

Validity and Reliability of the SF-36 Administered to Caregivers of Patients with Alzheimer's Disease: Evidence from a South American Sample

Gerardo Machnicki^a Ricardo F. Allegri^{a, b} Cristina G. Ranalli^a Cecilia M. Serrano^{a, b}
Carol Dillon^a Kathleen W. Wyrwich^c Fernando E. Taragano^b

^aMemory Research Center, Department of Neurology, Zubizarreta General Hospital, GCBA, and

^bDepartment of Neuropsychology (SIREN), CEMIC University, Buenos Aires, Argentina;

^cUnited BioSource Corporation, Bethesda, Md., USA

© S. Karger AG, Basel

**PROOF Copy
for personal
use only**

ANY DISTRIBUTION OF THIS
ARTICLE WITHOUT WRITTEN
CONSENT FROM S. KARGER
AG, BASEL IS A VIOLATION
OF THE COPYRIGHT.

Key Words

Quality of life • Caregiver • Alzheimer's disease • SF 36

Abstract

Background: Being a caregiver of a patient with Alzheimer's disease is associated with impaired health status and declines in health-related quality of life (HRQoL). This paper evaluates the reliability and validity of the Argentinean version of the Medical Outcomes Study Short-Form Health Survey (SF-36) among caregivers of patients with Alzheimer's disease. **Methods:** Forty-eight caregivers of Alzheimer's disease patients completed the SF-36, the Zarit Burden Interview (ZBI) and the Neuropsychiatric Inventory (NPI). Patients were evaluated for dementia severity using the Clinical Dementia Rating (CDR) and for cognitive status using the Mini Mental State Examination (MMSE). **Results:** The SF-36 scales demonstrated adequate-to-strong internal consistency (Cronbach's alpha range: 0.72 to 0.92). Correlations between the SF-36 scales and the ZBI were moderate to strong (range: -0.19 to -0.79, all $p < 0.01$ except for physical function). Significant correlations between the SF-36 scales and the CDR, MMSE and NPI were lower (range: -0.30 to -0.40, $p < 0.001$) and strongest in mental health-related scales of the SF-36.

The SF-36 demonstrated good factorial validity. **Conclusions:** The Argentinean translation of the SF-36 is reliable and valid for use to measure the HRQoL of caregivers of patients with Alzheimer's disease.

Copyright © 2009 S. Karger AG, Basel

Caregiving to elderly relatives by family members increased almost 3-fold in the last century due to the growing incidence and prevalence of chronic and degenerative disorders [1]. Dementia is one of the diseases that require caregiving due to cognitive and behavioral disturbances. Caregiver burden in dementia should be also of concern to developing countries with aging populations [2].

In the last 25 years, both theoretical frameworks and empirical evidence about caregiving have been developed. Caregiving is a multidimensional concept [3], with an impact on both physical and mental health, as well as on family finances and their structure of time. All these impacts have been shown to negatively affect the caregivers of patients with dementia [4–6], in some cases, more than in the caregivers of non-demented patients [7].

Health-related quality of life (HRQoL) is an important indicator of health. The HRQoL combines the presence

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

© 2009 S. Karger AG, Basel
1420-8008/09/0000-0000\$26.00/0

Accessible online at:
www.karger.com/dem

Ricardo F. Allegri, MD, PhD
Department of Neuropsychology (SIREN)
CEMIC University
Galvan 4102, 1431 Buenos Aires (Argentina)
Tel. +54 11 4546 8227, Fax +54 11 4433 3443, E-Mail rallegr@cemic.edu.ar

of symptoms with the subjective impact of those as perceived by the person experiencing those symptoms [8]. HRQoL instruments are often classified as generic or disease-specific. Generic instruments have the advantage of allowing comparisons across different diseased and healthy populations; while disease-specific measures are designed for measurement within a given disease area, the disease-specific measure may be more sensitive to detect differences and changes [8].

The HRQoL in caregivers has been studied, among other instruments, with the Medical Outcomes Study Short-Form Health Survey (SF-36) [9–13], a widely used generic HRQoL measure [14]. As with any other form of research, it is very important to use reliable and valid questionnaires when measuring the HRQoL. Assessing the reliability and validity is a continuous process, as questionnaires expand for use in different populations and geographical areas, and as instruments are created or modified [15].

To our knowledge, the psychometric characteristics of the SF-36 in current caregivers of patients with Alzheimer's disease in South America have not been examined. Berg-Weger et al. [16] examined the reliability and validity of the SF-36 in former caregivers of patients with Alzheimer's disease in Missouri, USA. Reliability was acceptable for the full scale and summary scales, and generally acceptable for the 8 scales. Factorial, convergent and discriminative validity results also supported the SF-36 as a promising tool for understanding post-caregiver health, although this study was limited due to only one setting. Building on the findings of Berg-Weger et al. in former and not current caregivers, the objective of this research was to determine the reliability and validity of the SF-36 questionnaire in caregivers of patients with Alzheimer's disease in Buenos Aires, Argentina.

Methods

Patient Population

The study population was derived from the Memory Clinic at Zubizarreta Hospital (Buenos Aires, Argentina), a reference hospital for patients with cognitive problems. Since 2001, the clinic has implemented a clinical database with routine data collection beginning at all initial visits of the patients and their primary caregivers. The clinical protocol records demographic information, clinical variables (functional status, neurological exams, and neuropsychiatric evaluation) and caregiver burden.

The SF-36 was administered for this cross-sectional study to caregivers of patients with Alzheimer's disease, and the responses of all caregivers since 2001 who completed the SF-36 were captured in the clinical database. Only primary caregivers, defined

as those in charge of the majority of caregiving tasks, were included, and all patients cared for by those caregivers had a diagnosis of probable Alzheimer's disease by NINCS-ADRDA criteria [17]. This study received local Institutional Review Board approval and appropriate informed consent was obtained from each participant dyad (patients and their caregivers).

Caregiver Measures

Caregivers completed 2 self-administered questionnaires in a private room at the clinic: the SF-36 (version 1.0, Argentinean translation) and the Zarit Burden Index (ZBI). The SF-36 is the most widely-used generic health measure in the world [18, 19]. Thirty-five of the 36 items load into 8 scales: physical functioning, role limitations due to physical health, role limitations due to physical problems (role physical), bodily pain, general health perceptions, vitality, social functioning, role limitations due to emotional problems (role emotional) and mental health. One additional item that measures health transition over the past year is asked but not included in any SF-36 scale score. Standardized scores that range from 0 (worst health) to 100 (best health) are obtained for each scale. The first 4 scales named above have primary associations with the physical component summary measure, while the last 4 have primary associations with the mental component summary measure. The reliability and validity of the SF-36 have been extensively studied in general and diseased populations [18]. The SF-36 has been validated successfully in Argentina [20].

The ZBI is a widely used scale of caregiver burden [21], and the Spanish adaptation for Argentina was used [22] in this study. It consists of 22 items that are self-administered. A single score is obtained ranging from 0 (lowest caregiver burden) to 88 (highest caregiver burden).

Alzheimer's Disease Patient Measures

Three types of Alzheimer disease patient assessments were considered: the Mini Mental Scale Examination (MMSE) [23], the Clinical Dementia Rating (CDR) [24] and the Neuropsychiatric Inventory (NPI) [25].

The MMSE is a widely used cognitive impairment measure. It contains 11 items and it is interview-administered. A 0 to 30 score is obtained, with lower scores indicating higher cognitive impairment [23]. The Spanish-for-Argentina version was used [26].

The CDR was used to assess the severity of dementia. It is comprised of 6 items with a 0 to 3 summary score. Higher scores represent greater severity [24].

The NPI assesses neuropsychiatric symptoms. It covers 12 items, including psychotic, frontal, affective/negative and other symptoms. The existence, frequency and severity caused by symptoms are considered. A total score ranging from 0 to 144 is obtained, with higher scores representing higher behavioral symptoms [25].

Statistical Analysis

All analyses were performed with SPSS version 12 (SPSS Inc., Chicago, Ill., USA) and Lisrel 8.80 (Scientific Software International, Inc., Lincolnwood, Ill., USA). Scoring was performed according to each questionnaire's guidelines. Imputations for the SF-36 missing data, if necessary, were performed according to the developer's guidelines [26].

Reliability

Reliability examines the extent to which the score of a scale performs consistently in a short period of time (evaluating if items perform consistently within a subscale). When a measure is collected at a single point in time, reliability is assessed through internal consistency reliability. This assessment evaluates the average of all inter-item correlations, among items that are supposed to measure the same underlying factor.

Cronbach's alpha statistic was used to assess the internal consistency reliability. This is the least conservative way to assess reliability; however, the lack of retest data in this study made it impossible to evaluate test-retest reliability or the stability of measures over time. A Cronbach's alpha value ≥ 0.70 is considered as evidence of acceptable reliability [15].

Validity

Validity refers to whether a scale is capturing what is intended to be measured. Construct validity can be assessed with 3 subtypes of validity concepts: (i) extreme groups (or discriminative), (ii) convergent and (iii) discriminative validity [8]. Convergent and discriminative validity were considered in this research. As the ZBI measures a construct related to HRQoL (how burdensome patient symptoms are for the caregiver), convergent validity was examined using the correlation between the SF-36 scale scores and the ZBI, and negative correlations were hypothesized. Also, given that the ZBI was the closest measure to the HRQoL, it was expected that the observed ZBI correlations should be the highest among all correlations in this analysis of the SF-36. Correlations coefficients between the ZBI and the 8 SF-36 scale scores were also analyzed.

Discriminant construct validity was examined using the correlation between the SF-36 scale scores and the MMSE, NPI and CDR. It was hypothesized that the SF-36 scale scores would be positively correlated to MMSE (higher MMSE means a patient with higher cognitive status, and this should be positively correlated to higher SF-36 scores), and negatively correlated to NPI and CDR (the higher the NPI and CDR, the worse the status of the patient from a behavioral point of view).

Given the small sample size and non-normality in the measures, a non-parametric measure of correlation (Spearman's ρ) was used for all correlation analyses.

Factorial Validity

Factorial validity examines whether the observed responses correspond to the latent (i.e., unobserved) constructs being postulated. In the case of the SF-36, the second-order factorial structure was examined. The second-order factor structure hypothesized that the 8 observed scales are indicators of the 2 latent constructs: physical and mental health. The factorial structure of the SF-36 has been examined both in the general population [27] and in former caregivers of patients with Alzheimer's disease [16], but not in current caregivers.

Factorial validity was evaluated with confirmatory factor analysis. This approach is appropriate because the objective is to assess if the observed measures correspond to a given factorial structure already observed or postulated [25]. The second-order factorial structure of the SF-36 has been previously determined [16, 27]. Measures of model fit examine if the hypothesized covariance structure can account for the observed covariance in the data. The confirmatory factor measurement model was estimated

Table 1. Caregivers: demographic and caregiving variables

Variable	Value
Mean age, years	58.8 \pm 14.9
29–65 years	32 (66.7%)
66–89 years	16 (33.3%)
Gender	
Female	41 (85.4%)
Male	7 (14.6%)
Mean education, years (range)	9.65 \pm 3.46 (4–17)
Relationship with patient	
Spouse	24 (50.0%)
Children	18 (37.5%)
Other	6 (12.5%)
Caregiver residing with patient	36 (75%)
Other paid caregiver	3 (6.2%)
Caregiving	
Mean hours per week (range)	33.7 \pm 18.3 (6–56)
Caregiver burden	
Mean Zarit score (range)	33.19 \pm 19.64 (2–84)

Values do not always add up to 48 due to missing values.

with the Satorra-Bentler χ^2 statistic (with $p > 0.05$ as evidence of no difference between the observed and hypothesized covariance structure), the root mean square error of approximation (RMSEA, ≤ 0.05 or with the 90% confidence interval (CI) containing 0.05 as evidence of good fit), the Comparative Fit Index (CFI) and Non-Normed Fit Index (NNFI) (both ≥ 0.9 for good fit). Given a small sample size and non-normal data, a robust maximum likelihood estimation with asymptotic covariance matrix was used.

Results

Study Sample

Fifty-two caregivers were enrolled, of which 48 caregivers had analyzable SF-36 responses. Missing SF-36 data were imputed for 3 caregivers and involved no more than 3 responses missing (of the 36 questions) in each case. The caregivers' characteristics are presented in table 1.

Two-thirds of the caregivers were in the 29–65 years age category, indicating that they were potentially economically active. Eighty-five percent were female. The mean time dedicated to caregiving was 6.83 h per week. The mean ZBI score was 33 on the 0–88 scale.

Table 2 shows the demographic and clinical measures of the patients. There was a near-even distribution of patient ages between the 58–75 and the 76–89 years age categories. Fifty percent of the patients were female, and the

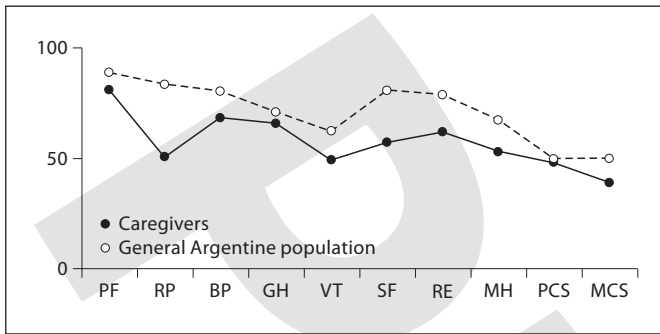


Fig. 1. SF-36 scale and summary scores: study population and Argentinean population norms [20]. PF = Physical functioning; RP = role limitations due to physical health; BP = bodily pain; GH = general health perceptions; VT = vitality; SF = social functioning; RE = role limitations due to emotional problems; MH = mental health; PCS = physical component summary; MCS = mental component summary.

mean time since diagnosis was 40 months. All patients showed some degree of dementia severity (50% of patients had a CDR score of 2 or more). Fifty-five percent of the patients had MMSE ratings ≤ 20 , and the mean NPI was 32.

Figure 1 shows the SF-36 mean scores for each scale and the summary scales. For comparison, the Argentinean population norms are also shown [20]. The HRQoL of the study sample was lower, particularly in the mental health-related scales and summary score.

Internal Consistency Reliability

Internal consistency reliability of the SF-36 scales was acceptable-to-good for all subscales (range: 0.72–0.92) (table 3). For reference, the results of the validation study of Berg-Weger et al. [16] are included. There was no clear pattern of similarities or differences between the Berg-Weger and colleagues reliability results and the current findings. In some scales, reliability in the current study was higher (role physical, role emotional, general mental health and bodily pain) but this was not always the case.

Concurrent and Discriminative Validity

Correlations between the SF-36 scales and the other measures are presented in table 4. As hypothesized, the correlations between the SF-36 mental health-related scales and the ZBI were highest; the higher SF-36 scales correlations for the ZBI were with social function ($r = -0.79$), mental health ($r = -0.47$) and vitality ($r = -0.47$), while all the others were in the $r = -0.42$ to -0.33 range, except for physical function where $r = -0.19$. All the cor-

Table 2. Patients: demographic and clinical variables

Variable	Value
Mean age, years	74.7 \pm 7.4
58–75 years	23 (47.9%)
76–89 years	25 (52.1%)
Gender	
Female	24 (50.0%)
Male	24 (50.0%)
Education, years (range)	8.06 \pm 3.96 (1–18)
<i>Clinical variables</i>	
Mean duration of illness in months (range)	39.91 \pm 32.81 (6–180)
Mean CDR score	1.6 \pm 0.9
0.5	6 (13.3%)
1	16 (35.6%)
2	11 (24.4%)
3	12 (26.7%)
Mean MMSE score	17.6 \pm 7.4
0–10	9 (20.0%)
11–20	16 (35.6%)
21–30	20 (44.4%)
Mean NPI score (range)	32.35 \pm 28.41 (0–84)
Values do not always add up to 48 due to missing values.	

Table 3. Reliability of the SF-36 scales

Scale	Caregivers in Buenos Aires (n = 48)	Former caregivers in Missouri [16] (n = 102)
Physical functioning	0.76	0.84
Role physical	0.92	0.69
Role emotional	0.87	0.64
Social functioning	0.78	0.88
General mental health	0.86	0.75
Vitality	0.79	0.84
Bodily pain	0.87	0.80
General health perception	0.72	0.78

relations between the SF-36 scales and the ZBI were statistically significant ($p < 0.01$), except for physical functioning.

Correlations between the MMSE, CDR, NPI and the SF-36 scales were low-to-moderate when statistically significant (absolute value range: 0.31–0.40). The correlations of the SF-36 scales with these measures were strongest between the mental health-related SF-36 scales. The correlations of the SF-36 role emotional and social functioning scales with the MMSE, CDR and NPI were the

Fig. 2. Confirmatory factor model: path diagram (standardized estimates) for the SF-36 domains. MH score = Mental health score; EV score = vitality score; RE score = role limitations due to emotional problems score; SF score = social functioning score; HP score = general health perceptions score; PF score = physical functioning score; RP score = role limitations due to physical health score; P score = bodily pain score.

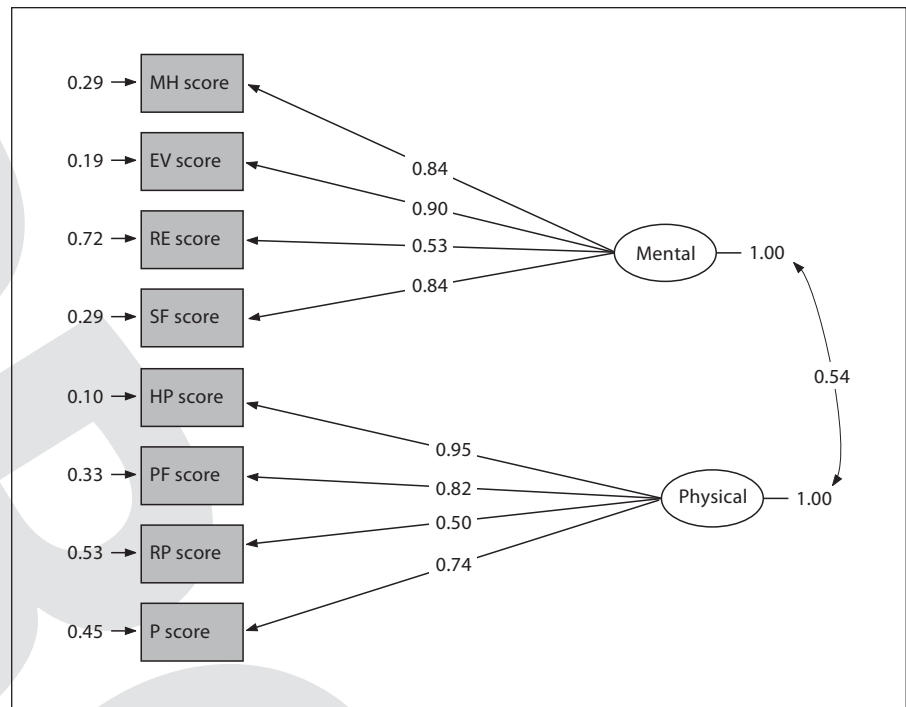


Table 4. Convergent and discriminative validity: Spearman's ρ

SF-36 scale	ZARIT	MMSE ^a	CDR ^a	NPI
Physical function	-0.19	0.13	-0.08	-0.03
Role physical	-0.33 ^b	0.38 ^b	-0.34 ^b	-0.29
Role emotional	-0.33 ^b	0.37 ^b	-0.38 ^b	-0.24
Social functioning	-0.79 ^b	0.30 ^b	-0.34 ^b	-0.40 ^b
Mental health	-0.47 ^b	0.31 ^b	-0.28	-0.35 ^b
Vitality	-0.47 ^b	0.17	-0.23	-0.19
Bodily pain	-0.44 ^b	0.20	-0.20	-0.24
General health perception	-0.42 ^b	0.10	-0.11	-0.27

^a n = 45 due to missing values; ^b statistically significant correlations, $p < 0.05$ or less.

highest and statistically significant ($p < 0.05$). In contrast, none of the correlations with physical function, vitality, pain and general health perception were statistically significant.

Factorial Validity

Figure 2 shows the standardized factor loadings of the confirmatory factor model. All the loadings exceeded the critical value of 0.60, except for the role emotional (loading = 0.53) and role physical scales (factor loading = 0.60).

Five of the SF-36 scale loadings were above 0.80. A correlation of $r = 0.54$ was observed for the hypothesized summary constructs of physical health and mental health. All the loadings and the correlation between physical and mental health were statistically significant ($p < 0.05$).

The model fit results were mixed. The Satorra-Bentler χ^2 statistic ($\chi^2 = 40.19$, d.f. = 19, $p < 0.031$) and the RMSEA (0.15, 90% CI: 0.087–0.22) yielded inadequate measures of fit. In contrast, the CFI = 0.93 and NNFI = 0.85 indicated a good fit between the model and the observed covariance between the model variables, which were the SF-36 scale scores.

Discussion

Caregiver burden has been recognized by the clinical community as an important and multidimensional problem. As such, it can be examined from different perspectives, including the HRQoL. This research supports the use of the SF-36, a widely used HRQoL instrument, as one tool that can be validly and reliably used when studying the burden of caregiving on the HRQoL in Argentina. The SF-36 can be used to assess the health of caregivers and to evaluate the impact of interventions aimed at reducing their burden, and subsequently, improving the HRQoL.

Reliability of the SF-36 was supported by Cronbach's alpha coefficients above 0.70, a critical threshold for acceptable reliability. Good reliability is a prerequisite for any valid research as it indicates an instrument measures an underlying construct with acceptable accuracy, and random error is not dominating the measurement process. Results for the 8 SF-36 scales were comparable and sometimes higher than those found in similar research by Berg-Weger et al. [16].

Concurrent and discriminative validity were supported by the correlations among the SF-36 scales and the measures of caregiver burden (ZBI), cognitive status (MMSE), dementia severity (CDR), and neuropsychiatric symptoms (NPI). Correlations were in a logical direction and the strengths of the correlations were as hypothesized. The strongest correlations were observed between the SF-36 social functioning, mental health, and vitality scales when compared with the ZBI. These 2 instruments (SF-36 and ZBI) are the closest among all the measures examined, as the ZBI focuses on the psychosocial burden of caregiving. Moreover, both of these measures were directly solicited and focus on the caregivers. The MMSE, CDR and NPI had comparable patterns of correlations with the SF-36 subscales, but correlations were smaller in magnitude than those involving the ZBI. All Alzheimer's disease patient measures and the ZBI were at least moderately correlated with the SF-36 role emotional and social functioning scales; 2 were correlated with the SF-36 role physical scale and 2 with the SF-36 mental health scale. Although Berg-Weger et al. [16] used different constructs, the strength of correlations in that validation study was similar to the relationships reported here. In summary, these correlations support the validity of SF-36 scales measured among current caregivers.

Factorial validity examines the strength of the assessed constructs by relating the measured variables to the hypothesized constructs, in this case, summary constructs for physical and mental health. Our confirmatory factor model yielded strong factor loadings for all scales except for the SF-36 role emotional scale, and a borderline value for the SF-36 role physical scale. The measurement challenges in these 2 scales are a well-known feature of the SF-36 version 1 that was used in this study [28]. The items in these 2 scales have only 'Yes/No' response options, and hence, yield a generally low variation across scores that limit correlational modeling. The SF-36 version 2 improved both of these 2 scales by rephrasing the items and adding 5 response levels, and therefore, allowing more variation within these 2 scales [29]. The results of the measurement model fit suggested adequate fit with

some of the indicators and poor fit with others. The correlation between the mental and physical construct was $r = 0.50$. In Berg-Weger et al. [16], this correlation was even higher ($r = 0.76$). These high correlations suggest a strong connection between physical and mental health as measured by the SF-36.

This study has strengths and limitations. The validity was examined using correlations and structural equation modeling techniques that shed light onto the relationship between the observed and latent measures. However, the very small sample size made it impossible to analyze the first-order loading factors (the loading of the 35 questions onto each of the 8 scales). Also, given the small sample size, the estimates should be considered with caution, and re-examined in larger, relevant samples of caregivers. This study was also limited given the cross-sectional design, making it impossible to calculate test-retest reliability of the SF-36 scales. Finally, and also due to the cross-sectional nature of this research, the responsiveness over time in the SF-36 scores was not evaluated.

Nonetheless, the evidence from reliability, the different types of validity, the factor loadings and some of the measurement model fit statistics suggest that the SF-36 is adequate to investigate the HRQoL of caregivers of patients with Alzheimer's disease. Further work should be undertaken in Latin America to confirm these results and to assess the several dimensions of caregiving burden in caregivers of patients with Alzheimer's disease and dementia.

Conclusions

The Argentinean version of the SF-36 is reliable and valid for use in caregivers of patients with Alzheimer's disease. This instrument demonstrated adequate reliability and concurrent, discriminative and factorial validity. The SF-36 could be used to track HRQoL outcomes in caregivers of patients with Alzheimer's disease.

Acknowledgements

The authors thank Dr. Kathleen Wyrwich for her helpful comments on this article. This research was jointly supported by a grant in aid for scientific research (Carrillo Oñativia) from the Ministry of Health of Argentina (C.D., and R.F.A.), and a research grant from the Secretary of Health of Buenos Aires Government (R.F.A., C.G.R. and C.S.). The views expressed in this publication are those of the authors and not necessarily those of the Ministry of Health of Argentina or those of the Secretary of Health of Buenos Aires Government.

References

- 1 Wagner L: Comparative Analysis of Caregiver Data for Caregivers to the Elderly. 1987 and 1997. Bethesda, National Alliance for Caregiving, 1997.
- 2 Fratiglioni L, De Ronchi D, Aguero-Torres H: Worldwide prevalence and incidence of dementia. *Drugs Aging* 1999;15:365–375.
- 3 Pearlin LI, Mullan JT, Semple SJ, Skaff MM: Caregiving and the stress process: an overview of concepts and their measures. *Gerontologist* 1990;30:583–594.
- 4 Leinonen E, Korpisammal L, Pulkkinen LM, Pukuri T: The comparison of burden between caregiving spouses of depressive and demented patients. *Int J Geriatr Psychiatry* 2001;16:387–393.
- 5 Thommessen B, Aarsland D, Braekhus A, Oksengaard AR, Engedal K, Laake K: The psychosocial burden on spouses of the elderly with stroke, dementia and Parkinson's disease. *Int J Geriatr Psychiatry* 2002;17:78–84.
- 6 Machinicki G, Allegri RF, Dillon C, Serrano C, Taragano FE: Cognitive, functional and behavioral factors associated with the burden of caring for geriatric patients with cognitive impairment or depression: evidence from a South American sample. *Int J Geriatr Psychiatry* 2009;24:382–389.
- 7 Gonzalez-Salvador MT, Arango C, Lyketsos CG, Barba AC: The stress and psychological morbidity of the Alzheimer patient caregiver. *Int J Geriatr Psychiatry* 1999;14:701–710.
- 8 Streiner DL, Norman GR: Health measurement scales: a practical guide to their development and use. Oxford, Oxford University Press, 2003.
- 9 Cannuscio CC, Jones C, Kawachi I, Colditz GA, Berkman L, Rimm E: Reverberations of family illness: a longitudinal assessment of informal caregiving and mental health status in the Nurses' Health Study. *Am J Public Health* 2002;92:1305–1311.
- 10 Clyburn LD, Stones MJ, Hadjistavropoulos T, Tuokko H: Predicting caregiver burden and depression in Alzheimer's disease. *J Gerontol B Psychol Sci Soc Sci* 2000;55:S2–S13.
- 11 Dilworth-Anderson P, Williams SW, Cooper T: Family caregiving to elderly African Americans: caregiver types and structures. *J Gerontol B Psychol Sci Soc Sci* 1999;54:S237–S241.
- 12 Fillit H, Cummings J: Practice guidelines for the diagnosis and treatment of Alzheimer's disease in a managed care setting: part II – pharmacologic therapy. *Alzheimer's Disease (AD) Managed Care Advisory Council. Manag Care Interface* 2000;13:51–56.
- 13 Neumann PJ, Kuntz KM, Leon J, Araki SS, Hermann RC, Hsu MA, et al: Health utilities in Alzheimer's disease: a cross-sectional study of patients and caregivers. *Med Care* 1999;37:27–32.
- 14 Cramer JA, Spilker B: *Quality of Life and Pharmacoeconomics: An Introduction*. Philadelphia, Lippincott-Raven, 1998.
- 15 Nunnally J, Bernstein I: *Psychometric Theory*. New York, McGraw-Hill, 1994.
- 16 Berg-Weger M, Rauch SM, Rubio DM, Tebb SS: Assessing the health of adult daughter former caregivers for elders with Alzheimer's disease. *Am J Alzheimers Dis Other Demen* 2003;18:231–239.
- 17 McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM: Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984;34:939–944.
- 18 Ware JE Jr, Snow KK, Kosinski M, Gandek B: *SF-36 Health Survey manual and interpretation guide*. Boston, The Health Institute, New England Medical Center, 1993.
- 19 Ware JE Jr, Sherbourne CD: The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473–483.
- 20 Augustovski F, Lewin G, García Elorrio E, Rubinstein A: The Argentine-Spanish SF-36 Health Survey was successfully validated for local outcome research. *J Clin Epidemiol* 2008;61:1279–1284.
- 21 Zarit SH, VandenBos GR: Effective evaluation of memory in older persons. *Hosp Community Psychiatry* 1990;41:9–10, 16.
- 22 Allegri RF, Sarasola D, Serrano CM, Taragano FE, Arizaga RL, Butman JL, Loñ L: Neuropsychiatric symptoms as a predictor of caregiver burden in Alzheimer's disease. *Neuropsychiatr Dis Treat* 2006;2:105–110.
- 23 Folstein MF, Folstein SE, McHugh PR: 'Minimal state'. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–198.
- 24 Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL: A new clinical scale for the staging of dementia. *Br J Psychiatry* 1982;140:566–572.
- 25 Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J: The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology* 1994;44:2308–2314.
- 26 Allegri RF, Ollari JA, Mangone CA, Arizaga RL et al: 'El Mini Mental State Examination' en la Argentina: Instrucciones para su administración. *Revista Neurológica Argentina* 1999;24:31–35.
- 27 Reed PJ: Medical Outcomes Study Short Form 36: testing and cross-validating a second-order factorial structure for health system employees. *Health Serv Res* 1998;33(5 Pt 1):1361–1280.
- 28 de Vet HC, Ader HJ, Terwee CB, Pouwer F: Are factor analytical techniques used appropriately in the validation of health status questionnaires? A systematic review on the quality of factor analysis of the SF-36. *Qual Life Res* 2005;14:1203–1218; discussion 1219–1221, 1223–1224.
- 29 Hawthorne G, Osborne R, Taylor A, Sansoni J: The SF-36 version 2: critical analyses of population weights, scoring algorithms and population norms. *Qual Life Res* 2007;16:661–673.