

Psychometric properties of the World Health Organization Disability Assessment Schedule used in the European Study of the Epidemiology of Mental Disorders

M. A. BUIST-BOUWMAN,^{1,2} J. ORMEL,¹ R. DE GRAAF,² G. VILAGUT,³ J. ALONSO,³
E. VAN SONDEREN,⁴ W. A. M. VOLLEBERGH,² AND THE ESEMED/MHEDEA 2000
INVESTIGATORS*

- 1 Graduate Research School of Behavior, Cognition and Neurosciences (BCN), University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
- 2 Netherlands Institute of Mental Health and Addiction, Utrecht, The Netherlands
- 3 Institut Municipal d'Investigació Mèdica (IMIM), Barcelona, Spain
- 4 Northern Centre for Healthcare Research (NCH), Groningen, The Netherlands

Abstract

This study assessed the factor structure, internal consistency, and discriminatory validity of the World Health Organization Disability Assessment Schedule (WHODAS) version used in the European Study of the Epidemiology of Mental Disorders (ESEMEd). In total 8796 adults were assessed using the ESEMEd WHODAS (22 severity and 8 frequency items). An Exploratory Factor Analysis (EFA) with promax rotation was done with a random 50%. The other half was used for confirmatory factor analysis (CFA) comparing models (a) suggested by EFA; (b) hypothesized a priori; and (c) reduced with four items. A CFA model with covariates was conducted in the whole sample to assess invariance across Mediterranean (Spain, France and Italy) and non-Mediterranean (Belgium, Germany and the Netherlands) countries. Cronbach's alphas and discriminatory validity were also examined. EFA identified seven factors (explained variance: 80%). The reduced model (six factors, four frequency items excluded) presented the best fit [Confirmatory Fit Index (CFI) = 0.992, Tucker–Lewis Index (TLI) = 0.996, Root Mean Square Error of Approximation (RMSEA) = 0.024]. The second-order factor structure also fitted well (CFI = 0.987, TLI = 0.991, RMSEA = 0.036). Measurement non-invariance was found for Embarrassment. Cronbach's alphas ranged from 0.84 for Participation to 0.93 for Mobility. Preliminary data suggest acceptable discriminatory validity. Thus, the ESEMEd WHODAS may well be a valuable shortened version of the WHODAS-II, but future users should reconsider the filter questions. Copyright © 2008 John Wiley & Sons, Ltd.

Key words: factor structure, internal consistency, validity, disability assessment schedule

*The ESEMEd/MHEDEA 2000 Investigators are: Jordi Alonso, Matthias Angermeyer, Sebastian Bernert, Ronny Bruffaerts, Traolach S. Brugha, Giovanni de Girolamo, Ron de Graaf, Koen Demyttenaere, Isabelle Gasquet, Josep Maria Haro, Steven J. Katz, Ronald C. Kessler, Viviane Kovess, Jean Pierre Lépine, Johan Ormel, Gabriella Polidori, and Gemma Vilagut.

Introduction

Due to improvements in public health, fewer diseases in higher-income countries are fatal nowadays (Mathers and Loncar, 2006). The drawback of this increased survival rate, however, is that an increasing number of

individuals experience limitations in their ability to function in everyday life (Mathers and Loncar, 2006). Apart from the difficulties at the individual level (Buist-Bouwman, 2006), diseases involving impairments lead to increased societal costs as well due to, for example, work loss (Goering et al., 1996; Kessler and Frank, 1997; Kouzis and Eaton, 1994), and the utilization of health and support services (Kouzis and Eaton, 1997).

To assess the disturbances in social adjustment and behavior in persons with mental and physical disorders, the World Health Organization Disability Assessment Schedule (WHODAS) was published in 1988 (<http://www.who.int/icidh/whodas/generalinfo.html>; assessed on December 30, 2007). In 1998, the WHODAS-II was launched which was a completely revised version of the WHODAS to reflect the WHO's current thinking about functioning and disability as described in the International Classification of Functioning, Disability and Health (ICF). In short, the ICF provides codes to describe the complete range of health states and the consequences of ill-health for both physical and mental disorders (WHO, 2001). Psychometric testing of the WHODAS-II suggest a two level hierarchical structure in which individual items load onto one of six domains which in turn load on a global disability latent variable. There is a growing number of studies evaluating the psychometric properties of the WHODAS-II (Chwastiak and Von Korff, 2003; Chisolm et al., 2005; Annicchiarico et al., 2004; Kim et al., 2005; Chopra et al., 2004; Gallagher and Mulvany, 2004; McKibbin et al., 2004; Van Tubergen et al., 2003), suggesting acceptable internal consistency, test-retest reliability, convergent validity and responsiveness to change.

Due to time constraints, a shortened version of the WHODAS-II was used in the European Study of the Epidemiology of Mental Disorders (ESEMEd) which sought to estimate the prevalence and associated disability using nationally representative, general population surveys in Europe. Compared to the original WHODAS-II, the ESEMEd WHODAS (1) is more time efficient, because less questions were asked and a filter question was employed for each multi item domain; (2) assesses overall role functioning (your normal daily activities) in stead of housework and employment separately; (3) employs one frequency item in each scale while the original WHODAS-II employs one frequency question per item.

This study examines the psychometric properties of the ESEMEd WHODAS. More specifically: (1) its

factor structure, (2) internal consistency of the disability factors identified; (3) the robustness of the factor structure across countries, and (4) its discriminative validity. The assessment of the psychometric properties of the ESEMEd WHODAS is important, because its strengths and weaknesses are essential in interpreting data in the studies it is and will be used in.

Methods

A complete description of the methods of the ESEMEd project is presented elsewhere (The ESEMEd investigators, 2002, 2004). Briefly, ESEMEd is a cross-sectional survey representative of the adult population of Belgium, France, Germany, Italy, the Netherlands, and Spain. In total 21 425 individuals aged 18 years and older, residing in private households, were interviewed between January 2001 and July 2003. The overall response rate of the study was 61.2%, ranging from 45.9% in France to 78.6% in Spain. The ethics committees in each participating country approved these procedures and informed consent was obtained from all respondents.

The survey interview

Screening section

The screening section (a lifetime psychiatric screening instrument), located at the beginning of the questionnaire, was administered to all 21 425 respondents (for detailed information about the questionnaire, see The ESEMEd investigators, 2004). It contains lifetime screening questions for some specific disorders of mood (i.e. depression and dysthymia) and anxiety (i.e. panic disorder, social phobia, specific phobia, agoraphobia and generalized anxiety depression). The complete list of psychiatric screening questions included in the ESEMEd questionnaire is presented elsewhere (The ESEMEd investigators, 2004). All participants responding positively to any of the screening questions had to complete the Composite International Diagnostic Interview (CIDI) section of the specific disorder prompted by that question. Based on the responses provided in the mood and anxiety sections, the respondent was identified as having depression or anxiety symptoms and was directed to the long or the short path accordingly. The respondents that followed the long path of the questionnaire were individuals that, based on their anxiety or depression symptoms could be considered as 'high risk individuals', and a random subsample (25%) of the respondents without symptoms

('low risk individuals'). The remaining 75% of respondents without symptoms not randomly selected for the long path followed the short path of the questionnaire. Only those who followed the long path of the interview ($n = 8796$) were assessed the ESEMeD WHODAS and therefore, were included in the analyses of this study (for a visual representation, see Figure 1).

World Health Organization Disability Assessment Schedule (WHODAS)

The ESEMeD WHODAS (see Appendix) includes six scales: (a) Role Functioning, (b) Cognition, (c) Mobility, (d) Self-care, and (e) Social Interaction, (f) Participation. The first disability scale (Role Functioning) in the ESEMeD WHODAS is a radical departure from the WHODAS-II scale as global questions about overall role functioning rather than separate questions about housework and employment were used. The four questions that assessed Role Functioning were explicitly designed to define 'normal activities' in such a way as to encompass whatever it is that the respondent might normally do whether they are a homemaker or an employed person. Also, these questions were all frequency questions while other domains consist mostly of severity items. The total score on the scale 'Role Functioning' was calculated in such a way that each day out of role is assigned a score of 1, each day of cutback in quantity or quality is assigned a score of 0.5, and each day of extreme effort is assigned a score of 0.25. The sum is then transformed to a 0–100 scale, which indicates the percent of days a respondent was completely out of role.

In the multi item domains Cognition, Mobility, Self-care, and Social Interaction, respondents were initially asked a single general question about whether they had any difficulty in the broad area of functioning in the past 30 days and, if so, they are asked about frequency

and severity of the problems. These domains are scored in a 0 to 100 scale, where 0 represents no disability and 100 represents maximum impairment.

The last three questions in the Participation scale, i.e. Embarrassment, Discrimination, and Family burden, were only assessed in respondents that reported at least some limitations in other domains of functioning. Those that were not asked the question were assigned '0'. It was reasoned that since they did not report any disability in the previous questions, and therefore were unlikely to experience embarrassment, discrimination, or family burden because of their health problems.

For all domains, respondents who gave refuse or do not know responses to individual items are assigned a score of zero in order to give conservative estimates about problems in functioning. The exact wordings of the questions are presented in the Appendix (Table A1).

Statistical analyses

Structural equation models for categorical and continuous variables were used to assess the psychometric properties of the ESEMeD WHODAS. First, an Exploratory Factor Analysis (EFA) with promax rotation was performed on a random 50% of the sample. Second, the remaining 50% of the sample was used to perform Confirmatory Factor Analyses (CFAs) to compare three different models: (a) the model suggested from the results of the EFA, (b) the theoretical WHO model that assumed six disability factors, and (c) a reduced model, in which the same six domains as in model b were tested, but without the frequency items FD10a (Cognition), FD12a (Mobility), FD14a (Self-care), and FD16a (Social Interaction). Third, a CFA with the country covariate (MIMIC model) was estimated to test for measurement non-invariance (i.e. direct relationship between the covariate and the items that are not

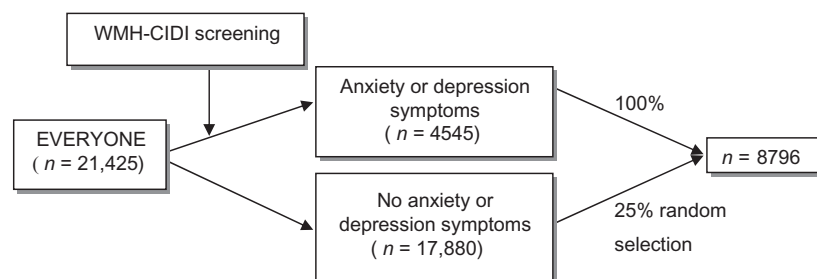


Figure 1. Assessment of ESEMeD WHODAS scales.

mediated by the factors) and population heterogeneity (i.e. relationship between the covariate and the factors) across countries, that were grouped into Mediterranean (Spain, France and Italy) and non-Mediterranean (Belgium, Germany and the Netherlands). Fourth, a single second-order factor, including the first-order factors, was specified and tested using CFA. Fifth, internal consistency of the ESEMeD WHODAS was determined by calculating Cronbach's alpha for each domain. Cronbach's alpha indicated good internal consistency for group comparisons if it is higher than 0.70 and good internal consistency for individual comparisons if it is 0.90 or above. Finally, the discriminating validity of the ESEMeD WHODAS was examined by comparing individuals (a) without any lifetime mental disorder; (b) any lifetime, but no 12-month mental disorder; (c) any non-severe 12-month mental disorder; and (d) any severe 12-month mental disorder. Mental disorders were assessed using a revised and further enhanced version of the CIDI, called WMH-CIDI, which was developed and adapted by the Coordinating Committee of the WHO-WMH 2000 Initiative. The WMH-CIDI includes fully structured questions on presence, persistence and intensity of clusters of psychiatric symptoms followed by probes for age of onset and lifetime course. The WMH-CIDI gives the Diagnostic and Statistical Manual of Mental Disorders – fourth edition (DSM-IV) diagnoses as well as International Classification of Disease (ICD)-10 diagnoses. For this article, we only included DSM-IV diagnoses.

As described by the ESEMeD investigators in 2004, five types of weights were applied in the study:

- (i) Weights to adjust for the different probabilities of selection of the individuals 18 years or older that lived in the same household. These weights applied only where household was implied in the sampling design, taking spouses into account where appropriate.
- (ii) Weight to restore under-sampling of hard-to-reach people that were randomly selected to be reapproached. This type of weight incorporates the information of the proportion of hard to reach individuals that were randomly selected to be reapproached in each country.
- (iii) Post-stratification weights were calculated and applied to the data in order to correct for imbalances of age and gender characteristics between the study sample and their respective general

populations. Population census distributions of age and gender were obtained for each country stratified by the primary- sampling units used in the specific country.

- (iv) Population projection weights were applied in order to restore the proportion of each country's population size when analyzing all the countries together, because currently, the different sample sizes across countries do not reflect the differences in the population sizes between these countries.
- (v) Weights to account for the different probabilities of selection for the 'long path' of the questionnaire.

EFA and CFA for categorical and continuous variables were performed with Mplus, version 4.2 (Muthén and Muthén, 1998–2006). Factors were selected in the EFA taking into account the residuals and the percent of variance explained and also based on the interpretability of the results. In CFA, robust weighted least squares estimator was applied (WLSMV in Mplus), that uses a diagonal weight matrix with robust standard errors and mean- and variance-adjusted χ^2 test statistic. Goodness-of-fit (GOF) was assessed with the following fit indices: (a) Confirmatory Fit Index (CFI), and (b) Tucker–Lewis Index (TLI) which both indicate good fit if their values exceed 0.95 (Hu and Bentler, 1999), (c) Root Mean Square Error of Approximation (RMSEA) which indicates adequate fit if it is less than 0.08, and good fit if the value is less than 0.05 (Browne and Cudeck, 1993), (d) Weighted Root Mean Square Residual (WMRM) indicates good fit if it is less than 1.0. A decisive argument for electing TLI and WMRM was that they are relatively independent of the sample size compared with other common indices. Other issues that were taken into account to assess model fit were: the degree of significance of the factor loadings and the fact that the residual variances did not take negative values for any of the items. Estimation methods implemented in M-PLUS for the analysis of complex survey data were used.

The internal consistency and validity analyses were performed with SASTM software version 9.01 (SAS institute, 2006) and SUDAAN software version 9, a statistical package used to estimate standard errors of data obtained from surveys with a complex sampling design (Research Triangle Institute, 2005). The Taylor series linearization method implemented in SUDAAN

was used, and the sampling strata and primary sampling units in each country were taken into account.

Results

Descriptive

Table 1 presents the activity limitations and participation restrictions in the ESEMeD sample. The activities in which many individuals experienced limitations in the past 30 days include reduced quantity of work (FD7; 13.9%) and being emotionally affected (FD18b; 14.3%), while activities in which few individuals experienced limitations are: controlling their emotions when being around people (FD17e; 1.5%) and experiencing discrimination or unfair treatment (FD21; 1.5%).

Factor structure of the ESEMeD WHODAS

Table 2 shows the results of the EFA on a random 50% of the total sample. Seven factors were found which explained 80% of the variance. Frequency and severity items tended to load on different factors. Two frequency items, FD10a (Cognition) and FD16a (Social Interaction), loaded on a separate factor altogether. FD14a (Self-care) presented small loading in all factors (<0.3). In the subsequent CFA, FD14a was kept with the fifth factor, because factor loading was highest for that factor and it was consistent with the a priori theory.

Comparing three competing models

Three CFAs were performed on the remaining 50% of the ESEMeD sample. The results presented in Table 3 suggest that the data fitted the EFA-model (developed in the EFA) well (CFI, TLI), acceptably (RMSEA), and not so good (WRMR). The theoretical WHO model fitted slightly worse, although still good according to the TLI. The CFI and RMSEA were acceptable but the WRMR was not good. The reduced model (six factors, without four frequency items) outperformed the other two models, as all fit indices indicated good fit.

Internal consistency was determined for the disability scales in the theoretical WHO model and the reduced model (Table 4) using the full sample of 8796 individuals, because the factor structure in the two subsamples was similar. All disability scales have good internal consistency for group comparisons. The disability scales based on the reduced model perform slightly less compared to the disability scales based on the theoretical WHO model, but are still very good. The slightly lower Cronbach's alpha of the factors

without the frequency items might be due to the fact that Cronbach's alpha tends to increase as the number of items increase.

Global disability latent variable

It is unclear whether one or two global disability latent variables underpin the observed associations between the disability scales. The two-dimensional model would encompass activity limitations (Cognition, Self-care, Mobility, Social Interaction) and participation restrictions (Role Functioning and Participation). The one- and two-dimensional models were tested on the three sets of factor scales as derived in the three competing models. The solution for the two-dimensional model (activity limitations, participation restrictions) was not positive definite in any of the three sets of factor scales, suggesting a very high correlation between the two dimensions. The one-dimensional fitted rather well in all three sets of factor scales as indicated by the CFI and TLI (Table 5). The RMSEA indicated acceptable fit in the exploratory and the theoretical WHO model, while it indicated good fit in the reduced model. None of the models (sets of factor scales) had acceptable WRMR as indicated by values above 1.5, but the reduced model performed best. In every model, Role Functioning presented a low R^2 , which indicates that the one-dimensional second-order factor of global disability only partly covers the role functioning domain.

Invariance and population heterogeneity across countries

The CFA model was estimated separately for Mediterranean and non-Mediterranean countries. The GOF indices were quite good and similar to those obtained for all the countries together (Table 6). Moreover, the factor loadings were all significant in the separate models. These results suggest that there is dimensional invariance (same number of factors in each group) and configural invariance (common factors associated with the same items across groups) (Gregorich, 2006). In the models including the covariate (MIMIC) presented in Table 6, the GOF indices are still good. The average scores on Mobility and Participation differ across groups because there is an indirect effect of the country (population heterogeneity) on these two factors. Moreover, measurement non-invariance was observed for the Embarrassment item, as the χ^2 test comparing the model including the direct effect on this item with the model that did not include the direct effect is

Table 1. Activity limitations and participation restrictions in the ESEMeD sample

Item ¹	Percentage with limitations	Mean (standard deviation)	
		Overall	With limitations
<i>Role Functioning</i>			
FD4 (unable to work) ²	8.3 (11.9)	1.1 (6.6)	11.7 (15.7)
FD7 (reduced quantity) ²	13.9 (18.8)	2.8 (10.6)	16.3 (17.7)
FD8 (reduced quality) ²	9.4 (12.8)	1.3 (7.3)	13.4 (16.4)
FD9 (effort) ²	10.3 (13.6)	1.5 (8.1)	14.2 (15.8)
<i>Cognition</i>			
FD10a (frequency) ²	5.3 (6.9)	0.5 (4.0)	10.0 (12.3)
FD11a (concentrating) ³	3.8 (4.7)	1.1 (0.5)	2.9 (0.9)
FD11b (understanding) ³	2.1 (2.6)	1.0 (0.3)	2.6 (0.7)
FD11c (remembering) ³	3.4 (4.6)	1.1 (0.4)	2.7 (0.9)
FD11d (learning) ³	2.3 (2.7)	1.0 (0.4)	3.0 (1.0)
<i>Mobility</i>			
FD12a (frequency) ²	12.7 (18.6)	2.2 (12.0)	17.6 (18.1)
FD13a (standing) ³	11.1 (16.4)	1.3 (1.4)	3.4 (1.5)
FD13b (moving in home) ³	8.5 (12.1)	1.2 (0.9)	2.9 (1.2)
FD13c (walking) ³	11.0 (16.4)	1.3 (1.5)	3.8 (1.5)
<i>Self-care</i>			
FD14a (frequency) ²	3.2 (4.9)	0.6 (6.3)	18.4 (18.0)
FD15a (washing) ³	2.6 (4.0)	1.1 (0.7)	3.4 (1.4)
FD15b (dressing) ³	2.6 (4.0)	1.1 (0.6)	3.2 (1.3)
FD15c (being alone) ³	1.6 (2.4)	1.0 (0.5)	3.5 (1.8)
<i>Social Interaction</i>			
FD16a (frequency) ²	3.6 (5.0)	0.5 (5.1)	14.3 (15.4)
FD17a (conversing) ³	1.8 (2.3)	1.0 (0.3)	2.8 (0.9)
FD17b (unknown people) ³	1.7 (2.2)	1.0 (0.4)	3.1 (1.5)
FD17c (friendship) ³	1.6 (2.0)	1.0 (0.3)	2.8 (0.9)
FD17d (new friends) ³	2.1 (2.9)	1.1 (0.6)	3.6 (1.6)
FD17e (control emotions) ³	1.5 (1.9)	1.0 (0.3)	2.7 (0.8)
<i>Participation</i>			
FD18b (emotionally affected) ³	14.3 (19.2)	1.2 (0.9)	2.5 (1.0)
FD18c (family financial burden) ³	6.6 (9.1)	1.1 (0.7)	2.7 (1.2)
FD18d (joining activities) ³	10.1 (13.8)	1.2 (1.0)	3.0 (1.6)
FD18e (barriers in world around) ³	7.7 (11.1)	1.1 (0.9)	2.9 (1.7)
FD20 (embarrassment) ³	5.5 (7.8)	1.1 (0.6)	2.7 (1.4)
FD21 (discrimination) ³	1.5 (1.8)	1.0 (0.4)	2.8 (1.5)
FD22 (family burden) ³	6.3 (8.6)	1.1 (0.6)	2.6 (1.2)

¹See Appendix for exact wording of items.²Mean is based on the number of days (0–30) a person reported having difficulties.³Mean is based on the following values: (1) no difficulties, (2) mild difficulty, (3) moderate difficulty, (4) severe difficulty, and (5) spontaneous 'could not do'.

Table 2. Exploratory factor analysis on the items from the ESEMeD WHODAS

	Component						
	I	II	III	IV	V	VI	VII
<i>Role Functioning</i>							
FD4 (unable to work)	0.869	−0.023	−0.079	−0.026	0.032	0.013	−0.014
FD7 (reduced quantity)	0.676	0.035	−0.049	0.013	0.011	0.051	0.005
FD8 (reduced quality)	0.911	0.035	0.046	−0.042	−0.044	0.001	0.003
FD9 (effort)	0.841	−0.007	0.024	0.023	−0.052	0.001	0.043
<i>Cognition</i>							
FD10a (frequency)	0.074	0.251	0.462	−0.020	−0.054	−0.077	−0.018
FD11a (concentrating)	0.043	0.884	−0.018	−0.048	0.085	0.072	−0.011
FD11b (understanding)	0.058	0.973	−0.133	0.031	−0.039	−0.033	0.041
FD11c (remembering)	−0.060	0.867	0.082	0.094	−0.082	0.070	0.058
FD11d (learning)	−0.040	0.806	0.035	−0.001	0.069	0.152	−0.019
<i>Mobility</i>							
FD12a (frequency)	0.175	−0.013	0.228	0.567	−0.032	−0.150	0.024
FD13a (standing)	−0.024	0.061	−0.051	0.693	0.219	0.041	0.166
FD13b (moving in home)	0.026	0.072	−0.093	0.563	0.286	0.056	0.223
FD13c (walking)	−0.043	0.021	−0.038	0.719	0.194	0.096	0.162
<i>Self-care</i>							
FD14a (frequency)	0.262	−0.081	0.268	0.023	0.294	−0.076	−0.050
FD15a (washing)	−0.007	−0.018	0.008	0.071	0.864	0.070	0.109
FD15b (dressing)	−0.003	0.091	−0.018	0.112	0.864	0.040	−0.018
FD15c (being alone)	−0.015	−0.099	−0.006	0.058	0.675	0.099	0.368
<i>Social Interaction</i>							
FD16a (frequency)	0.061	−0.122	0.728	−0.056	−0.029	0.136	0.015
FD17a (conversing)	0.000	0.163	0.011	−0.029	0.115	0.851	−0.053
FD17b (unknown people)	0.009	0.032	0.022	−0.024	0.179	0.789	0.084
FD17c (friendship)	0.021	0.075	0.005	0.082	−0.064	0.852	0.065
FD17d (new friends)	−0.010	−0.061	0.087	0.049	0.036	0.786	0.222
FD17e (control emotions)	0.020	0.216	0.012	−0.078	−0.063	0.760	0.117
<i>Participation</i>							
FD18b (emotionally affected)	−0.017	0.254	0.024	0.081	−0.115	0.081	0.647
FD18c (family financial burden)	0.028	−0.040	−0.014	−0.042	−0.002	−0.078	0.942
FD18d (joining activities)	0.047	−0.040	−0.005	0.178	−0.075	0.228	0.700
FD18e (barriers in world around)	0.015	−0.147	0.004	0.211	−0.122	0.136	0.796
FD20 (embarrassment)	0.004	0.223	0.015	−0.040	0.197	−0.025	0.602
FD21 (discrimination)	−0.021	0.059	0.015	−0.145	0.162	−0.114	0.897
FD22 (family social life burden)	0.038	0.109	0.023	−0.002	0.131	0.189	0.601
Cumulative percent of variance explained	53.0%	63.6%	70.9%	75.0%	79.0%	82.1%	84.7%

Note: $\chi^2 = 91.187$ (df = 27); RMSEA = 0.024.

Table 3. Comparison of fit indices of the CFAs for (a) the new model as found in the EFA, (b) the model as it was a priori designed in ESEMeD, and (c) the reduced model in which frequency items were left out

	WHODAS models		
	Exploratory	A priori	Reduced
χ^2 (df)	411.609 (31)	787.403 (32)	118.049 (33)
CFI	0.958	0.916	0.992
TLI	0.977	0.956	0.996
RMSEA	0.053	0.073	0.024
WRMR	1.925	2.649	0.969

Table 4. Internal consistency (Cronbach's alpha) of the domains of functioning identified in the ESEMeD WHODAS

Domains of functioning	Cronbach's alphas of WHODAS models	
	A priori	Reduced
Global disability	0.94	0.93
Role Functioning	0.90	0.90
Cognition	0.89	0.88
Mobility	0.94	0.93
Self-care	0.92	0.89
Social Interaction	0.91	0.91
Participation	0.84	0.84

Table 5. Comparison of fit indices of the global disability factor for (a) the new model as found in the EFA, (b) the model as it was a priori designed in ESEMeD, and (c) the reduced model in which frequency items were left out

	WHODAS models		
	Exploratory	A priori	Reduced
χ^2 (df)	436.830 (23)	544.486 (25)	165.699 (25)
CFI	0.954	0.942	0.987
TLI	0.966	0.961	0.991
RMSEA	0.064	0.069	0.036
WRMR	2.728	2.880	1.556
<i>Latent variable R²</i>			
Role Functioning	0.204	0.201	0.158
Cognition	0.679	0.669	0.665
Mobility	0.687	0.696	0.715
Self-care	0.820	0.823	0.816
Social Interaction	0.667	0.708	0.712
Participation	0.869	0.882	0.908
FD10a/FD16a	0.393	—	—

statistically significant [$\chi^2(1) = 56.014$, p -value < 0.001], indicating that the model improves when the direct effect is included, although the GOF indices were very similar in both models (and better than the prespecified cutoff point).

Validity of the ESEMeD WHODAS

To examine the discriminatory validity of the ESEMeD WHODAS, the following groups known to differ in

activity limitations were used: (1) individuals without any DSM-IV lifetime mental disorder; (2) individuals with a DSM-IV lifetime, but no 12-month DSM-IV mental disorder; (3) individuals with a 12-month DSM-IV mental disorder that was (a) not severely, or (b) severely interfering (according to the Sheehan disability scales) with work, social life, or personal relationships. It was hypothesized that the first group functioned best on all disability scales measured by the ESEMeD

Table 6. Comparison of MIMIC models that (a) include only the indirect effects of the covariate on each of the factors, and (b) include the indirect effects and an additional direct effect of country on the Embarrassment item

	Model with country covariate Estimate (SE)	Final model with direct effect Estimate (SE)
<i>Indirect effects parameters:</i>		
Role Functioning	−0.07 (0.12)	−0.07 (0.12)
Cognition	0.003 (0.06)	0.003 (0.06)
Mobility	0.22 (0.05) ¹	0.22 (0.05) ¹
Self-care	−0.06 (0.09)	−0.05 (0.09)
Social Interaction	0.12 (0.07)	0.117 (0.07)
Participation	0.04 (0.04)	0.09 (0.04) ¹
<i>Direct effect parameter:</i>		
Embarrassment item	–	−0.41 (0.06) ¹
χ^2 (DF) of comparison of the two models ²		56.014 (1), <i>p</i> -value < 0.001
<i>Fit indices</i>		
χ^2 (DF) ³	160.494 (30)	152.617 (30)
CFI	0.992	0.992
TLI	0.996	0.996
RMSEA	0.022	0.022
WRMR	1.201	1.168

¹ Estimate (EST)/standard error (SE) > 1.96 (statistically significant estimate).

² Chi-squared (χ^2) value for mean and variance-adjusted Weighted Least Squares estimator (WLSMV) cannot be used for difference testing. To compare the two nested models, the DIFFTEST command implemented in Mplus has been used.

³ Chi-squared (χ^2) value for mean and variance-adjusted Weighted Least Squares estimator (WLSMV) used.

WHODAS and those with severely interfering 12-month mental disorder the worst. The data clearly supported the hypothesis (Table 7). The level of activity limitations and participation restrictions increased with the level of severity as defined by the groups. Individuals without any lifetime mental disorder functioned better than any other group whereas those with a severely interfering 12-month mental disorder did worst.

To further examine the validity of the ESEMeD WHODAS we also compared individuals with physical disorders who reported (1) no, (2) moderate, or (3) severe interference of their physical disorders with work, social life, or their personal relationships (results not presented here). A similar gradient in WHODAS-II activity limitations and participation restrictions was found.

Discussion

The ESEMeD WHODAS is a modified version of the WHODAS-II, in which filter questions were used and

fewer questions were included to reduce respondent burden. The modifications seem to have come at little cost in terms of psychometric properties. The structure of the ESEMeD WHODAS is quite similar to the structure found in the original WHODAS-II. We found one uni-dimensional global disability latent variable that was an accurate summary of the subscale scores. The Global Disability variable consisted of six distinct subscales: Role Functioning, Cognition, Mobility, Self-care, Social Interaction, and Participation which is in accordance with the WHODAS-II. Excluding four frequency items (Cognition, Mobility, Self-care, and Social Interaction) from the questionnaire improved the underlying factor structure. Also, discriminatory validity and internal consistency of the ESEMeD WHODAS are acceptable and the factor structure is robust over Mediterranean and non-Mediterranean countries.

However, some issues regarding the ESEMeD WHODAS need to be addressed. First, Role Functioning retained much unique variance not shared with Global Disability. This may be related to the fact that

Table 7. Activity and participation limitations in respondents with gradient on severity of mental disorders

	No lifetime mental disorder		Any lifetime mental disorder (no 12 months)		Any non-severe 12 months mental disorder		Any severe 12 months mental disorder	
	Percentage	Mean (95% confidence interval, CI)	Percentage	Mean (95% CI)	Percentage	Mean (95% CI)	Percentage	Mean (95% CI)
Role Functioning	60.4	7.3 (6.6–8.0)	55.3	9.7 (8.3–11.1)	64.5	15.4 (1.5–19.3)	61.5	23.1 (20.2–26.0)
Cognition	2.8	0.3 (0.1–0.4)	5.5	0.4 (0.3–0.5)	12.1	1.0 (0.6–1.4)	22.7	4.4 (3.4–5.5)
Mobility	10.1	3.5 (2.9–4.1)	13.5	3.4 (2.6–4.2)	18.3	5.8 (3.1–8.5)	26.8	9.4 (7.6–11.2)
Self-care	2.2	0.8 (0.5–1.1)	3.2	0.6 (0.4–0.8)	6.3	2.1 (0.5–3.6)	9.1	2.9 (1.7–4.1)
Social Interaction	1.6	0.3 (0.1–0.4)	4.0	0.5 (0.3–0.7)	5.7	1.0 (0.1–2.0)	12.8	3.3 (2.3–4.3)
Participation	17.1	2.3 (2.0–2.6)	23.9	3.3 (2.8–3.8)	38.0	6.3 (4.4–8.3)	49.6	11.7 (10.1–13.4)

Note: Severity of the disorder is based on question of interference asked in each of the disorder sections: 'You mentioned having the mental health problems I just asked you about. How much did your mental health problems and these other problems interfere with either your work, your social life, or your personal relationships during that episode – not at all, a little, some, a lot, or extremely?'

the Role Functioning questions are all frequency questions whereas the other domains consist exclusively of intensity questions. Therefore, it might be better to not include the Role Functioning in the overall disability scale but to use it separately as an index of limitations in role functioning. Second, the findings do not exclude a possible hierarchy between Cognition, Mobility, Self-care, and Social Interaction on the one hand and Role Functioning and Participation on the other hand as the latter two regard a more general level of functioning. Third, the item FD20 (Embarrassment) is performing differently in Mediterranean and non-Mediterranean countries. This is most likely an internal problem of the ESEMeD questionnaire due to a problem in the translation from English into French. At this point we do not recommend removing it from the questionnaire, but advise adjusting for non-invariance of this item by allowing the direct relation of the country to the item FD20, i.e. including the direct effect in the model. Fourth, the prevalence of disability in some domains is unexpectedly low. This may be due to the nature of the sample, a general population sample, but it is also possible that the filter questions have been too stringent. This may have reduced the ability of the ESEMeD WHODAS to differentiate individuals with relatively mild levels of functional limitations from individuals who have no functionally limitations at all. An additional limitation concerning the very skewed data is that most statistical procedures assume a normal distribution. Solutions to this problem may be to (a) use statistical procedures that do not rely on a normal distribution, but can handle very skewed and even zero-inflated data, (b) categorize limitations in 'zero', 'some', and 'many', or (c) dichotomize at the 90th percentile of the population scores. The most profitable strategy depends on the purpose of the study. For further studies we suggest excluding the filter questions to get more elaborate data on disability, especially in the non-disabled to mildly disabled range. Finally, the results presented in this study stimulate, hopefully, further psychometric research. We suggest using a longitudinal design in order to be able to evaluate responsiveness to change and test–retest reliability. In addition, we suggest dropping the filter questions of the ESEMeD WHODAS as it is not clear yet whether they have sufficient sensitivity. The data suggest that the current filter questions are too restrictive, i.e. have rather low sensitivity. Dropping the filter questions would yield a dataset that would allow the application of Item Response Theory

(which has several potential advantages over the Classical Test Theory methods) to assess additional psychometric properties of the ESEMeD WHODAS. Dropping the filter questions would imply an increase of the average administration time but this could be compensated by dropping the frequency questions which do not seem to add much to the severity questions.

Conclusion

The current study suggests that the ESEMeD WHODAS may well be a valuable shortened version of the WHODAS-II to measure activity limitations and participation restriction. The use of filter questions and frequency items is not recommended in future studies. Also needed are longitudinal research and the application of Item Response Theory to confirm our conclusions and extend them by assessing responsiveness to change, test-retest reliability, and additional validity aspects.

Acknowledgements

The (ESEMeD) project is carried out in conjunction with the World Health Organization World Mental Health (WMH) Survey Initiative. We thank the WMH staff for assistance with instrumentation, fieldwork, and data analysis. These activities were supported by the United States National Institute of Mental Health (R01MH070884), the John D. and Catherine T. MacArthur Foundation, the Pfizer Foundation, the US Public Health Service (R13-MH066849, R01-MH069864, and R01 DA016558), the Fogarty International Center (FIRCA R01-TW006481), the Pan American Health Organization, Eli Lilly and Company, Ortho-McNeil Pharmaceutical, Inc., GlaxoSmithKline, and Bristol-Myers Squibb. A complete list of WMH publications can be found at <http://www.hcp.med.harvard.edu/wmh/>. In Europe, this project was funded by the European Commission (Contract QLG5-1999-01042 and SANCO contract 2004123); the Piemont Region (Italy), Fondo de Investigacion Sanitaria, Instituto de Salud Carlos III, Spain (FIS 00/0028), Ministerio de Ciencia Y Tecnologia, Spain (SAF 2000-158-CE), Department de Sanitat, Generalitat de Catalunya, Spain, other local agencies, an unrestricted educational grant from Glaxo-SmithKline and a personal grant from GlaxoSmithKline Netherlands.

Declaration of Interests

All authors state that they do not have conflicts of interest.

References

Annicchiarico R, Gibert K, Cortes U, Campana F, Caltagirone C (2004). Qualitative profiles of disability. *J Rehabil Res Dev* 41(6): 835–846.

Browne MW, Cudeck R (1993). Alternative ways of assessing model fit. In Bollen KA, Long JS (eds) *Testing Structural Equation Models*. Newbury Park, CA: Sage. pp 136–162.

Buist-Bouwman MA, De Graaf R, Vollebergh W, Alonso J, Bruffaerts R, Ormel J et al. (2006). Functional disability of mental disorders and comparison with physical disorders. A study among the general population of six European countries. *Acta Psychiatr Scand* 113: 492–500.

Chisolm TH, Abrams HB, McArdle R, Wilson RH, Doyle PJ (2005). The WHO-DAS II: psychometric properties in the measurement of functional health status in adults with acquired hearing loss. *Trends Amplif* 9(3): 111–126.

Chopra PK, Couper JW, Herrman H (2004). The assessment of patients with long-term psychotic disorders: application of the WHO Disability Assessment Schedule II. *Aust NZ J Psychiatry* 38(9): 753–759.

Chwastiak LA, Von Korff M (2003). Disability in depression and back pain: evaluation of the World Health Organization Disability Assessment Schedule (WHO DAS II) in a primary care setting. *J Clin Epidemiol* 56(6): 507–514.

Gallagher P, Mulvany F (2004). Levels of ability and functioning: using the WHODAS II in an Irish context. *Disabil Rehabil* 26(9): 506–517.

Goering P, Lin E, Campbell D, Boyle M, Offord D (1996). Psychiatric disability in Ontario. *Can J Psychiatry* 41: 564–571.

Gregorich SE (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Med Care* 44(11 Suppl. 3): S78–S90.

Hu L, Bentler PM (1999). Cutoff criterion for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling* 6: 1–55.

Kessler RC, Frank R (1997). The impact of psychiatric disorders on work loss days. *Psychol Med* 27: 861–873.

Kim JM, Stewart R, Glozier N, Prince M, Kim SW, Yang SJ, Shin IS, Yoon JS (2005). Physical health, depression and cognitive function as correlates of disability in an older Korean population. *Int J Geriatr Psychiatry* 20: 160–167.

Kouzis A, Eaton W (1994). Emotional disability days: prevalence and predictors. *Am J Public Health* 84: 1304–1307.

Kouzis A, Eaton W (1997). Psychopathology and the development of disability. *Soc Psychiatry Psychiatr Epidemiol* 32: 379–386.

McKibbin C, Patterson TL, Jeste DV (2004). Assessing disability in older patients with schizophrenia: results from the WHODAS-II. *J Nerv Ment Dis* 192(6): 405–413.

Mathers CD, Loncar D (2006). Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med* 3(11): 2011–2030.

Muthén LK, Muthén BO (1998–2006). *Mplus Statistical Analysis with Latent Variables, User's Guide*, 4th edition. Los Angeles, CA: Muthén & Muthén.

- Research Triangle Institute (2005). SUDAAN, Release 9.0. Research Triangle Park, NC: Research Triangle Institute.
- SAS Institute Inc (2006). Base SAS® 9.1.3 Procedures Guide, 2nd edition, volumes 1, 2, 3, and 4. Cary, NC: SAS Institute Inc.
- The ESEMeD/MHEDEA 2000 Investigators (2002). The European Study of the Epidemiology of Mental Disorders (ESEMeD/MHEDEA 2000) project: rationale and methods. *Int J Methods Psychiatr Res* 11: 55–67.
- The ESEMeD/MHEDEA 2000 Investigators (2004). Sampling and methods of the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand* 109(s420): 8–20.
- van Tubergen A, Landewe R, Heuft-Dorenbosch L, Spoorenberg A, van der Heijde D, van der Tempel H, van der Linden S (2003). Assessment of disability with the World Health Organisation Disability Assessment Schedule II in patients with ankylosing spondylitis. *Ann Rheum Dis* 62: 140–145.
- World Health Organization (2001). International Classification of Functioning, Disability and Health (ICF). Geneva: World Health Organization.

Correspondence: J. (hans) Ormel, University Medical Center Groningen, University of Groningen, P.O. Box 30.001, 9700 RB Groningen, The Netherlands.
Telephone: (+31)-50-3612078
Fax: (+31)-50-3619722
Email: J.Ormel@med.umcg.nl

Appendix

Table A1. ESEMeD WHODAS

Scale	Question	Description
Role Functioning	FD4	Beginning yesterday and going back 30 days, how many days out of the past 30 were you <i>totally unable</i> to work or carry out your normal activities?
	FD7	How many days out of the past 30 were you able to work and carry out your normal activities, but had to cut down on what you did or not get as much done as usual?
	FD8	How many days out of the past 30 did you cut back on the <i>quality</i> of your work or how <i>carefully</i> you worked because of problems with either your physical health, your mental health, or your use of alcohol or drugs?
	FD9	How many days out of the past 30 did it take an extreme effort to perform up to your usual level at work or at your other normal daily activities because of problems with either your physical health, your mental health, or your use of alcohol or drugs?
Cognition	FD10	Was there ever a time in the past 30 days when health-related problems caused difficulties with either your concentration, memory, understanding, or ability to think clearly?
	FD10a	How many days did you have these difficulties during the past 30 days?
	FD11a	During those days, how much difficulty did you have in each of the following areas: Concentrating on doing something for 10 minutes – none, mild, moderate, or severe difficulty?
	FD11b	During those days, how much difficulty did you have in understanding what was going on around you? – none, mild, moderate, or severe difficulty?
	FD11c	During those days, how much difficulty did you have in remembering to do important things? – none, mild, moderate, or severe difficulty?
	FD11d	During those days, how much difficulty did you have in learning a new task – for example, learning how to get to a new place? – none, mild, moderate, or severe difficulty?
Mobility	FD12	Was there ever a time in the past 30 days when health-related problems caused you difficulties with mobility, such as standing for long periods, moving around inside your home, or getting out of your house?
	FD12a	How many days did you have these difficulties during the past 30 days?

Table A1. *continued*

Scale	Question	Description
Self-care	FD13a	During those days, how much difficulty did you have in each of the following areas: standing for long periods, such as 30 minutes? – none, mild, moderate, or severe difficulty?
	FD13b	During those days, how much difficulty did you have in each of the following areas: Moving around inside your home? – none, mild, moderate, or severe difficulty?
	FD13c	During those days, how much difficulty did you have in walking a long distance such as (a kilometer/half a mile)? – none, mild, moderate, or severe difficulty?
	FD14	Was there ever a time in the past 30 days when health-related problems caused you difficulties with self care, such as washing your whole body, getting dressed, or feeding yourself?
	FD14a	How many days did you have these difficulties during the past 30 days?
	FD15a	During those days, how much difficulty did you have in each of the following areas: washing your whole body? – none, mild, moderate, or severe difficulty?
	FD15b	During those days, how much difficulty did you have in getting dressed? – none, mild, moderate, or severe difficulty?
	FD15c	During those days, how much difficulty did you have in staying by yourself for a few days? – none, mild, moderate, or severe difficulty?
Social Interaction	FD16	Was there ever a time in the past 30 days when health-related problems caused you difficulties either getting along with people, maintaining a normal social life, or participating in social activities?
	FD16a	How many days did you have these difficulties during the past 30 days?
	FD17a	During those days, how much difficulty did you have in each of the following areas: starting and maintaining a conversation? – none, mild, moderate, or severe difficulty?
	FD17b	During those days, how much difficulty did you have in dealing with people you did not know well? – none, mild, moderate, or severe difficulty?
	FD17c	During those days, how much difficulty did you have in maintaining friendships? – none, mild, moderate, or severe difficulty?
	FD17d	During those days, how much difficulty did you have in making new friends? – none, mild, moderate, or severe difficulty?
	FD17e	During those days, how much difficulty did you have in controlling your emotions when you were around people? – none, mild, moderate, or severe difficulty?
	FD18b	How much were you emotionally affected by your health during the past 30 days? – none, mild, moderate, or severe difficulty?
Participation	FD18c	How much was your health a drain on the financial resources of you or your family during the past 30 days? – none, mild, moderate, or severe difficulty?
	FD18d	During the past 30 days, 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 because of your health? – none, mild, moderate, or severe difficulty?
	FD18e	How much of a problem did you have because of barriers or hindrances in the world around you during the past 30 days? – none, mild, moderate, or severe difficulty?
	FD20	How much embarrassment did you experience because of your health problems during the past 30 days? – none, a little, some, a lot, or extreme embarrassment?
	FD21	How much discrimination or unfair treatment did you experience because of your health problems during the past 30 days? – none, a little, some, a lot, or extreme unfair treatment?
	FD22	How much did your health-related difficulties interfere with the life and activities of your close friends and family members during the past 30 days? – not at all, a little, some, a lot, or extremely?