

Minimal clinically important difference and minimal detectable change of the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) amongst patients with chronic musculoskeletal pain

Clinical Rehabilitation

1–6

© The Author(s) 2020



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0269215520942573

journals.sagepub.com/home/cre

Niina Katajapuu^{1,2}, Ari Heinonen¹
and Mikhail Saltychev³

Abstract

Objectives: The aim of this study is to estimate a minimal clinically important difference (MCID) and a minimal detectable change (MDC) of the 12-item WHODAS 2.0 amongst patients with chronic musculoskeletal pain.

Design: Cross-sectional cohort study.

Setting: Outpatient Physical and Rehabilitation Medicine clinic.

Subjects: A total of 1988 consecutive patients with musculoskeletal pain.

Interventions: A distribution-based approach was employed to estimate a minimal clinically important difference, a minimal detectable change, and a minimal detectable percent change (MDC%).

Results: The mean age of the patients was 48 years, and 65% were women. The average intensity of pain was 6.3 (2.0) points (0–10 numeric rating scale) and the mean WHODAS 2.0 total score was 13 (9) points out of 48. The minimal clinically important difference ranged between 3.1 and 4.7 points. The minimal detectable change was 8.6 points and minimal detectable % change was unacceptably high 66%.

Conclusions: Amongst patients with chronic musculoskeletal pain, the 12-item WHODAS 2.0 demonstrated a high minimal detectable change of almost nine points. As the minimal detectable change exceeded the level of minimal clinically important difference, nine points were considered to be the amount of change perceived by a respondent as clinically significant.

Keywords

Whodas, minimal clinically important difference, minimal detectable change, musculoskeletal pain

Received: 9 January 2020; accepted: 25 June 2020

¹Faculty of Sports and Health Sciences, University of Jyväskylä, Jyväskylä, Finland

²Faculty of Health and Wellbeing, Turku University of Applied Sciences, Turku, Finland

³Department of Physical and Rehabilitation Medicine, Turku University Hospital and University of Turku, Turku, Finland

Corresponding author:

Niina Katajapuu, Turku University of Applied Sciences, Joukahaisenkatu 3, Turku, 20520, Finland.

Email: niina.katajapuu@turkuamk.fi

Introduction

The World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) is a generic tool to assess functioning in diverse situations.¹⁻³ While the WHODAS 2.0 has widely been used in clinical practice and research, the interpretation of results, obtained from the WHODAS 2.0 responses, have not been well defined.^{4,5}

The interpretation of test results relies heavily on such characteristics as minimal clinically important difference (“MCID”) and minimal detectable change (“MDC”). While changes in test score may be statistically significant, they are not necessarily perceived by patients as clinically significant. This is especially true when the results are obtained from a large sample – a very small difference may become statistically significant while its practical importance perceived by patients is negligible. Minimal clinically important difference describes the smallest amount of change or difference that might be considered important by patients or clinicians.⁶ There are two common methods to calculate a minimal clinically important difference: an anchor-based and a distribution-based. There is no general agreement on which method is preferable. Probably, they both have their pros and cons in different particular situations. The anchor can be either an objective or subjective measure (e.g. question about mild improvement noticed by a patient or a clinician). An anchor-based method reflects the patient’s or clinician’s point of view. In turn, a distribution-based method is based explicitly on the statistical variability of obtained scores. A minimal detectable change is the smallest amount of change or difference that is not the result of measurement error.⁷ In an clinically ideal world, the minimal clinically important difference must exceed the level of minimal detectable change. If minimal clinically important difference is less than minimal detectable change, then the observed result below the level of minimal detectable change may be caused by chance and not by the true difference in scores even if the result exceeds the level of minimal clinically important difference.⁸

Thus, without knowledge on minimal clinically important difference and minimal detectable

change, the clinical meaning of the WHODAS 2.0 total score estimates remains unclear. The minimal clinically important difference of the 12-item WHODAS 2.0 has been established by a single study amongst patients with anxiety and stress disorders.⁹ In that study, an anchor-based method has been used, and the minimal clinically important difference has been estimated around three points for a less strict model and six to seven points for a stricter model. The minimal detectable change of the WHODAS 2.0 has also recently been reported by a single study amongst institutionalized ambulatory older adults as 10 points.¹⁰ So far, there have not been reports on the minimal clinically important difference or minimal detectable change of the WHODAS 2.0 amongst patients with musculoskeletal health conditions. Due to the WHODAS 2.0 score’s multidimensionality and, thus, potentially high level of minimal detectable change, the trustworthiness of the WHODAS 2.0 total score has been questioned.^{2,11}

The objective of this study was to estimate the minimal clinically important difference and minimal detectable change of 12-item WHODAS 2.0 amongst people with chronic musculoskeletal pain.

Methods

Data for this study were derived from the Turku ICF Study (T54/2012) approved by the ethics committee of Hospital District of Southwest Finland (ETMK 60/180/2012). Participants provided their written informed consent for participation. This was a cross-sectional study amongst 3150 consecutive patients who were seen in an outpatient Physical and Rehabilitation Medicine clinic of university hospital between April 2014 and February 2017. The survey was sent to the patients and filled up before a physician appointment. The survey included 12-item WHODAS 2.0 questionnaire and questions on demographics, pain intensity, and perceived general health.

Self-administered 12-item WHODAS 2.0

The self-administered 12-item WHODAS 2.0 contains 12 items covering the most common limitations of functioning appearing in general population

(Appendix A). The questionnaire covers limitations during the last 30 days. A Likert-like scale is used to define the severity of limitation with 0 denoting “no limitation” and 4 denoting “extreme limitation or inability to function.” The total score is the sum of all 12 items where a score of 48 points represents the worst possible restriction.¹

Independent variables

Age was defined in full years at the time of visiting the clinic. *Pain intensity* was assessed using a 11-point numeric rating scale with 0 denoting “no pain” and 10 denoting “worst possible pain.” *Educational level* was dichotomized “further education” (equivalent “further education or higher” in UK) versus “no further education” (equivalent of “primary and secondary education” in UK). *Body mass index* was calculated as a body mass divided by a squared body height (kg/m^2). *Perceived general health* status was assessed on a 4-point scale where 0 indicated best possible and 3 worst possible health. Main diagnoses were defined using the International Classification of Diseases, 10th edition

Statistical analysis

The results were reported as means, standard deviations, and standard errors, medians, ranges, and interquartile ranges when appropriate. The internal consistency was assessed by Cronbach’s alpha considering $\alpha \geq 0.9$ excellent, $0.8 \leq \alpha < 0.9$ good, $0.7 \leq \alpha < 0.8$ acceptable, $0.6 \leq \alpha < 0.7$ questionable, $0.5 \leq \alpha < 0.6$ poor, and $\alpha < 0.5$ unacceptable.

To describe the variability between an individual’s observed score and the true score, standard error of measurement (SEM) was calculated as $\text{SEM} = \text{SD} \times \sqrt{1 - r_{xx}}$ where r_{xx} is reliability coefficient of the test – in this case, Cronbach’s alpha.¹² Since the data were cross-sectional and no patients’ opinion on perceived change in functioning was available as an anchor, a distribution-based approach was employed to estimate minimal clinically important difference for the WHODAS 2.0. Three different formulas were used for the task:^{13–18}

- 1) Minimal clinically important difference = standard error of measurement
- 2) Minimal clinically important difference = $0.5 \times$ standard deviation
- 3) Minimal clinically important difference = $0.33 \times$ standard deviation

The minimal detectable change was calculated as $1.96 \times$ standard error of measurement $\times \sqrt{2}$. The minimal detectable change was also expressed as a percentage (“MDC%”) – an estimate that is independent of the units of measurement. Representing the relative amount of random measurement error, the minimal detectable % change was calculated as (minimal detectable change /observed mean WHODAS total score) $\times 100$. The minimal detectable % change $< 30\%$ was considered acceptable and $< 10\%$ excellent^{19,20}

All the analyses were conducted using Stata/IC Statistical Software: Release 15. College Station (StataCorp LP, TX, USA).

Results

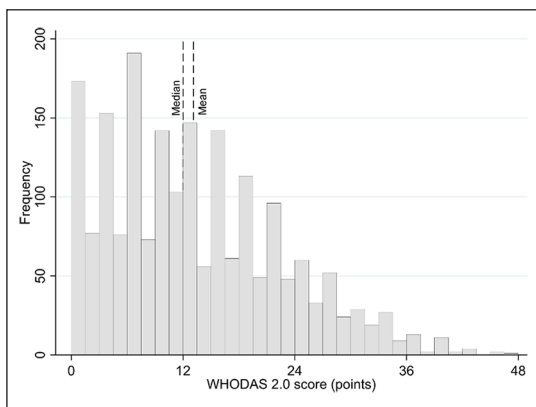
Of 3150 patients visiting the clinic, 1988 (63%) participated the study. The patients were 47.6 (6.3) years old and 1297 (65%) were women. The average intensity of pain was 6.3 (2.0) points on a numeric rating scale. The general health median was 1 (range 0 to 4, IQR 1 to 2) (Table 1). The majority of the patients were referred to the clinic due to non-specific chronic pain in their low back, neck, extremities, or soft tissue in general. Due to the national guidelines, patients with rheumatoid arthritis, severe osteoarthritis, or fractures were referred to other specialized clinics. Thus, only one patient had a main diagnosis of rheumatoid arthritis, 0.3% had diagnoses of traumas, and 2% had diagnoses of primary osteoarthritis. Most of the patients ($n = 1746$, 88%) had a main diagnosis “M” – “Diseases of the musculoskeletal system and connective tissue.” The most frequent single diagnoses were “M54 Dorsalgia” ($n = 781$, 39%) and “M79 Other soft tissue disorders” ($n = 202$, 10%). The patients’ characteristics are presented in Table 1.

The distribution of WHODAS 2.0 total score was abnormal with shift towards mild disability

Table 1. Demographic characteristics and the WHODAS 2.0 total score.

Variable	Total
Age (mean and standard deviation), years	47.6 (6.3)
Body mass index (mean and standard deviation), kg/cm ²	27.4 (5.7)
Pain (mean and standard deviation), points	6.3 (2.0)
Educational level (absolute proportions and percentage)	
No further education	1258 (67%)
Further education or higher	609 (33%)
WHODAS 2.0 (mean and standard deviation), points	13.1 (9.4)
WHODAS 2.0 (median, range, and interquartile range [IQR]), points	12 (0 to 48, IQR 6 to 19)

WHODAS: World health Organization Disability Assessment Schedule.

**Figure 1.** Histogram of the WHODAS 2.0 total score distribution.

WHODAS: World Health Organization Disability Assessment Schedule.

levels (Figure 1). However, the median and mean estimates were alike and thus, the distribution was considered close to normal enough to proceed with calculations based on mean and standard deviation. The Cronbach's alpha was good 0.89. The mean WHODAS 2.0 total score was 13.1 (9.4) and median 12 (Inter quartile range 6–19, range 0–48) points. Based on three different calculation formulas, the minimal clinically important difference estimates for the WHODAS 2.0 were 3.10 (calculated as standard error of measurement), 3.09 (calculated as one third of standard deviation), to 4.68 (calculated as half of standard deviation) points. The minimal detectable change was 8.6 points exceeding the level of minimal clinically important

difference and minimal detectable percentage change was unacceptably high 66%.

Discussion

Amongst almost 2000 patients with chronic musculoskeletal pain, the 12-item WHODAS 2.0 total score demonstrated a minimal clinically important difference of three to five points with minimal detectable change of up to 9 points exceeding a minimal clinically important difference almost twice. The minimal detectable % change showed that almost 70% change in WHODAS 2.0 total score should be expected before patients or clinicians might detect the change clinically.

While the large study sample advocates for the trustworthiness of the findings, the generalization of the results may be compromised by the study's cross-sectional design. Indeed, there were not longitudinal data to re-check the estimates by using an anchor-based approach with patients' real responses on the changed clinical situation. It has to be kept in mind that minimal clinically important difference and minimal detectable change are always statistical approximations, which could be different in real-world circumstances. The WHODAS 2.0 scores were distributed abnormally in the studied sample with most of the patients perceived only mild limitations of functioning. Therefore, caution is needed when generalizing the results amongst populations with more severe limitations.

Consistent with the results of this study, a study by Silva et al. has recently set the minimal detectable change of 12-item WHODAS in institutionalized

elderly to 9.6 points.¹⁰ Respectively, the size of minimal clinically important difference seen in the present study was similar to the estimates reported previously by a study amongst patients with anxiety and stress disorders.⁹ The results indirectly support previous reports on the potential unreliability of WHODAS 2.0 total score due to multidimensionality and a significant floor effect. A recent review suggested that 12-item WHODAS 2.0 is a multidimensional scale and it might be more useful when used to create a functioning profile than when providing a single total score.² A substantial floor effect of the 12-item WHODAS 2.0 has been seen in two studies.^{21,22} All these previous findings may explain the high estimates of minimal clinically important difference and minimal detectable change seen in the present study.

Further research in different populations is recommended. To confirm the results by using an anchor-based approach, longitudinal design is needed. The WHODAS 2.0 can be scored using a simple addition of individual items' scores (used in this study) or a more complex scheme taking into account the weights of different domains included into the WHODAS 2.0. Using that second scheme might affect the observed estimates and this possibility could be investigated by further research.

Clinical Messages

- Amongst patients with chronic musculoskeletal pain, the 12-item WHODAS 2.0 demonstrated a high minimal detectable change of almost nine points.
- As the minimal detectable change exceeded the level of minimal clinically important difference, nine points were considered to be the amount of change perceived by a respondent as being clinically significant.

Acknowledgements

None to declare

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Niina Katajapuu  <https://orcid.org/0000-0002-7416-8928>

Mikhail Saltychev  <https://orcid.org/0000-0003-1269-4743>

References

1. Üstün TB, Kostanjsek N and Chatterji S. *Measuring Health and Disability: Manual for WHO Disability Assessment Schedule (WHODAS 2.0)*. Geneva: World Health Organization, 2010.
2. Saltychev M, Katajapuu N, Barlund E, et al. 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. *Disabil Rehabil* 2019; 1–6.
3. Carozzi NE, Kratz AL, Downing NR, et al. Validity of the 12-item World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) in individuals with Huntington disease (HD). *Qual Life Res* 2015; 24: 1963–1971.
4. Von Korff M, Katon WJ, Lin EHB, et al. Functional outcomes of multi-condition collaborative care and successful ageing: results of randomised trial. *BMJ* 2011; 343: d6612–d6612.
5. Ferraz DD, Trippo KV, Duarte GP, et al. The effects of functional training, bicycle exercise, and exergaming on walking capacity of elderly patients with Parkinson disease: a pilot randomized controlled single-blinded trial. *Arch Phys Med Rehabil* 2018; 99: 826–833.
6. Guyatt G, Walter S and Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987; 40(2): 171–178.
7. Beaton DE, Bombardier C, Katz JN, et al. A taxonomy for responsiveness. *J Clin Epidem* 2001; 54: 1204–1217.
8. de Vet HCW and Terwee CB. The minimal detectable change should not replace the minimal important difference. *J Clin Epidem* 2010; 63(7): 804–805.
9. Axelsson E, Lindsater E, Ljotsson B, et al. The 12-item self-report World Health Organization Disability Assessment Schedule (WHODAS) 2.0 administered via the Internet to individuals with anxiety and stress disorders: a psychometric investigation based on data from two clinical trials. *JMIR Ment Health* 2017; 4: e58.
10. Silva AG, Cerqueira M, Raquel Santos A, et al. Inter-rater reliability, standard error of measurement and minimal detectable change of the 12-item WHODAS 2.0 and four performance tests in institutionalized ambulatory older adults. *Disabil Rehabil* 2019; 41(3): 366–373.

11. Saltychev M, Mattie R, McCormick Z, et al. Confirmatory factor analysis of 12-Item World Health Organization Disability Assessment Schedule in patients with musculoskeletal pain conditions. *Clin Rehabil* 2017; 31(5): 702–709.
12. Kohn CG, Sidovar MF, Kaur K, et al. Estimating a minimal clinically important difference for the EuroQol 5-Dimension health status index in persons with multiple sclerosis. *Health Qual Life Outcomes* 2014; 12: 66.
13. Gilbert C, Brown MCJ, Cappelleri JC, et al. Estimating a minimally important difference in pulmonary arterial hypertension following treatment with sildenafil. *Chest* 2009; 135(1): 137–142.
14. Guyatt GH, Osoba D, Wu AW, et al. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc* 2002; 77(4): 371–383.
15. Le QA, Doctor JN, Zoellner LA, et al. Minimal clinically important differences for the EQ-5D and QWB-SA in Post-traumatic Stress Disorder (PTSD): results from a Doubly Randomized Preference Trial (DRPT). *Health Qual Life Outcomes* 2013; 11: 59.
16. Pickard AS, Neary MP and Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. *Health Qual Life Outcomes* 2007; 5: 70.
17. Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008; 61(2): 102–109.
18. Shiklar R, Harding G, Leahy M, et al. Minimal important difference (MID) of the Dermatology Life Quality Index (DLQI): results from patients with chronic idiopathic urticaria. *Health Qual Life Outcomes* 2005; 3: 36.
19. Smidt N, van der Windt DA, Assendelft WJ, et al. Interobserver reproducibility of the assessment of severity of complaints, grip strength, and pressure pain threshold in patients with lateral epicondylitis. *Arch Phys Med Rehabil* 2002; 83(8): 1145–1150.
20. Chen CH, Lin SF, Yu WH, et al. Comparison of the test-retest reliability of the balance computerized adaptive test and a computerized posturography instrument in patients with stroke. *Arch Phys Med Rehabil* 2014; 95(8): 1477–1483.
21. Katajapuu N, Laimi K, Heinonen A, et al. Floor and ceiling effects of the World Health Organization Disability

Assessment Schedule 2.0 among patients with chronic musculoskeletal pain. *Int J Rehabil Res* 2019; 42(2): 190–192.

22. Gaskin CJ, Lambert SD, Bowe SJ, et al. Why sample selection matters in exploratory factor analysis: implications for the 12-item World Health Organization Disability Assessment Schedule 2.0. *BMC Med Res Methodol* 2017; 17: 40.

Appendix A. The 12-item WHODAS 2.0

In the past 30 days, how much difficulty did you have in:

1. Standing for long periods such as 30 minutes?
2. Taking care of your household responsibilities?
3. Learning a new task, for example, learning how to get to a new place?
4. How much of a problem did you have in joining in community activities (e.g. festivities, religious or other activities) in the same way as anyone else can?
5. How much have you been emotionally affected by your health problems?
6. Concentrating on doing something for 10 minutes?
7. Walking a long distance such as a kilometer (or equivalent)?
8. Washing your whole body?
9. Getting dressed?
10. Dealing with people you do not know?
11. Maintaining a friendship?
12. Your day-to-day work?

(0) None; (1) Mild; (2) Moderate; (3) Severe; (4) Extreme/Cannot do