Journal of Clinical Epidemiology 59 (2006) 541–543

BRIEF REPORT

Dimensional structure of the SF-36 in neurological patients

A.J. Dallmeijer[a,b,*], J. Dekker[a,b], D.L. Knol[b,c], S. Kalmijn[d,e], V.P.M. Schepers[e], V. de Groot[a,b], E. Lindeman[e], A. Beelen[f], G.J. Lankhorst[a,b], on behalf of the FuPro study group

[a]Department of Rehabilitation Medicine, VU University Medical Center, P.O. Box 7057, Amsterdam 1007 MB, The Netherlands
[b]Institute for Research in Extramural Medicine, VU University Medical Center, Amsterdam, The Netherlands
[c]Department of Clinical Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands
[d]Julius Center for Health Sciences and Primary Care, University Medical Center, Utrecht, The Netherlands
[e]Rehabilitation Center De Hoogstraat, Utrecht, and Rudolf Magnus Institute of Neuroscience, University Medical Center, Utrecht, The Netherlands
[f]Department of Rehabilitation, Academic Medical Center, Amsterdam, The Netherlands

Accepted 16 October 2005

Health-related quality of life is an important outcome in rehabilitation. Assessment of the impact of disability or diseases on quality of life is essential for the evaluation of rehabilitation treatments and needs. One of the most used generic health status measures is the Medical Outcome Study 36-Item Short Form Health Survey (SF-36), consisting of eight multi-item subscales and two higher order summary scales, the physical (PCS) and mental component summary scale (MCS) [1,2]. The broad application of the SF-36 in different populations and patient groups requires that the psychometric criteria underlying the dimensional structure of the questionnaire are met in each population [3]. For the eight subscales of the SF-36, it is important to test the assumption that items of a subscale measure the same underlying construct (dimensionality). For the summary scales it is important that the assumed two-factor structure can also be found in each population in which the SF-36 is applied, and that the underlying factor structure is similar to the factor structure found in the original data set used to determine the weighting factors for the calculations of the summary scales. Psychometric studies generally confirm the eight-dimensional structure of the SF-36 subscales and support the two-factor structure of the physical and mental summary measures across different patient groups and translations of the SF-36 [4–7]. The few studies that examined the dimensional structure of the SF-36 within patient groups with neurologic diseases reported less positive findings [8–10]. It is therefore necessary to further investigate the assumptions regarding dimensional structure of the SF-36, before using and comparing results of the SF-36 in neurologic patients.

As part of a 3-year follow-up study on functional prognosis in neurologic disorders, we investigated the dimensional structure of the SF-36 and its two summary measures in patients with stroke ($n = 198$), multiple sclerosis (MS, $n = 151$), and amyotrophic lateral sclerosis (ALS, $n = 188$). Principal component analysis (PCA) was applied, followed by orthogonal (Varimax) rotation in each patient group separately. To investigate the eight-dimensional structure of the SF-36, the number of factors to be rotated was restricted to the eight (originally proposed) factors. Items were considered to load on (were related to) a factor if factor loadings exceeded 0.40 on that factor and were lower on the other factors. Because all items had ordinal scores and several subscale scores showed a skewed distribution, PCA was carried out using polychoric correlations instead of Pearson correlations [11].

The percentage of total explained variance by eight extracted factors was 67, 78, and 73%, for the stroke, MS, and ALS groups, respectively. The number of extracted factors with eigenvalues exceeding 1.0 was nine for stroke, six for MS, and eight for ALS. All items of the Physical Functioning, Role Physical (RP), Bodily Pain, Social Functioning (SF), and Role Emotional subscales showed factor loadings exceeding 0.4 with the supposed factor in all groups (Table 1). However, items of the subscales General Health (GH) (in all groups), Vitality (VT) (in all groups), and Mental Health (in stroke) loaded on more than one factor (Table 1), showing factor loadings lower than 0.4 on the expected factor and higher factor loadings on other factors. In all three groups, the items of the General Health subscale measuring negative health status loaded on another factor than the items measuring positive health status. Also, the items of the Vitality subscale split up in items measuring positive and negative health status in all groups. These observations were also reported in earlier studies [8–10,12], indicating that the General Health and Vitality subscales are not measuring one single underlying construct in these patient groups. It is noteworthy

* Corresponding author. Tel.: 31204443063; fax: 31204440787.
E-mail address: A.Dallmeijer@vumc.nl (A.J. Dallmeijer).

doi: 10.1016/j.jclinepi.2005.11.001

Table 1
Results of principal component analysis of the SF-36 subscales

Domain	Number of Items	Range of factor loading of items with corresponding subscale (number of items with factor loadings >0.4)		
		Stroke	MS	ALS
Physical functioning	10	0.52–0.88 (10)	0.67–0.88 (10)	0.58–0.88 (10)
Role physical	4	0.57–0.78 (4)	0.70–0.79 (4)	0.63–0.80 (4)
Bodily pain	2	0.91–0.91 (2)	0.86–0.88 (2)	0.73–0.81 (2)
General health	5	<0.10–0.78 (3)	0.30–0.82 (3)	0.27–0.81 (4)
Vitality	4	0.18–0.63 (2)	0.22–0.72 (3)	0.27–0.77 (2)
Social functioning	2	0.48–0.50 (2)	0.45–0.52 (2)	0.66–0.77 (2)
Role emotional	3	0.80–0.83 (3)	0.76–0.84 (3)	0.71–0.78 (3)
Mental health	5	<0.10–0.79 (3)	0.55–0.74 (5)	0.62–0.79 (5)

Range of factor loadings of items with corresponding subscales are given: the number of items with a factor loading higher than 0.4 are given in parentheses.

that items measuring positive health status loaded on one factor and items measuring negative health status on the other, suggesting that our subjects had difficulties in changing between these reversed answering categories.

The factor structure of the physical and mental summary scales was studied using PCA on the subscale scores followed by orthogonal (Varimax) rotation on two factors. To support the two factor structure of the SF-36, three criteria should be met [2,4,6]: (1) eigenvalues of the first two factors >1, (2) the explained total variance >60%, (3) factor loadings (of subscales with summary scales) are in agreement with the originally hypothesized factor content. Eigenvalues of the first two factors exceeded 1.0 before rotation in the stroke and ALS group, but in MS, the second factor had a eigenvalue lower than 1.0 (0.87). The percentage of the total variance that was explained by two factors after rotation was satisfying in the MS group (69%) but was below 60% in the stroke (56%) and ALS groups (55%). Deviations from the hypothesized factor loadings were observed in several subscales (RP in stroke, SF scales in MS, GH in stroke and ALS, VT scale in MS), and the factor content differed between the groups. Although it is unclear to what extent deviations from the original factor structure are acceptable, the summary scores did not, in general, meet the criteria on several points. Other authors also found proportions of explained variance below 60% and deviating factor contents for the subscales with the physical and mental summary scores in patients with MS (9), stroke (8), and ALS (10) and raised questions about the use of the SF-36 summary scales in these populations. Our findings indicate that the use of the two summary measures in patients with stroke, MS, and ALS should be reconsidered.

To summarize, our results showed that the eight-dimensional structure of the SF-36 was generally confirmed, but results also suggest that the General Health and Vitality subscales are not suitable for use in patients with stroke, MS and ALS, due to lack of unidimensionality. Further research is required to determine whether the Mental Health subscale can be validly used in stroke patients. Application of the two summary scales should be reconsidered in these patient groups because the hypothesized physical and mental factor structure was not sufficiently supported. The number of subjects in the three diagnostic groups was at the lower end of the recommended number of cases for performing principal component analysis. Although this may limit the conclusions of this study to some extent, it is clear that caution should be taken when using the SF-36 in patients with stroke, MS, and ALS. In addition, the findings emphasize the importance of testing psychometric assumptions when using generic questionnaires in different patient groups.

Acknowledgments

This project was funded by the Netherlands Organization for Health Research and Development (Grant: 1435.0001). FuPro study group: G.J. Lankhorst, J. Dekker, A.J. Dallmeijer, M.J. IJzerman, H. Beckerman, V. de Groot: VU University Medical Center Amsterdam (project coordination); A.J.H. Prevo, E. Lindeman, V.P.M. Schepers: University Medical Center, Utrecht; H.J. Stam, E. Odding, B. van Baalen: Erasmus Medical Center, Rotterdam; A. Beelen: Academic Medical Center, Amsterdam, The Netherlands.

References

[1] Ware JE, Snow KK, Kosinski M, Gandek B. SF-36 Health survey manual and interpretation guide. Boston, MA: New England Medical Center, The Health Institute; 1993.

[2] Ware JE, Kosinski M, Keller SK. Physical and Mental Health Summary Scales: A user's manual. Boston, MA: New England Medical Center, The Health Institute; 1994.

[3] Nunnally JC, Bernstein IA. Psychometric theory. 3th ed. New York: McGraw-Hill; 1994.

[4] McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993;31(3):247–63.

[5] Gandek B, Ware JE Jr, Aaronson NK, Alonso J, Apolone G, Bjorner J, et al. Tests of data quality, scaling assumptions, and reliability of the SF-36 in eleven countries: results from the IQOLA

Project. International Quality of Life Assessment. J Clin Epidemiol 1998;51(11):1149–58.

[6] Ware JE Jr, Kosinski M, Gandek B, Aaronson NK, Apolone G, Bech P, et al. The factor structure of the SF-36 Health Survey in 10 countries: results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol 1998;51(11):1159–65.

[7] McHorney CA, Ware JE Jr, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care 1994;32(1):40–66.

[8] Hobart JC, Williams LS, Moran K, Thompson AJ. Quality of life measurement after stroke: uses and abuses of the SF-36. Stroke 2002;33(5):1348–56.

[9] Hobart J, Freeman J, Lamping D, Fitzpatrick R, Thompson A. The SF-36 in multiple sclerosis: why basic assumptions must be tested. J Neurol Neurosurg Psychiatry 2001;71(3):363–70.

[10] Jenkinson C, Hobart J, Chandola T, Fitzpatrick R, Peto V, Swash M. Use of the short form health survey (SF-36) in patients with amyotrophic lateral sclerosis: tests of data quality, score reliability, response rate and scaling assumptions. J Neurol 2002;249(2):178–83.

[11] Kim JO, Mueller ChW. Introduction to factor analysis what it is and how to do it. Beverly Hills, CA: Sage Publications; 1996.

[12] Moorer P, Suurmeije T, Foets M, Molenaar IW. Psychometric properties of the RAND-36 among three chronic diseases (multiple sclerosis, rheumatic diseases and COPD) in The Netherlands. Qual Life Res 2001;10(7):637–45.