



Validation of the Brazilian Version of WHODAS 2.0 in Patients with Mental Disorders: Should the 12-Item Scale be an Alternative to 36-Item Scale in DSM-5?

Fábio Brasil^{1,†}, Andreia MB Brasil², Cassyano J Correr¹

Abstract

The World Health Organization Disability Assessment Schedule (WHODAS) 2.0 Brazilian version was self-applied in 500 psychiatric and 50 non-psychiatric patients in an outpatient medical centre. The instrument was reapplied in 50 psychiatric patients whose treatment with psychotropics was not altered, and in 100 whose it was altered. Confirmatory factor analysis confirmed the theoretical dimensional model proposed by the WHO for the WHODAS 2.0 36-item scale, and a simpler alternative model for the WHODAS 2.0 12-item scale. All the items of the WHODAS 2.0 was proven to be valid and reproducible for use in people with mental disorders from Brazil. The WHODAS 2.0 scores differentiate psychiatric from non-psychiatric patients and exhibited a significant correlation with the Global Assessment of Functioning scale. However, the 12-item scale was more responsive to changes in the patients' treatment than the 36-item scale, and this should be considered for Diagnostic and Statistical Manual of Mental Disorders.

Keywords

Quality of life, Disability evaluation, Psychometrics, World health organization, Diagnostic and statistical manual of mental disorders

Introduction

Diagnosing a mental disorder is just as important as measuring the impact it will have on the life on the patient. To do this, the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) opted for the Global Assessment of Functioning scale (GAF), which measures the general psychofunctionality in scale from 0 (worst) to 100 (best) using integers [1], although the measurement obtained by the GAF is biased because it is limited only to the evaluator's opinion, as well as every clinician-reported outcome (ClinRO) measure [2]. Patient-reported outcome (PRO) measures do not exhibit this bias, since they are filled out by the patients themselves (self-applied)

or are applied during interviews; however they are limited only to the patient's opinion [3]. Therefore, it is important to note that all these instruments are subjective.

The World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) is the official PRO measure of the World Health Organization (WHO) for measuring functionality and incapacity related to any disease or state of health [4]. The WHODAS 2.0 full questionnaire entails 36 items divided into six domains (D1 to D6) which are responded to using a five-point Likert scale according to the degree of difficulty in carrying out certain everyday activities: from 1 ("no difficulty") to 5 ("extreme difficulty") [5]. The WHODAS

¹Post-Graduate Program in Pharmaceutical Sciences, Federal University of Paraná (UFPR), Curitiba, PR, Brazil

²Post-Graduate Program in Feeding and Nutrition, Federal University of Paraná (UFPR), Curitiba, PR, Brazil

[†]Author for correspondence: Fábio Brasil, Setor de Ciências da Saúde, Universidade Federal do Paraná, Av. Lothário Meissner 632, Curitiba, PR, Brazil, Tel/Fax: 55-41-33604098, email: tenfabiobrasil@gmail.com

2.0 abbreviated questionnaire uses 12 items selected from the full instrument, two from each domain, and represents 81% of the variance of the whole 36-item scale [4] (the literature advises 75% or far above for abbreviated versions of an instrument [6]); it has been recommended for situations where time or funding is limited [5]. The fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), current world reference for diagnosing in psychiatry, included the WHODAS 2.0 36-item scale in replacement of the GAF [7].

The WHODAS 2.0 has been validated for Brazil in reproductive-aged women, but not in the male population [8]. Furthermore, people with mental illnesses exhibit unique and pathological latent psychological traits which justify the need for an independent validation process for this special population [9]. The objective of this study was to validate the official Brazilian Portuguese version of the WHODAS 2.0 in men and women with mental disorders, to verify its precision, and to compare the responsiveness to clinical changes between the 36-item scale, the 12-item scale, and the GAF.

Methods

This observational study was conducted in a private medical centre in the city of Curitiba, Paraná, Brazil. This medical centre was chosen because it has an ample outpatient clinic for general psychiatry that works alongside an outpatient clinic that treats chronic degenerative diseases.

Patients typically schedule their visits five to ten weeks in advance, so random selection of patients needed to be conducted in two steps: half of the sample was selected between February 1, 2016 and March 31, 2016, and the other half between April 1, 2016 and May 31, 2016.

The survey was conducted in accordance with the norms established in the Helsinki Declaration and was approved by the Universidade Federal do Paraná Ethics Committee under CAAE record number 50937015.4.0000.0102.

■ Psychometric instruments

Authorization was requested from the WHO to use the official Brazilian Portuguese version of the WHODAS 2.0, published by the Universidade Federal do Triângulo Mineiro [5], which was applied to all the participants. The scores were calculated by the individual summary of items and subsequent conversion to a scale of 0 to 100.

The same procedure was conducted to calculate the scores relating to the set of 12 items that comprise the abbreviated questionnaire.

The GAF scores, which were calculated for all the psychiatric participants by their own attending doctors, were transcribed in an inverse scale to facilitate comparison with the WHODAS 2.0 scores.

■ Psychiatric participants

The ideal sample for factor analysis is equal to or greater than 10 subjects per questionnaire item [10]. Of the 4,112 different patients (2,796 women) above 18 years of age with appointments scheduled for the psychiatric outpatient clinic during the study period, 250 women and 250 men were randomly chosen. When a selected individual did not agree to participate in the study, did not appear for his or her appointment, was functionally illiterate, or was not using psychotropic medications, a new participant was chosen from the remaining patients of the same sex as a replacement.

The recommended sample size for reproducibility and responsiveness assessment is at least 50 subjects [10]. The first 25 women and 25 men who did not have their psychopharmaceutical treatment altered during their appointment were invited to return to the clinic one week later and respond to the WHODAS 2.0 again; if they did not appear, the next patient was asked, maintaining gender equality among the participants. Using the same criteria, the first 50 women and 50 men whose psychopharmaceutical treatment was quantitatively or qualitatively altered during their appointments were invited to return to the clinic five weeks later for another appointment; in this case, they responded to the WHODAS 2.0, and the GAF was recalculated by the attending physician.

■ Non-psychiatric participants

In validation studies, the control groups should be composed of a minimum of 50 subjects [10]. Of the 2,068 different patients (1,530 women) above 18 years of age with appointments scheduled for the chronic diseases outpatient clinic during the study period, 25 women and 25 men were randomly chosen. When the selected individual did not agree to participate in the study, did not appear for his or her appointment, was functionally illiterate, or was using psychopharmaceuticals, a new participant was chosen from the remaining patients of the same sex as a replacement.

■ **Statistical analysis**

To verify the discriminant validity in relation to mental health, the WHODAS 2.0 scores for the group of psychiatric patients and the group of non-psychiatric patients were compared using Student's t test. With the goal of instituting the convergent validity, Pearson's correlation coefficients were verified between the WHODAS 2.0 scores and GAF scores.

To establish the validity of the construct in the psychiatric population, confirmatory factor analysis (CFA) was conducted using the maximum likelihood method, in which the items in the WHODAS 2.0 were tested with regard to the factorial model at two hierarchical levels proposed by the WHO: the first relating to general incapacity and the second related to the distribution of the six theoretical domains [4]. An alternative model was proposed for the 12-item scale. The values considered acceptable for the fit indexes were as follows: $0 \leq \chi^2/\text{df} \leq 3$ with $0.01 \leq p \leq 1$; $0 \leq \text{standardized root mean square residual (SRMR)} \leq 0.1$; $0 \leq \text{root mean square error of approximation (RMSEA)}$ and confidence interval of 90% (CI 90%) ≤ 0.08 ; $0.95 \leq \text{comparative fit index (CFI)} \leq 1$ [11]. Factor loadings below 0.5 indicated the need to exclude the corresponding item and repeat the CFA [12].

Next, to complete validation of the construct, internal consistency was verified using Cronbach's alpha coefficient for the WHODAS 2.0 scales as a whole and for its individual domains, with values above 0.7 [10].

The reproducibility (precision) of individual items was tested in those patients for whom the WHODAS 2.0 was reapplied without changes in their treatment involving psychoactive medications; this was done by calculating intraclass correlation coefficients (ICC) using the two-way random effects model with absolute agreement, the values of which should exceed the cutoff point of 0.7 [10]. Responsiveness (accuracy) was tested comparatively with the GAF by calculating the areas under the receiver operating characteristics (ROC) curves in patients who completed the WHODAS 2.0 a second time after changes in psychoactive drug treatment; in this case the WHODAS 2.0 scores were inverted to result in convex curves.

The statistical calculations, including the generation of random numbers for simple random selection, were carried out using

Statistical Package for the Social Sciences software version 21.0 with the Analysis of Moment Structures extension.

Results

Of the 500 selected psychiatric patients, 48 (20 females) were excluded from the survey because after they began, they were unable to answer all the items in the WHODAS 2.0; the reasons were as follows: hurry (n=16), disinterest (n=13), weariness (n=8), nervousness (n=6), and confusion (n=5). All 50 selected non-psychiatric patients were able to complete the questionnaire in its entirety. The withdrawals by psychiatric patients were attributed to their psychic fragility, because both groups were approached in the same way by a single trained health staff.

The principal diagnoses, codified by the 10th revision of the International Classification of Diseases [13], of the psychiatric patients were as follows: F01-F09 (n=9), F10-F19 (n=51), F20-F29 (n=29), F30-F39 (n=108), F40-F48 (n=136), F50-F59 (n=53), F60-F69 (n=44), F84 (n=6), F90-F98 (n=16); and of the non-psychiatric patients were as follows: B15-B19 (n=2), D50-D89 (n=4), E00-E89 (n=8), I00-I99 (n=10), J00-J99 (n=6), K00-K95 (n=5), L00-L99 (n=3), M00-M99 (n=8), N00-N99 (n=4). The descriptive variables of the total population of the study, including the scores for the psychometric instruments (which showed normal distribution according to the Kolmogorov-Smirnov test), are presented in **Table 1**.

Both the 36-item total scores and those scores calculated considering only the 12 items that integrate the abridged WHODAS 2.0 questionnaire were able to distinguish the psychiatric patients from the non-psychiatric patients (p<0.01) and women from men (p<0.01) using the t-test. The WHODAS 2.0 scores of the 36-item scale yielded, in the group of psychiatric patients in modulus, a Pearson correlation coefficient of 0.88 (p<0.01) with the scores of the 12-item scale, and 0.59 (p<0.01) with the GAF scores. For the 12-item scale, scores had a correlation of 0.66 (p<0.01) with the scores for the GAF.

CFA proved the factorial structure proposed by the WHO for the WHODAS 2.0 36-item scale for the psychiatric population (**Figure 1**), with satisfactory values for the fit indices: $\chi^2/\text{df}=1.68$ (p=0.03); SRMR=0.05; RMSEA=0.04 (CI 90% 0.02-0.05); CFI=0.97. The factor loadings

Table 1: Descriptive variables of the sample population.

Variables	Psychiatric outpatients (n = 452; 230♀)	Non-psychiatric outpatients (n = 50; 25♀)
Age (years)	40.03 (14.72) (18-78)♀ 41.23 (13.97) (18-74)♂	41.76 (15.18) (21-79)♀ 43.60 (17.11) (18-76)♂
Study time (years)	12 (3-24)♀ 11 (3-23)♂	12 (4-21)♀ 12 (8-17)♂
36-item WHODAS 2.0 scores (scale from 0 to 100)	33.75 (22.09) (1.39-76.39)♀ 23.53 (17.62) (1.39-75.69)♂	38.69 (9.00) (25.69-56.94)♀ 28.53 (8.11) (18.06-49.31)♂
12-item WHODAS 2.0 scores (scale from 0 to 100)	33.28 (21.56) (2.08-89.58)♀ 25.44 (16.47) (2.08-85.42)♂	37.42 (16.72) (16.67-79.17)♀ 32.25 (11.95) (12.50-66.67)♂
GAF scores (inverted scale from 0 to 100)	41.00 (26.26) (10-95)♀ 35.97 (22.90) (10-95)♂	-

The variables age and scores are described as average; standard deviation, minimum and maximum range in parentheses: μ (σ) (min-max).
 The variable years of study is presented as median; minimum and maximum range in parentheses: med (min-max).
 ♀ = Women; ♂ = Men; WHODAS = World Health Organization Disability Assessment Schedule; GAF = Global Assessment of Functioning.

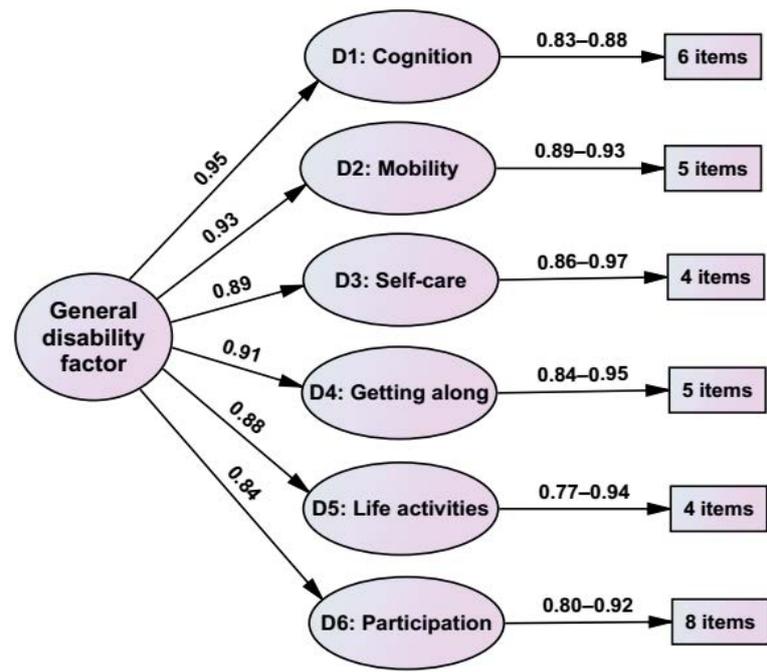


Figure 1: Confirmatory factor analysis of the Brazilian version of the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0).

ranged from 0.77 to 0.97, so there was no need to delete items (Table 2). The alpha coefficients of the WHODAS 2.0 36-item scale domains ranged from 0.76 to 0.88, total=0.92.
 The two-level hierarchical factorial structure used to validate the WHODAS 2.0 full questionnaire

(Figure 1) and the factorial structure that provides single-construct model were not satisfactory when applied to the 12-item set comprising the abbreviated questionnaire. However, the alternative construct model in that the 12 items are divided into their six respective domains, and these domains are all correlated

Table 2: Factor loadings and reproducibility of the items of the World Health Organization Disability Assessment Schedule 2.0 in Brazilian psychiatric outpatients.

Items	Factor loadings		Intraclass correlation coefficients
	36-item scale	12-item scale	
D1.1: Concentrating on doing something for ten minutes?	0.88	0.92	0.81
D1.2: Remembering to do important things?	0.85	–	0.79
D1.3: Analysing and finding solutions to problems in day-to-day life?	0.84	–	0.74
D1.4: Learning a new task, for example, learning how to get to a new place?	0.87	0.88	0.82
D1.5: Generally understanding what people say?	0.84	–	0.76
D1.6: Starting and maintaining a conversation?	0.83	–	0.75
D2.1: Standing for long periods such as 30 minutes?	0.93	0.95	0.86
D2.2: Standing up from sitting down?	0.90	–	0.84
D2.3: Moving around inside your home?	0.89	–	0.82
D2.4: Getting out of your home?	0.89	–	0.78
D2.5: Walking a long distance such as a kilometre [or equivalent]?	0.91	0.92	0.85
D3.1: Washing your whole body?	0.97	0.92	0.86
D3.2: Getting dressed?	0.90	0.96	0.84
D3.3: Eating?	0.87	–	0.80
D3.4: Staying by yourself for a few days?	0.86	–	0.75
D4.1: Dealing with people you do not know?	0.87	0.91	0.78
D4.2: Maintaining a friendship?	0.95	0.93	0.82
D4.3: Getting along with people who are close to you?	0.86	–	0.76
D4.4: Making new friends?	0.93	–	0.85
D4.5: Sexual activities?	0.84	–	0.82
D5.1: Taking care of your household responsibilities?	0.80	0.84	0.77
D5.2: Doing most important household tasks well?	0.78	–	0.79
D5.3: Getting all the household work done that you needed to do?	0.80	–	0.81
D5.4: Getting your household work done as quickly as needed?	0.77	–	0.75
D5.5: Your day-to-day work/school?	0.94	0.89	0.75
D5.6: Doing your most important work/school tasks well?	0.94	–	0.77
D5.7: Getting all the work done that you need to do?	0.85	–	0.83
D5.8: Getting your work done as quickly as needed?	0.83	–	0.79
D6.1: How much of a problem did you have in joining in community activities (for example, festivities, religious or other activities) in the same way as anyone else can?	0.83	0.79	0.74
D6.2: How much of a problem did you have because of barriers or hindrances in the world around you?	0.81	–	0.75
D6.3: How much of a problem did you have living with dignity because of the attitudes and actions of others?	0.80	–	0.78
D6.4: How much time did you spend on your health condition, or its consequences?	0.80	–	0.76
D6.5: How much have you been emotionally affected by your health condition?	0.87	0.95	0.81
D6.6: How much has your health been a drain on the financial resources of you or your family?	0.91	–	0.84
D6.7: How much of a problem did your family have because of your health problems?	0.92	–	0.83
D6.8: How much of a problem did you have in doing things by yourself for relaxation or pleasure?	0.82	–	0.80

Confirmatory factor analysis by the maximum likelihood method.

Intraclass correlation coefficients by the two-way random effects model (interval of one week).

The items were reproduced from: <http://www.who.int/classifications/icf/whodasii/en/> [Accessed November 15, 2017].

within a single hierarchical level, presented adequate values for the fit indexes: $\chi^2/df=1.84$ ($p=0.08$); SRMR=0.04; RMSEA=0.02 (IC 90% 0.01-0.03); CFI=0.98. The factor loadings of the WHODAS 2.0 12-item scale ranged from 0.79 to 0.96 (Table 2) and the alpha coefficients of the domains ranged from 0.75 to 0.83; total=0.85.

The individual ICCs for the items in the WHODAS 2.0 between application and reapplication without change in treatment ranged from 0.74 to 0.86 (Table 2). The areas under the ROC curve, between application and reapplication with change in treatment were as follows for the psychometric scales: 36-item WHODAS 2.0=0.63 ($p<0.01$); 12-item WHODAS 2.0=0.70 ($p<0.01$); GAF=0.80 ($p<0.01$) (Figure 2).

Discussion

The patients who refused to participate or dropped out of the study tend to be different from those who agreed and remained [12-14]; however this situation should not be considered as selection bias, because the viability and acceptability of the instrument determine its

clinical universe [15,16]. Based on the results, it was seen that the Brazilian Portuguese version of the WHODAS 2.0 PRO measure met all the statistical requirements to be considered valid and reliable (reproducible and responsive), and can be used safely in Brazil in patients with mental disorders and can even distinguish them from patients with non-psychiatric illnesses. Validation of the construct was particularly successful in the 36-item scale, since the CFA confirmed the theoretical factorial structure proposed by the WHO [4] (Figure 1); consequently, the procedure also determined the validity of the content. The 12-item scale did not follow the factorial structure suggested by the WHO, however a simpler empirical construct model that also correctly addresses the items to their respective domains was confirmed in the population sample, and this fact validated the scale. The difference between both structures is due to the absence of the items with lower factor loadings in the abbreviated instrument (Table 2), which strengthens the six domains or specific factors (second level), resulting in the weakening of the general factor (first level) until its dissipation. At least 810 studies were

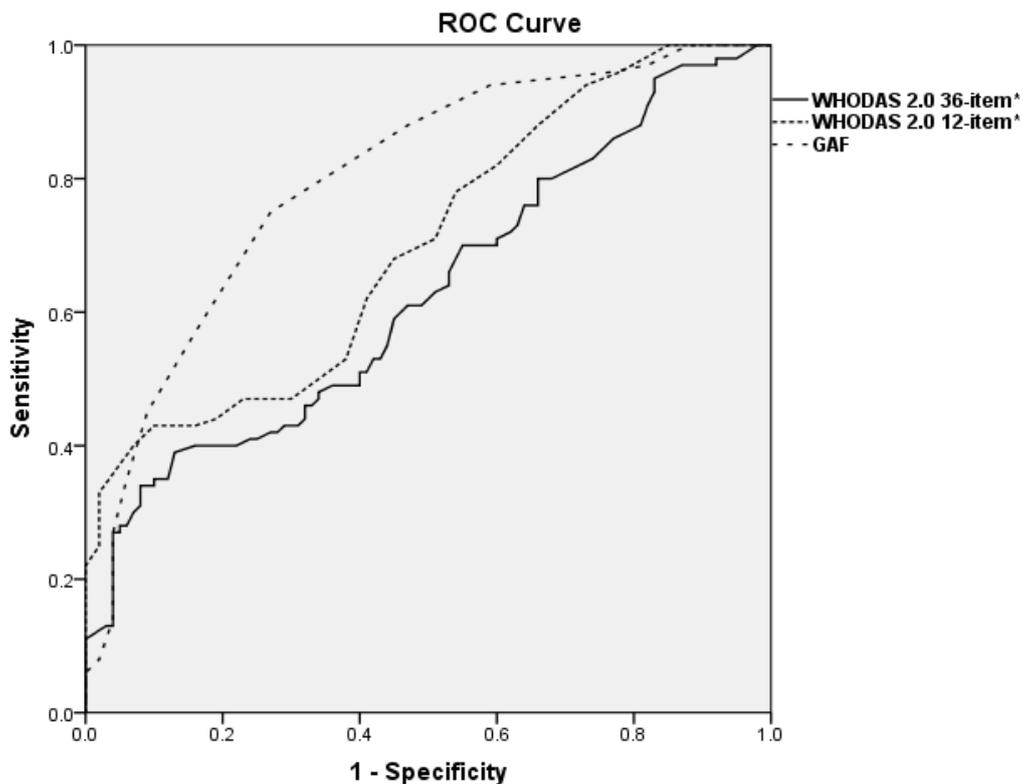


Figure 2: Responsiveness to changing treatment with psychopharmaceuticals in psychiatric patients.

*Inverted scale; ROC = Receiver Operating Characteristics; WHODAS = World Health Organization Disability Assessment Schedule; GAF = Global Assessment of Functioning.

conducted in 94 countries using WHODAS 2.0, about 40% in psychiatry, they demonstrated heterogeneous factorial structures, since the number of factors irregularly ranged from 1 to 7 for both the 36-item and the 12-item scale; however none of them evaluated the differences between the two versions regarding the responsiveness to the treatment of psychiatric patients [8,17].

Although the GAF scores are considered reliable, even when established by different evaluators (as long as they are experienced) [18], this study found that the responsiveness of the GAF was too convex and linear (Figure 2), which indicates a subjective tendency among the physician evaluators to overvalue the clinical improvement of their patients when treatment is modified. In situations like these, the subjectivity of the evaluator is always unwelcome, whilst patient's subjectivity becomes welcome to the extent that it represents their own aspiration for improvement and well-being [3,19]; moreover, ClinRO measures cannot directly measure symptoms, which are experienced just by patients [2]. Considering this perspective, the editors of the DSM-5 were correct in replacing the GAF with the WHODAS 2.0 as their gold standard for assessing patients' functionality [1,7], this replacement was the theoretical reason for verifying convergent validity between a PRO and a ClinRO measure, which might seem unprofessional. Nevertheless, recklessly, there is no mention in translated versions of the manual (such as the Brazilian version) [20] that the WHODAS 2.0 should be validated in the target culture to be used appropriately.

The absolute scores for the PRO measures are often not interpretable; what really matters

for individual assessment of the patient or the outcomes of clinical studies is the responsiveness triggered by an intervention or an established treatment, in other words, the changes in the scores [21,22]. In this regard, the WHODAS 2.0 12-item scale was shown in this study to be superior to the 36-item scale (Figure 2), suggesting that it is more appropriate for evaluations of the efficacy of psychotropic drugs and for pharmacoeconomic studies. Moreover, the abbreviated questionnaire had a stronger correlation with the GAF and has the advantage of rapid application and is less prone to errors resulting from fatigue or inattention in the respondents. Therefore, the editors of the DSM-5 should consider the WHODAS 2.0 12-item as a sensible alternative to the WHODAS 2.0 36-item scale in new versions of this manual. Furthermore, despite isolated studies with people aged 15 to 17 years [23-25] and the release of an unofficial version directed at children and adolescents [26], one limitation of the WHODAS 2.0 is the fact that it is only recommended by the WHO for application in adults [4,5] and, consequently, use of the GAF should not yet be discontinued.

Acknowledgements

We wish to thank the patients and staff at the Instituto Setti.

Funding

No financial support was received.

Conflict of Interest

The authors have no conflicts of interest.

References

- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders – Text Revision: DSM-IV-TR. American Psychiatric Publishing, Arlington, USA (2000).
- Powers JH, Patrick DL, Walton MK, et al. Clinician-reported outcome assessments of treatment benefit: report of the ISPOR Clinical Outcome Assessment Emerging Good Practices Task Force. *Value. Health* 20(1), 2-14 (2017).
- Snyder CF, Aaronson NK, Choucair AK, et al. Implementing patient-reported outcomes assessment in clinical practice: A review of the options and considerations. *Qual. Life. Res* 21(8), 1305-1314 (2012).
- Üstün TB, Chatterji S, Kostanjsek N, et al. Developing the World Health Organization Disability Assessment Schedule 2.0. *Bull. World. Health. Organ* 88(11), 815-823 (2010).
- World Health Organization. Avaliação de Saúde e Deficiência: Manual do WHO Disability Assessment Schedule (WHODAS 2.0). Universidade Federal do Triângulo Mineiro, Uberaba, BRA (2015).
- Hagtvet K, Sipos K. Creating short forms for construct measures: The role of exchangeable forms. *Pedagogika* 66(6), 689-713 (2016).
- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: DSM-5. American Psychiatric Publishing, Arlington, USA (2013).
- Silveira C, Parpinelli MA, Pacagnella RC, et al. Validation of the 36-item version of the WHO Disability Assessment Schedule 2.0 (WHODAS 2.0) for assessing women's disability and functioning associated with maternal morbidity. *Rev. Bras. Ginecol. Obstet* 39(2), 44-52 (2017).
- Salvador-Carulla L. Routine outcome assessment in mental health research. *Curr. Opin. Psychiatry* 12(2), 207-210 (1999).
- Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J. Clin. Epidemiol* 60(1), 34-42 (2007).
- Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation

- models: test of significance and descriptive goodness-of-fit measures. *Meth. Psychol. Res* 8(2), 23-74 (2003).
12. Costello AB, Osborne JW. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Pract. Assess. Res. Eval* 10(7), 1-9 (2005).
 13. World Health Organization. International Statistical Classification of Diseases and Related Health Problems 10th Revision. World Health Organization, Geneva, CHE (2016).
 14. Nohr EA, Frydenberg M, Henriksen TB, et al. Does low participation in cohort studies induce bias? *Epidemiology* 17(4), 413-418 (2006).
 15. Lewis CC, Fischer S, Weiner BJ, et al. Outcomes for implementation science: An enhanced systematic review of instruments using evidence-based rating criteria. *Implement. Sci* 10(155), 1-17 (2015).
 16. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: A quarter century on. *Lancet* 390(10092), 22-28 (2017).
 17. Federici S, Bracalenti M, Meloni F, et al. World Health Organization disability assessment schedule 2.0: An international systematic review. *Disabil. Rehabil* 39(23), 2347-2380 (2017).
 18. Pedersen G, Hagtvet KA, Karterud S. Generalizability studies of the Global Assessment of Functioning—Split version. *Compr. Psychiatry* 48(1), 88-94 (2007).
 19. Say R, Murtagh M, Thomson R. Patients' preference for involvement in medical decision making: a narrative review. *Patient. Educ. Couns* 60(2), 102-114 (2006).
 20. American Psychiatric Association. Manual Diagnóstico e Estatístico de Transtornos Mentais: DSM-5. Artmed, Porto Alegre, BRA (2014).
 21. Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J. Clin. Epidemiol* 61(2), 102-109 (2008).
 22. Brasil F, Pontarolo R, Correr CJ. Patient reported outcomes measures (PROMs) in diabetes: Why are they still rarely used in clinical routine? *Diabetes Res. Clin. Pract* 97(1), e4-e5 (2012).
 23. Andrews G, Kemp A, Sunderland M, et al. Normative data for the 12 item WHO Disability Assessment Schedule 2.0. *PLoS One* 4(12), e8343 (2009).
 24. Betancourt TS, McBain R, Newnham EA, et al. A behavioral intervention for war-affected youth in Sierra Leone: A randomized controlled trial. *J. Am. Acad. Child Adolesc. Psychiatry* 53(12), 1288-1297 (2014).
 25. Kimber M, Rehm J, Ferro MA. Measurement invariance of the WHODAS 2.0 in a population-based sample of youth. *PLoS One* 10(11), e142385 (2015).
 26. Scorza P, Stevenson A, Canino G, et al. Validation of the "World Health Organization Disability Assessment Schedule for Children, WHODAS-Child" in Rwanda. *PLoS One* 8(3), e57725 (2013).