2.0 using item response theory. The presence of DIF was assessed both numerically and graphically.

When considering differences in both difficulty and discrimination, a statistically significant DIF was observed for all the items except for the item "washing" (p -values <0.001). The DIF was mostly uniform. Items "standing", "household responsibilities", "learning a new task", "emotional affectedness", "concentrating", "washing" and "dressing" showed better precision amongst older patients. Items "walking", "dealing with strangers", "maintaining a friendship" and "day-to-day work" were more informative amongst younger patients.

The 12-item WHODAS 2.0 showed a significant age-related DIF in 11 out of 12 items among people with musculoskeletal pain. Difficulty and discrimination parameters were alike for all 12 items. The DIF was mostly uniform. Items "standing", "household responsibilities", "learning a new task", "emotional affectedness", "concentrating", "washing" and "dressing" showed better precision amongst older patients. Items "walking", "dealing with strangers", "maintaining a friendship" and "day-to-day work" were more informative amongst younger patients. These results can be of use when applying the WHODAS 2.0 to people with musculoskeletal complaints, especially when the studied groups are predominated by people of certain age.

Asiasanat: tuki- ja liikuntaelinten sairaudet, poikkileikkaustutkimukset, psykometriikka, validointitutkimukset

Introduction

It is usually assumed that the response to a questionnaire explicitly reflects only the level of measurable phenomenon. However, in reality, the response may be affected by the intrinsic characteristics of the respondents, for example, by sex, educational level or socioeconomic status (1). If these intrinsic factors significantly affect the response, then the measure may show differences between two groups even though the level of the entity being measured is the same, or, respectively, no differences may be seen even if the real levels of a trait to be measured are different. This phenomenon is called "differential item functioning" (DIF) and is usually studied by performing an analysis based on item response theory (IRT) (2). It has previously been suggested that age can affect the perception of the level of disability (3-5). Age-related DIF has been observed when evaluating various subjective measures of functioning, among others National Health Interview Survey Disability Supplement, PROMIS® physical functioning items, and EQ-5D-5L (3-5).

The WHO Disability Assessment Schedule (WHODAS 2.0) is a general scale to assess the level of disability in different populations and health disorders (6). The easy-to-use 12-item version of the WHODAS 2.0 has gained popularity especially as a screening tool or for research purposes (6, 7). The 12-item WHODAS 2.0 measures the level of disability in six major areas of functioning: cognition, mobility, self-care, getting along, life activities, and participation (6).

The psychometric properties of WHODAS 2.0 have well been studied in different health disorders, and it has been found to be valid and reliable scale (6). Of all the psychometric properties of the WHODAS 2.0, the potential presence of DIF has probably been studied the least. So far, all that is known about the DIF of the WHODAS 2.0 items is that it can be present in some health conditions like cancer and schizophrenia (8, 9) and absent in others like osteoarthritis, myocardial infarction and mild brain injury (10-12). Even though the DIF of other questionnaires measuring disability in musculoskeletal diseases has previously been studied, the possible DIF of WHODAS 2.0 (including age-related DIF) had not been investigated before among people with musculoskeletal complaints (13-15).

Knowledge about the possible DIF of the WHODAS 2.0 can be important, at least in certain situations, for clinicians, researchers and decision-makers. An example can be a situation when there is a need to assess the level of disability in populations with predomination of people of a certain age. Potential age-related DIF should also be taken into account when the level of disability is assessed repeatedly in a few-year interval as aging may affect the response even if the real disability level remains unchanged. The aim of this study was to detect a possible age-related DIF of the WHODAS 2.0 amongst people with musculoskeletal disorders.

Materials and methods

This was a cross-sectional study of consecutive patients with chronic musculoskeletal pain at the Physical and Rehabilitation Medicine (PRM) outpatient clinic of a university hospital between April 2014 and February 2017. The study was based on a self-administered version of the 12-item WHODAS 2.0 that patient had filled out before a physician appointment. The university hospital Ethics Committee approved the study.

Sex was defined as men vs. women. Age was defined in full years at the time of the response. Educational level was defined as high school vs. no high school. Marital status was defined as single vs. co-habiting. Occupational status was dichotomized using the International Standard Classification of Occupations (ISCO) as managers and professionals vs. technicians and manual workers. Pain intensity was assessed using a 11-point numeric rating scale (NRS) from zero 'no pain' to 10 'worst possible pain'. BMI was calculated as body weight in kg divided by a squared height in meters (kg/m²). The level of leisure-time physical activity (lately "physical activity") was assessed by using a four-item questionnaire with a five-grade scale. The patients were asked to mark the amount of their weekly physical activity 'as usual' during the last year or since the onset of the disease. The respondents compared their physical activity to the following items: 1) walking with normal speed; 2) fast walking; 3) jogging; or 4) running. The respondents approximated the amount of physical activity to "none"; "<½ hour"; "around 1 hour"; "2 to 3 hours"; and ">4 hours". The responses were converted into metabolic equivalent of task (MET) and expressed as MET-hour/week. To obtain METs from a questionnaire, the minutes of physical activity were multiplied by the following coefficients: walking with normal speed – by 2.3; fast walking – by 3.6; jogging – by 7.0; and running by 8.0. The total amount of physical activity per one week was the sum of all four items.

When filling out the WHODAS 2.0, the respondents were asked to rate how much difficulty they had experienced during the past 30 days in performing 12 different activities (6). These activities were: standing, household responsibilities, learning a new task, joining in community activities, emotional

affection, concentrating, walking, washing, getting dressed, dealing with unknown people, maintaining a friendship, and day-to-day work. The responses were given on a scale from zero (no disability) to four (extreme disability or inability to perform an activity) (6). The total score of WHODAS 2.0 questionnaire was calculated as a sum of all 12 items divided by 48, multiplied by 100, and presented as a percentage where 100% represents the worst possible disability level.

Statistical analysis

The descriptive statistics were reported as means and standard deviations (SD) or as frequencies and percentage, when appropriate. Two groups of equal size were formed based on age and entitled “younger patients” and “older patients”. The differences between age groups were evaluated by using independent t-test or chi-square test, when appropriate.

Using item response theory, difficulty and discrimination parameters were calculated for each of 12 items of the WHODAS 2.0. Difficulty was understood here as the level of disability (compared to the average level in the sample) needed to give a particular response to an item. In turn, discrimination was understood here as the steepness of the regression curve describing the amount of information obtainable from the response to an item depending on the level of perceived disability. Discrimination parameter was interpreted as: 0.01–0.34 ‘very low’, 0.35–0.64 ‘low’, 0.65–1.34 ‘moderate’, 1.35–1.69 ‘high’ and >1.7 ‘very high’ (16). Here, information was defined as inverse variance. In other words, here, information was understood as the preciseness of the score obtained from a response to an item. The information was explored creating item information curves. The presence of DIF was confirmed if there was statically significant difference (likelihood ratio test resulted in p -value <0.05) between a fully constrained model and a model taking into account differences between items in both difficulty and discrimination. The DIF was considered uniform if the direction of the difference in information curves was similar for the entire scale of disability. In turn, the non-uniform DIF was present if the curves of two groups were overlapping, in other words, the direction of difference varied at the

different levels of disability. The level of statistical significance was set as <0.05 . All the reported p -values were two-tail.

All the analyses were conducted using Stata/IC Statistical Software: Release 18. College Station (StataCorp LP, Texas, USA).

Results

Of 1,739 patients, 1,123 (65%) were women (Table 1). The average age was 46.6 (14.7) years, the average severity of pain 6.3 (2.0) points and the average WHODAS 2.0 composite score was 28.2% (19.5%). There were some statistically significant differences between two age groups.

Table 1. Descriptive characteristics.

Variable	Entire sample		Younger		Older		<i>p</i> -value ^a
	N	%	N	%	N	N	
Sex							0.5590
Men	616	35	302	35	314	36	
Women	1,123	65	567	65	556	64	
Total	1,739	100	869	100	870	100	
Educational status							0.0020
No high school	1,112	67	538	63	574	71	
High school	552	33	312	37	240	29	
Marital status							0.0020
Single	409	25	178	22	231	28	
Cohabiting	1,227	75	645	78	582	72	
Occupational status							0.2910
Managers and professionals	623	47	327	46	296	48	
Service and physical work	706	53	391	54	315	52	
Main diagnosis							<0.001
M54 Dorsalgia	698	40	394	45	304	35	
M79 Soft tissue disorders	182	10	100	12	82	9	
M51 Intervertebral disc disorders	122	7	61	7	61	7	
M75 Shoulder lesions	89	5	42	5	47	5	
Other	643	37	271	34	372	43	
	Mean	SD	Mean	SD	Mean	SD	
Age, years	46.6	14.7	34.9	8.7	58.4	9.0	<0.0001
WHODAS score, %	28.2	19.5	27.3	19.6	29.1	19.4	0.0438
Pain, points	6.3	2.0	6.1	2.0	6.5	1.9	<0.0001
BMI, kg/m ²	27.4	5.7	26.7	6.0	28.1	5.4	<0.0001
Physical activity, MET-h/wk	904.4	868.6	1,145.1	971.5	664.0	671.1	<0.0001

^a t-test for continuous variables, Chi² test for categorical variables

Difficulty parameters were alike for all 12 items (Table 2). Discrimination parameter varied from moderate 0.68 to very high 1.76 (Table 3). The 95% CIs of discrimination parameter were overlapping for all 12 items.

Table 2. Difficulty parameter by age group

Response	Younger		Older		Younger		Older		Younger		Older		Younger		Older			
	Mean	95 CI%	Mean	95 CI%	Mean	95 CI%	Mean	95 CI%	Mean	95 CI%	Mean	95 CI%	Mean	95 CI%	Mean	95 CI%		
S1 Standing for long periods				S5 Emotionally affected by health problems				S9 Dressing										
1 vs 0	0.75	0.38	1.12	-0.19	-0.47	0.09	-1.43	-1.67	-1.20	-1.14	-1.37	-0.90	0.25	0.01	0.49	0.14	-0.05	0.32
2 vs 1	-0.21	-0.56	0.13	-0.45	-0.72	-0.17	0.06	-0.12	0.25	0.24	0.06	0.43	1.23	0.92	1.55	1.19	0.94	1.44
3 vs 2	0.45	0.11	0.78	0.47	0.22	0.73	0.44	0.24	0.63	0.57	0.38	0.76	1.68	1.27	2.09	1.47	1.16	1.77
4 vs 3	1.58	1.18	1.99	1.03	0.76	1.31	2.80	2.37	3.22	3.11	2.62	3.61	3.76	2.86	4.67	3.38	2.67	4.08
S2 Household responsibilities				S6 Concentrating				S10 Dealing with strangers										
1 vs 0	-0.72	-0.87	-0.57	-0.74	-0.91	-0.57	0.84	0.60	1.09	0.87	0.66	1.09	1.92	1.55	2.30	2.19	1.75	2.63
2 vs 1	0.00	-0.13	0.14	-0.11	-0.25	0.04	0.67	0.44	0.91	0.99	0.76	1.21	1.39	1.06	1.73	1.41	1.06	1.77
3 vs 2	0.77	0.62	0.92	0.74	0.60	0.89	1.62	1.31	1.94	1.58	1.29	1.88	1.43	1.02	1.83	1.95	1.47	2.42
4 vs 3	1.86	1.62	2.10	1.84	1.61	2.08	2.54	2.03	3.06	3.03	2.43	3.64	2.46	1.89	3.03	2.70	1.96	3.43
S3 Learning a new task				S7 Walking a long distance				S11 Maintaining a friendship										
1 vs 0	2.28	1.80	2.75	1.97	1.55	2.40	0.58	0.32	0.84	0.63	0.26	1.00	0.68	0.51	0.86	1.18	0.92	1.45
2 vs 1	1.62	1.20	2.05	1.25	0.89	1.61	0.35	0.09	0.62	0.05	-0.32	0.41	1.09	0.89	1.29	1.34	1.06	1.62
3 vs 2	1.48	0.96	2.00	1.64	1.18	2.10	0.96	0.66	1.26	0.35	-0.01	0.71	1.23	0.99	1.47	1.64	1.29	1.99
4 vs 3	3.38	2.44	4.32	2.89	2.11	3.67	1.18	0.84	1.52	1.01	0.64	1.38	2.35	1.97	2.73	3.33	2.60	4.07
S4 Joining in community activities				S8 Washing ^a				S12 Day-to-day work										
1 vs 0	0.26	0.11	0.40	0.27	0.11	0.43	0.61	0.40	0.82	0.44	0.26	0.62	-1.06	-1.28	-0.84	-0.52	-0.77	-0.27
2 vs 1	0.66	0.49	0.82	0.53	0.36	0.69	1.11	0.86	1.35	1.04	0.82	1.26	-0.19	-0.39	0.00	-0.26	-0.49	-0.03
3 vs 2	0.91	0.73	1.10	0.90	0.72	1.08	1.51	1.19	1.83	1.44	1.16	1.72	0.26	0.06	0.46	0.68	0.44	0.91
4 vs 3	1.77	1.52	2.02	1.70	1.46	1.95	3.09	2.46	3.73	2.27	1.85	2.70	0.64	0.43	0.84	0.99	0.73	1.24

^a Likelihood ratio test 0.0474, all the others <0.0001

Table 3. Discrimination parameter of WHODAS 2.0 items by age group

WHODAS item	Younger			Older		
	Mean	95 CI%		Mean	95 CI%	
S1 Standing for long periods	0.66	0.56	0.76	0.89	0.74	1.03
S2 Household responsibilities	1.65	1.41	1.89	1.76	1.48	2.04
S3 Learning a new task	0.89	0.71	1.06	0.93	0.75	1.12
S4 Joining in community activities	1.68	1.41	1.95	1.71	1.41	2.00
S5 Emotionally affected by health problems	1.08	0.92	1.24	1.20	1.00	1.39
S6 Concentrating	1.05	0.88	1.22	1.23	1.01	1.44
S7 Walking a long distance	0.90	0.76	1.04	0.67	0.56	0.79
S8 Washing	1.05	0.88	1.22	1.25	1.04	1.46
S9 Dressing	0.74	0.62	0.86	1.02	0.85	1.19
S10 Dealing with strangers	1.16	0.94	1.39	1.07	0.85	1.29
S11 Maintaining a friendship	1.54	1.28	1.79	1.08	0.89	1.28
S12 Day-to-day work	1.13	0.96	1.30	0.99	0.83	1.15

When taking into account differences in both difficulty and discrimination, a statistically significant DIF was observed (p -values <0.001 as evaluated by likelihood ratio test) for all the items except for the item “washing” ($p=0.061$). Figure 1 showed item information curves for all 12 items. The DIF was mostly uniform or almost uniform for all the WHODAS 2.0 items. Items “standing”, “household responsibilities”, “learning a new task”, “emotional affectedness”, “concentrating”, “washing” and “dressing” showed better precision amongst older patients, while items “walking”, “dealing with strangers”, “maintaining a friendship” and “day-to-day work” were more informative amongst younger patients.

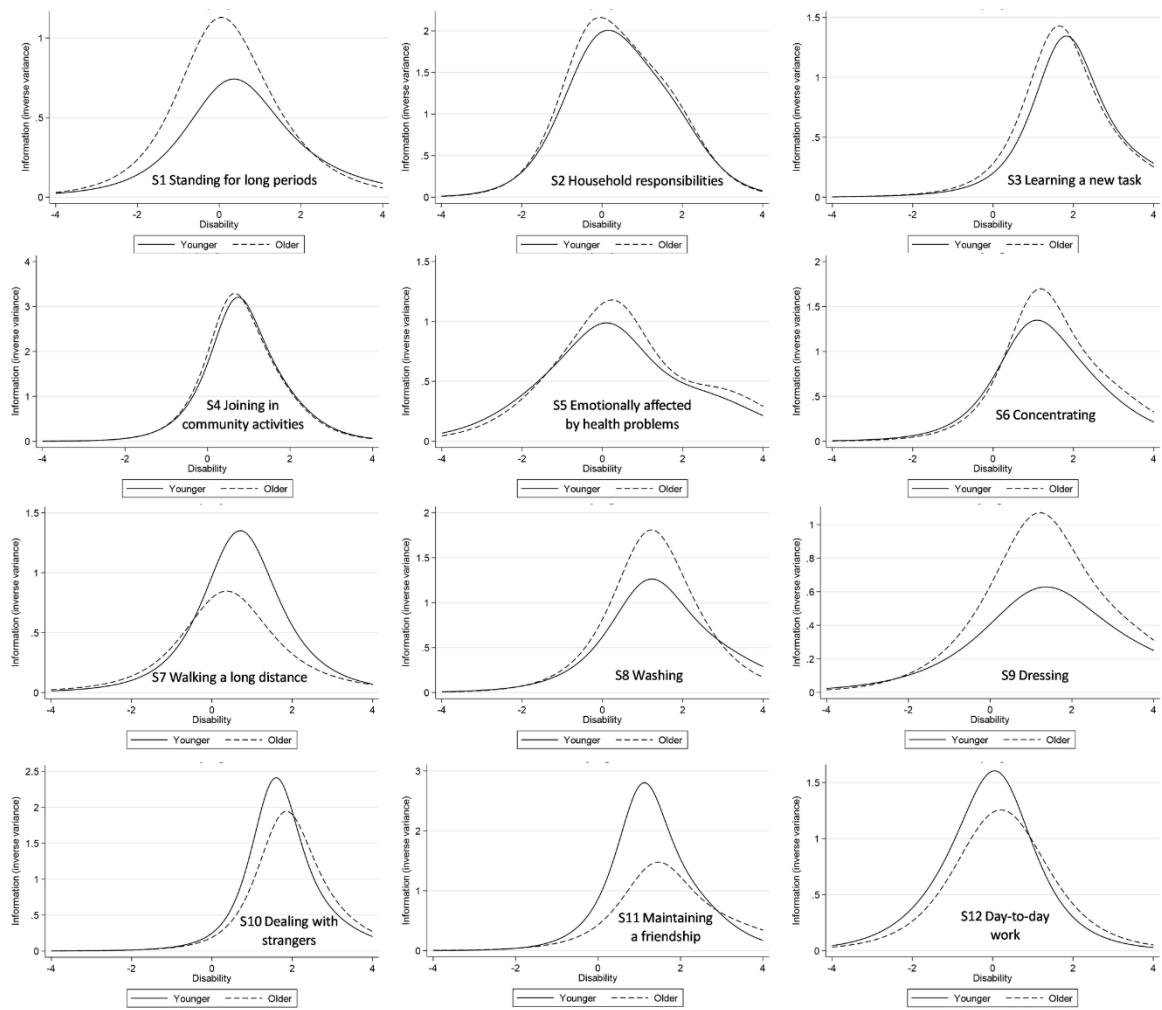


Figure 1. Item information curves of WHODAS 2.0 items by age group.

Discussion

This study amongst 1,739 patients with musculoskeletal pain showed significant age-related DIF between younger and older subgroups in most of the items of 12-item WHODAS 2.0. Difficulty and discrimination parameters were alike for all 12 items. When taking into account differences in both difficulty and discrimination, a statistically significant DIF was observed for all the items except for the item "washing". The DIF was mostly uniform. Items "standing", "household responsibilities", "learning a new task", "emotional affectedness", "concentrating", "washing" and "dressing" showed better precision amongst older patients. Items "walking", "dealing with strangers", "maintaining a friendship" and "day-to-day work" were more informative amongst younger patients.

The generalizability of the results might be weakened by the fact that the study was performed in a highly specialized health care unit. Therefore, the situation may differ in other settings, like e.g., primary health care. Although the sample size was sufficient to produce statistically significant results, the group of less than 2,000 patients is hardly representative of the entire population suffering from musculoskeletal diseases. In addition, the sample was heterogeneous in terms of diseases and the sample size was not sufficient to perform subgroup analyses. It also has to be mentioned that the IRT is always an approximation.

The age-related DIF of WHODAS 2.0 among people suffering from chronic musculoskeletal disorders is not well studied. Therefore, a direct comparison between current results and previous knowledge is difficult. Indirectly, the results can be mirrored in studies where the WHODAS was used in other diseases and in different age groups. For example, Berrio et al. have seen the age-related DIF in a single item 'getting along' when the 36-item WHODAS 2.0 was used in schizophrenia patients (9). Also, Svanborg et al. have studied the 36-item WHODAS 2.0 characteristics among patients with mental disorders and reported DIF for one item, but in that case, it was "sexual activities" (17). Vaganian et al. have observed the age-related DIF in the 'self-care' domain of the 12-item WHODAS 2.0 in patients with cancer (8). On the other hand, Young et al. have studied the 32- and 36-item WHODAS 2.0

properties in multiple sclerosis patients without seeing significant DIF (18). Also, Kirchberger et al. did not observe DIF in the WHODAS 2.0 among patients with myocardial infarction (11). In this way, the only general observation regarding the possible phenomenon of DIF in WHODAS 2.0 is that it has not been studied in depth enough so far. Based on previous studies, however, it can be assumed that DIF can occur in certain situations and this possibility must be taken into account when using this measure. People can experience a chronic musculoskeletal disorder at very different ages. It is reasonable to assume that people of different ages may experience the types and severity of limitations differently at different ages.

Further research on the possible DIF of WHODAS 2.0 among people with different musculoskeletal disorders is needed. It is desirable that sample sizes are big enough to include subgroup analyses based on different disorders and descriptive characteristics. Also, a longitudinal repeated measures research may show the stability of DIF over time and possible changes in its magnitude and direction through different ages.

Conclusions

The 12-item WHODAS 2.0 showed a significant age-related DIF in 11 out of 12 items among people with musculoskeletal pain. Difficulty and discrimination parameters were alike for all 12 items. The DIF was mostly uniform. Items “standing”, “household responsibilities”, “learning a new task”, “emotional affectedness”, “concentrating”, “washing” and “dressing” showed better precision amongst older patients. Items “walking”, “dealing with strangers”, “maintaining a friendship” and “day-to-day work” were more informative amongst younger patients. These results can be of use when applying the WHODAS 2.0 to people with musculoskeletal complaints, especially when the studied groups are predominated by people of certain age.

References

1. Katajapuu N, Laimi K, Heinonen A, Saltychev M. Gender-related differences in psychometric properties of WHO Disability Assessment Schedule 2.0. *Int J Rehabil Res.* 2019;42(4):316-21.
2. Steven PR, Dennis AR. *Handbook of Item Response Theory Modeling : Applications to Typical Performance Assessment.* New York: Routledge; 2015.
3. Fleishman JA, Spector WD, Altman BM. Impact of Differential Item Functioning on Age and Gender Differences in Functional Disability. *The Journals of Gerontology: Series B.* 2002;57(5):S275-S84.
4. Paz SH, Spritzer KL, Morales LS, Hays RD. Age-related Differential Item Functioning for the Patient-Reported Outcomes Information System (PROMIS®) Physical Functioning Items. *Prim Health Care.* 2013;3(131).
5. Penton H, Dayson C, Hulme C, Young T. An Investigation of Age-Related Differential Item Functioning in the EQ-5D-5L Using Item Response Theory and Logistic Regression. *Value in Health.* 2022;25(9):1566-74.
6. World Health Organisation. *Measuring Health and Disability: Manual for WHO Disability Assessment Schedule (WHODAS 2.0).* Geneva, Switzerland: WHO Press, World Health Organization; 2010.
7. Saltychev M, Katajapuu N, Bärlund E, Laimi K. Psychometric properties of 12-item self-administered World Health Organization disability assessment schedule 2.0 (WHODAS 2.0) among general population and people with non-acute physical causes of disability – systematic review. *Disability and Rehabilitation.* 2021;43(6):789-94.
8. Vaganian L, Bussmann S, Boecker M, Kusch M, Labouvie H, Gerlach AL, et al. An item analysis according to the Rasch model of the German 12-item WHO Disability Assessment Schedule (WHODAS 2.0). *Qual Life Res.* 2021;30(10):2929-38.

9. Berrío Á I, Gómez-Benito J, Guilera G. Differential Item Functioning in the WHODAS 2.0 Scale in Schizophrenia: An Application of the Rasch Trees Method Based on Demographic and Clinical Covariates. *Assessment*. 2022;29(8):1858-68.
10. Kutlay Ş, Küçükdeveci AA, Elhan AH, Öztuna D, Koç N, Tennant A. Validation of the World Health Organization disability assessment schedule II (WHODAS-II) in patients with osteoarthritis. *Rheumatology International*. 2011;31(3):339-46.
11. Kirchberger I, Braitmayer K, Coenen M, Oberhauser C, Meisinger C. Feasibility and psychometric properties of the German 12-item WHO Disability Assessment Schedule (WHODAS 2.0) in a population-based sample of patients with myocardial infarction from the MONICA/KORA myocardial infarction registry. *Population Health Metrics*. 2014;12(1):27.
12. Snell DL, Siegert RJ, Silverberg ND. Rasch analysis of the World Health Organization Disability Assessment Schedule 2.0 in a mild traumatic brain injury sample. *Brain Inj*. 2020;34(5):610-8.
13. Crins MHP, Terwee CB, Ogreden O, Schuller W, Dekker P, Flens G, et al. Differential item functioning of the PROMIS physical function, pain interference, and pain behavior item banks across patients with different musculoskeletal disorders and persons from the general population. *Qual Life Res*. 2019;28(5):1231-43.
14. Christiansen DH, McCray G, Winding TN, Andersen JH, Nielsen KJ, Karstens S, et al. Measurement properties of the musculoskeletal health questionnaire (MSK-HQ): a between country comparison. *Health and Quality of Life Outcomes*. 2020;18(1).
15. Nishigami T, Mibu A, Tanaka K, Yamashita Y, Watanabe A, Tanabe A. Psychometric properties of the Japanese version of short forms of the Pain Catastrophizing Scale in participants with musculoskeletal pain: A cross-sectional study. *Journal of Orthopaedic Science*. 2017;22(2):351-6.
16. Baker FB. The basics of item response theory: ERIC; 2001. 6 p.
17. Svanborg C, Amer A, Nordenskjöld A, Ramklint M, Soderberg P, Tungstrom S, et al. Evidence for validity of the Swedish self-rated 36-item version of the World Health Organization Disability

Assessment Schedule 2.0 (WHODAS 2.0) in patients with mental disorders: a multi-centre cross-sectional study using Rasch analysis. *J Patient Rep Outcomes*. 2022;6(1):45.

18. Young CA, Rog DJ, Sharrack B, Constantinescu C, Kalra S, Harrower T, et al. Measuring disability in multiple sclerosis: the WHODAS 2.0. *Qual Life Res*. 2023;32(11):3235-46.