ORIGINAL REPORT

QUALITY OF LIFE AMONG SWEDISH PATIENTS AFTER STROKE: PSYCHOMETRIC EVALUATION OF SF-36

Ann-Helene Almborg, MSc[1, 2] and Stig Berg, PhD[1]

From the [1]Institute of Gerontology, School of Health Sciences, Jönköping University, Jönköping and [2]The Vårdal Institute, The Swedish Institute for Health Sciences, Lund, Sweden

Objective: **Few studies have evaluated the scale assumptions of the Medical Outcomes Study 36-item Short-Form (SF-36) with stroke survivors. The aims of this study were to evaluate the scale assumptions of SF-36 using Swedish patients after stroke and to compare patients' quality of life with that of a healthy population.**

Design: **Cross-sectional study.**

Subjects: **SF-36 was tested in 188 patients (mean age 74 years) with acute stroke consecutively enrolled from a stroke unit in southern Sweden during 2003–05.**

Methods: **Data were collected by interview during a home visit 2–3 weeks after discharge. Psychometric analyses were conducted, and stroke survivors' quality of life was compared with a Swedish normal population.**

Results: **The internal consistency reliability was > 0.70 for all scales. There were notable floor and/or ceiling effects for 3 scales. For 7 scales, there was the expected association with the 2 summary scales. Compared with a normal Swedish population, stroke has a negative effect on health-related quality of life, especially for patients aged 45–54 years.**

Conclusion: **SF-36 functions well as a measure of health-related quality of life in Swedish patients after stroke, but the 2 summary scales have shortcomings. Our findings support good divergent validity of SF-36 for discriminating health-related quality of life of stroke groups and normal populations.**

INTRODUCTION

In Sweden, approximately 30,000 people suffer from stroke every year (1). In most countries, stroke is the third largest cause of death, and it often leads to functional disability, which negatively affects quality of life. Therefore, assessment of quality of life in patients after stroke is important in clinical practice, research and health policy evaluations. The Medical Outcomes Study 36-item Short-Form (SF-36) is a widely used standardized, generic self-report of health status for evaluating physical and mental health-related quality of life (HRQoL). It is recommended for use in clinical practice, clinical research, general population surveys and health policy evaluations (2, 3), although there might be some problems using "standard" HRQoL measures to assess people with disabilities (4).

The SF-36 consists of 36 items; 35 of the items group into 8 multi-item scales that collectively measure HRQoL: Physical Functioning (PF) (10 items), Role Physical (RP) (4 items), Bodily Pain (BP) (2 items), General Health (GH) (5 items), Vitality (VT) (4 items), Social Functioning (SF) (2 items), Role Emotional (RE) (3 items), and Mental Health (MH) (5 items). The remaining item concerns the experience of change in general health during the last year (2). The 36 items are measured on Likert-type scales and the response choices vary between 2 and 6 levels. The scores from the 8 scales are transformed, with each scale ranging from 0 (worst health) to 100 (best health); the item scores are averaged to create each scale score. After the 8 scores are weighted (2), 2 summary scales, the Physical Component Summary (PCS) and the Mental Component Summary (MCS), are calculated and standardized to a mean value of 50 and a standard deviation (SD) of 10.

The psychometric properties of SF-36 have been evaluated with normal populations in Sweden (2, 5) and several other countries. A few studies (6–10) have evaluated the psychometric assumptions of the SF-36 regarding data quality, scaling, reliability and construct validity among patients after stroke, but no studies have evaluated these assumptions in Swedish patients after stroke. Hobart et al. (7) found in a stroke group, that 2 scales (GH, SF) had low Cronbach's alpha, limited convergent and discrimant validity. Three scales (RP, RE, BP) had considerable floor and/or ceiling effects, which limit the usefulness in clinical groups. The 2 component summary scales were not supported in their study, as the correlations between the scales and the 2 components were inconsistent. The aims of this study were to evaluate the psychometric assumptions of the SF-36 regarding data quality, scaling, reliability and construct validity in Swedish patients after stroke and to estimate the impact of stroke on HRQoL by comparison with the Swedish normative population study.

METHODS

Patients

We considered for enrolment all patients from 2 Swedish municipalities (62,000 inhabitants) admitted to the stroke unit at a hospital in southern Sweden between 1 October 2003 and 30 November 2005. The stroke unit

cared for approximately 95% of all patients with acute stroke in the 2 municipalities. Patients with a medical diagnosis of stroke (International Classification of Diagnosis (ICD)-10: I61, I63 and I64) were consecutively enrolled. Patients with severe aphasia, cognitive impairment, or difficulty understanding the Swedish language were excluded. Cognitive impairment was defined as a score < 24 on the Mini Mental State Examination (MMSE) (11). If there was any doubt that the patients could understand the questions or if they could not express their answers due to aphasia they were excluded. During the first 5 months of enrolment, patients < 65 years of age and patients with earlier stroke episodes were excluded, but in order to increase the size of the study group these exclusion criteria were dropped starting on 1 March 2004. A total of 321 patients were admitted during the study period. Thirty-eight died before the interview, 23 declined participation, 28 had severe aphasia, 38 had cognitive impairment, 2 could not be contacted, and 4 had difficulties understanding the Swedish language. The final sample (*n* = 188) comprised 105 males (60%).

At the hospital, patients were provided oral and written information about the study. Patients were included only if they gave oral consent to participate. Ethical approval for the study was obtained from the Research Ethics Committee, Linköping University (Dnr 03-328).

Interviews and assessments

Patients were interviewed in their homes or in care homes 2–3 weeks after discharge. All interviews were conducted by the first author using the Swedish version of SF-36 (2). Cognitive function was assessed using the MMSE and Personal-Activity of Daily Life (P-ADL) performance was measured with the Barthel Index (BI). The BI can vary between 0 and 100 points, where 100 is totally independent functioning. The measurements were made 2–3 weeks after discharge. Medical and demographic data were collected by reviewing medical records and during the interview.

Statistical analyses

Tests of scaling assumptions with the summated-rating method were performed using descriptive statistics for item responses, equivalence of items' mean score, SD, skewness, and item-scale correlation (item-own-scale correlation and item-other-scale correlation). Different minimum values have been recommended for item-own-scale correlation to examine item convergent validity; these minimum values vary between 0.20 and 0.40 (12, 13). In this study we used $r \geq 0.40$. Discriminant validity of items and rates of scaling success were analysed by comparing the differences between item-own and item-other scale correlations. Differences of more than 2 standard errors of the correlation coefficient ($2 \times 1/\sqrt{n}$) were considered significant (5). Cronbach's alpha was used to examine the scales' internal consistency reliability. Descriptive statistics are reported for each scale score. A principal component analysis with oblimin rotation was used to analyse construct validity for the 2 summary scales (PCS and MCS).

The Student's *t*-test was used to compare mean values between genders. Effect size (ES) was calculated to compare the mean values of the whole stroke group and the mean values of the age groups 45–64 years, 65–74 years and ≥ 75 years with the similar groups in Swedish normative population study (2). Two patients younger than 45 years were excluded from the analysis by age group. The Cohen's effect size (d) was defined as the differences between the 2 samples mean values divided by the pooled SD (14). An effect size of 1 is equivalent to a change of 1 SD between the groups. When using this approach an ES of 0.2 is considered as small effect, 0.5 medium and 0.8 large effect. The Pearson correlations coefficient was used to compare BI and the SF-36 and educational level and the SF-36. Analyses were performed using of Statistical Package for the Social Sciences (SPSS) 13.0 (version 13.0). Statistical significance level was set at *p* < 0.05 (2-tailed).

RESULTS

The mean age (SD) of the total sample was 74.0 years (11.2), with a range of 32–92 years. The mean age (SD) was 75.5 years (11.3) for females and 72.8 years (10.9) for males. Most

patients lived with someone (59%) (Table I). Seventy-five percent had completed only elementary school. Brain infarction (I63) was the most common reason (93%) for stroke; for 80% of patients this was the first-ever stroke. The average hospital stay was 20.6 days (SD = 23.8, range = 2–130). The patients P-ADL performance measured by BI had a mean value of 92.3 (SD = 14.4, range = 40–100) scores.

The results of our tests of scaling assumptions with the summated-rating method, including the response option frequencies and missing data. The percentage of missing data ranged between 1.6% and 2.7%. The response option frequencies ranged between 0.5% and 88.1%, and there was both positive and negative skewness for items. There was substantial positive skewness for all items in the RP scale, indicating poorer health, and there was negative skewness for all items in the MH, indicating better health.

Correlation coefficients for the item-own-scale correlations exceeded 0.40 for 94% of the items. Two items, "I seem to get sick easier than others" (11a) and "I expect my health to get worse" (11c), in the GH scale correlated with a value > 0.30. The item-own-scale correlation for "Overall rating of general health" (item 1) within the GH scale was 0.56, but this item had a higher correlation with the PF (0.60) and VT (0.57) scales. The correlation coefficient for the item "I feel worn out" (9g) and its own scale (VT) was 0.44; it had a slightly higher correlation with the MH scale (0.49). Eight items (3a, 9a, 9e, 9h, 9i, 11a, 11b, and 11d) were correlated with the expected scale, but these correlations were not significant, as the item-own-scale correlations did not exceed by 2 standard errors the item-other-scale correlation.

Table I. *Patients' characteristics (n = 188)*

	n (%)
Age, years	
≤ 64	39 (20.7)
65–74	49 (26.1)
≥ 75	100 (53.2)
Education, years	
Elementary school ≤ 7	141 (75.0)
Secondary school 8–9	16 (8.5)
High school 10–12	20 (10.6)
University > 12	11 (5.9)
Living arrangement	
Living alone	77 (41.0)
Living together	111 (59.0)
Own home	180 (94.8)
Nursing home	8 (5.2)
Diagnosis, ICD	
I61 – brain haemorrhage	13 (6.9)
I63 – brain infarction	174 (92.6)
I64 – not specified stroke	1 (0.5)
First ever stroke	151 (80.3)
Prior stroke	37 (19.7)
Lengths of stay at hospital, days	
≤ 5	42 (22.3)
6–10	53 (28.2)
11–30	53 (28.2)
≥ 31	40 (21.3)

ICD: International Classification of Diagnoses.

Table II. *Results of item scaling tests and reliability estimates*

Scale	K*	Range of item		Item scaling tests		Scale correlation	
		Item convergent validity†	Item discriminant validity‡	Success/ total§	Scaling success (%)	Inter-reliability coefficients‖	Inter-scale correlations#
PF	10	0.61–0.85	0.03–0.49	79/80	99	0.94	0.20–0.54
RP	4	0.57–0.74	0.10–0.43	32/32	100	0.82	0.21–0.45
BP	2	0.80	0.15–0.44	16/16	100	0.87	0.19–0.43
GH	5	0.31–0.64	0.01–0.60	30/40	75	0.72	0.23–0.57
VT	4	0.44–0.59	0.15–0.58	24/32	75	0.74	0.34–0.59
SF	2	0.63	0.15–0.46	16/16	100	0.77	0.19–0.44
RE	3	0.69–0.80	0.14–0.49	24/24	100	0.87	0.23–0.51
MH	5	0.57–0.74	0.17–0.54	39/40	97.5	0.84	0.30–0.59

*Number of items and number of item convergent validity tests per scale.
†Correlation between items and expected scale, corrected for overlap.
‡Correlations between items and other scales.
§Number of significantly correlations to expected scale / total number of correlations.
‖Cronbach's alpha.
#Pearson's correlation coefficient.
PF: Physical Functioning; RP: Role Physical; BP: Bodily Pain; GH: General Health; VT: Vitality; SF: Social Functioning; RE: Role Emotional; MH: Mental Health.

Table II shows the results of our evaluation of item discriminate validity and scaling success rates. There were significant item-scale correlations (corrected for overlap) in 4 of the 8 scales, with a scaling success of up to 100%. The scaling success rating was 99% for the PF scale, 97.5% for the MH scale, and 75% for the GH and VT scales.

The internal consistency reliability for the 8 scales was greater than 0.70 (Cronbach's alpha). For 5 scales, Cronbach's alpha was greater than 0.80, and for the PF scale Cronbach's alpha was greater than 0.90. The scale intercorrelations were lower than the respective alpha values ($r=0.19$–0.57) (Table II).

For 7 of 8 scales, scores varied between 0 and 100 (Table III). There was positive skewness (+1.91) for the score distribution of the RP scale and negative skewness (–1.00) for the MH scale; thus, respondents' answers were towards poorer function respectively to better function. Skewness for the other 6 scales ranged between –0.59 and –0.04. Four scales had no floor and ceiling effects. Two scales had floor effects (RP, 70.8%; RE, 33.5%) and 2 scales had ceiling effects (RE, 42.2%; BP, 36.8%).

Factor analysis showed that 2 factors explained 57% of the variance: the MCS explained 45% and the PCS explained 12% (Table IV). As expected, in the SF-36 model, the PF, GH, and RP scales loaded on the PCS, while the BP scale loaded on the MCS instead of the PCS. Also as expected, the RE, MH, VT, and SF scales loaded on the MCS.

Comparisons by gender showed that males had significantly higher HRQoL in PF ($p=0.004$) and VT ($p=0.013$) than females. The stroke group's lowest mean score was for the RP scale (15.5 scores) and the highest mean score was for the MH scale (75.4 scores). The calculation of effect sizes for the whole stroke group showed large effect for 5 scales (PF, RP, SF, RE, PCS), medium effect for the VT scale, small effect for 3 scales (GH, MH, MCS) and no effect for the BP scale. The results of the middle age group (45–64 years) showed large effect for 4 scales, medium effect for 4 scales and small effect for 2 scales (Table V and Fig. 1).

The results showed moderate correlations between BI and the PF ($r=0.63$, $p=0.01$) and the PSC ($r=0.49$, $p=0.01$) and for the remaining scales the correlations were weak. There is weak correlation between educational level and the PF ($r=0.15$, $p=0.05$).

Table III. *Descriptive statistics and features of score distributions for short form-36 (SF-36)*

	PF	RP	BP	GH	VT	SF	RE	MH
Items	10	4	2	5	4	2	3	5
Levels	21	5	11	21	21	9	4	26
Mean	49.5	15.5	65.7	58.5	52.8	66.2	53.6	75.4
95% CI	45.0–54.1	11.3–19.8	61.1–69.5	55.4–61.6	49.4–56.1	62.1–70.4	47.2–60.0	72.4–78.4
Median	50.0	0.0	62.0	62.0	55.0	75.0	66.7	80.0
Range	0–100	0–100	0–100	5–100	0–100	0–100	0–100	0–100
SD	31.4	29.2	31.4	20.9	22.9	28.7	44.2	20.5
% floor†	6.5	70.8	1.6	0.5	2.2	4.3	33.5	0.5
% ceiling‡	4.9	6.5	36.8	1.1	1.6	22.2	42.2	9.2
Skewness	–0.04	1.91	–0.26	–0.23	–0.30	–0.59	–0.12	–1.00

†Percentage of subjects with worst possible score.
‡Percentage of subjects with best possible score.
PF: Physical Functioning; RP: Role Physical; BP: Bodily Pain; GH: General Health; VT: Vitality; SF: Social Functioning; RE: Role Emotional; MH: Mental Health; 95% CI: 95% confidence interval.

Table IV. *Scale validity and correlations with rotated principal components*

	Rotated principal components	
	Mental†	Physical†
Physical Functioning	0.33	0.89
Role Physical	0.40	0.65
Bodily Pain	0.58	0.35
General Health	0.55	0.78
Vitality	0.74	0.63
Social Functioning	0.60	0.43
Role Emotional	0.76	0.25
Mental Health	0.85	0.40
Eigenvalue	3.58	0.95
Variance, %	45	12

†Correlation between each short form-36 (SF-36) scale and rotated (oblimin) principal component. Strong association, $r \geq 0.70$; moderate to substantial association, $0.30 < r > 0.70$; weak association, $r \leq 0.30$.

DISCUSSION

SF-36 is widely used to measure HRQoL. Its 8 scales and 2 summary components have demonstrated good psychometric properties in studies of general populations and specific illnesses and disabilities. The present study in patients after stroke confirms to a great extent scale assumptions of SF-36, but demonstrates some differences. Our findings also show that stroke has a marked impact on quality of life, especially in patients aged 45–64 years.

Convergent validity was good for 94% of the items, whereas 2 items, "Overall rating of general health" (1) and "I feel worn out" (9g), had stronger correlations to other scales than to their own scale. This could indicate that the perceptions of physical impairment in patients after stroke have a stronger correlation to their overall ratings of general health. In the same way the perceptions of mental health are associated with the patients' feelings of being worn out. It is possible that these 2 items did not correlate with the intended scales due to the frequent symptoms of fatigue and physical disabilities, particularly so soon after discharge. Further research is needed to evaluate item-scale correlations for stroke groups at different intervals following stroke onset.
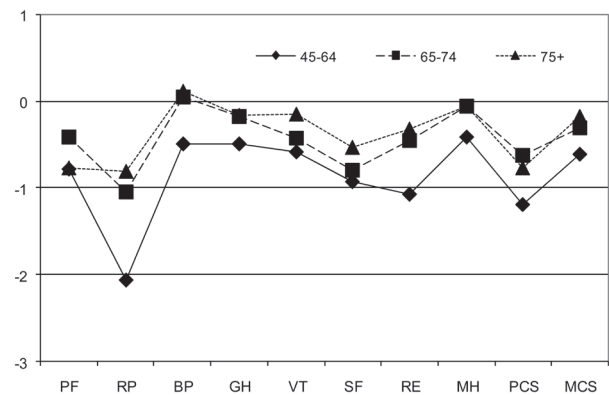
Effect size (d)

Fig. 1. Effect size of mean scores for patients after stroke in Sweden aged 45–64, 65–74 and 75 years and older from the Swedish population normative mean scores. PF: Physical Functioning; RP: Role Physical; BP: Bodily Pain; GH: General Health; VT: Vitality; SF: Social Functioning; RE: Role Emotional; MH: Mental Health; PCS: physical component summary; MCS: memtal component summary.

Scaling success ratings were high except for 2 scales, GH and VT, which had 75% scaling success. These findings are partly consistent with a previous study of patients after stroke (7), which demonstrated lesser scaling success for the SF (43%) and GH (29%) scales. The more limited scaling success does not support item discriminate validity for these scales (GH and VT), because all items in the scale did not have significant correlations to the own scale. If the item-own-scale correlation is more than 2 standard errors distant from the item-other-scale correlation, it is considered significant. This means that each item measures only one health concept (2, 15). The non-significant correlations could be due to the small sample size, as the standard error would be smaller with a larger sample (15).

The internal consistency was good for 3 scales (> 0.70) and very good or excellent for 5 scales (> 0.80). These internal consistency ratings are better than those reported in most other studies of patients after stroke, where correlations lower than 0.70 have been reported for the GH (7, 8), SF (7), and VT (6,

Table V. *Effect sizes (ES) for the stroke groups compared with the normative Swedish population groups*

	Age groups, years							
	45–64		65–74		≤75		Total group	
SF-36	ES (d)	CI	ES (d)	CI	ES (d)	CI	ES (d)	CI
PF	−0.78	−0.45 to −1.46	−0.42	−0.13 to −0.71	−0.77	−0.52 to −1.02	−1.28	−1.13 to −1.43
RP	−2.07	−1.73 to −2.40	−1.05	−0.75 to −1.34	−0.90	−0.64 to −1.15	−1.61	−1.46 to −1.76
BP	−0.50	−0.17 to −0.83	0.05	0.34 to −0.24	0.12	0.36 to −0.12	−0.15	0.0 to −0.30
GH	−0.49	−0.16 to −0.82	−0.18	0.12 to −0.47	−0.16	0.08 to −0.40	−0.49	−0.35 to −0.64
VT	−0.58	−0.25 to −0.91	−0.42	−0.13 to −0.71	−0.16	0.08 to −0.40	−0.62	−0.47 to −0.76
SF	−0.93	−0.60 to −1.26	−0.79	−0.50 to −1.08	−0.51	−0.26 to −0.75	−0.95	−0.81 to −1.10
RE	−1.07	−0.74 to −1.40	−0.45	−0.16 to −0.74	−0.31	−0.06 to −0.55	−0.88	−0.73 to −1.02
MH	−0.42	−0.09 to −0.75	−0.05	0.24 to −0.34	−0.06	0.19 to −0.30	−0.26	−0.12 to −0.41
PCS	−1.19	−0.85 to −1.52	−0.63	−0.34 to −0.92	−0.80	−0.54 to −1.06	−1.29	−1.14 to −1.44
MCS	−0.60	−0.27 to −0.93	−0.30	−0.01 to −0.59	−0.16	0.10 to −0.42	−0.38	−0.24 to −0.53

SF-36: short form-36; CI: confidence interval for effect size; PF: Physical Functioning; RP: Role Physical; BP: Bodily Pain; GH: General Health; VT: Vitality; SF: Social Functioning; RE: Role Emotional; MH: Mental Health; PCS: physical component summary; MCS: memtal component summary.

8) scales. We found that the other-scales correlation for each scale was lower than the reliability coefficient, supporting the idea that each scale measures a distinct concept.

Floor and ceiling effects are a problem in health-related measurements; they limit the usefulness of a scale for evaluating outcomes (16). There was a considerable floor effect for the RP and RE scales and a considerable ceiling effect for the BP and RE scales; this was also reported for previous studies of patients after stroke (7, 8, 10). In generally well-functioning populations, notable floor effects do not exist, while ceiling effects are more common (2, 5). The ceiling and floor effects in the RP and RE scales may relate to the fact that these scales have only 4 or 5 levels, making fine discriminations between individuals difficult. Another possible reason for the floor effects is that measurements were performed 2–3 weeks after discharge; therefore, the patients had limited opportunity to participate in their usual daily activities.

For 7 scales, the hypothesized association to each dimension was fulfilled, with moderate or strong associations. The BP scale had divergent loading and was moderately associated with MCS instead of PCS. This probably indicates that patients after stroke consider bodily pain more of a mental characteristic. The MCS explained most of the variance in the SF-36 scales, which is opposite to the results from Hobart et al. (7). Our findings demonstrate shortcomings in summarizing to PCS and MCS, which are in line with earlier studies among stroke survivors (7) and in general groups (17).

We found stroke to have a negative impact on HRQoL, especially in the middle-aged group. In the 2 other age groups, for the 2 physical scales PF and RP and for the PCS scale, HRQoL was poorer than for the normal population. Scores for the mental scales VT, SF, and RE were also lower than for the same age groups in the Swedish normal population. These findings are similar to earlier research comparing patients after stroke to a normal population in the UK (18). It is well known that stroke frequently causes impairment of physical, cognitive, emotional, behavioural and social functions, which affect stroke survivors' performance of daily activities, and consequently, their traditional roles and habits. Previous research has reported that limitations on daily activities (19), leisure activities (20) and community reintegration (19) have a negative effect on the HRQoL of stroke survivors. Psychosocial factors seem to be more essential than physical abilities in this respect (21). A possible explanation for our findings of low HRQoL in role and functioning scales is that this study was performed shortly after discharge, allowing only a short time for recovery of roles and habits.

Our findings support physical and mental health as latent factors in HRQoL among patients after stroke (7). The results also suggest that the impact of stroke on mental HRQoL is greater for patients aged 45–64 years than for older patients after stroke. This may be due to the fact that stroke occurs suddenly and unexpectedly, and middle-aged patients have not learned coping strategies for their new life situation and daily activities so soon after discharge. Older patients are more likely to have had previous illness, which might have required development of coping strategies. Another possible explanation

is that elderly anticipate health problems to a greater extent and regard them as a normal part of life and common among their peers. Older patients suffer more often from co-morbidity, which might diminish the relative impact of stroke.

Our findings of low HRQoL for the MH and VT scales could be explained by the fact that depressive symptoms and fatigue are frequent among patients after stroke. This is in accordance with other studies among patients after stroke, which have also reported that depression (22), fatigue (22) and social support (23) are predictors of HRQoL.

A strength of this study is that patients were enrolled consecutively from the same stroke unit during a 2-year period; consequently, there was no known selection bias. The patients were assessed between 2 weeks and 5 months post-stroke, which may affect the results, although around 80% were interviewed within 6 weeks after the stroke. Those with a longer hospital stay probably have more functional problems, which might influence the HRQoL. A further strength is the paucity of external and internal missing data. A limitation is the small sample size, which influences the size of the standard error and scaling success ratings. A strength of the study is that the patients were interviewed, as recommended by O'Mahoney et al. (10), who did not advocate sending postal questionnaires to elderly patients after stroke because of the risk that all questions due to cognitive dysfunction and other impairments might not be answered correctly or not at all. This is different from the study of the normative study population, which was conducted with self-reported questionnaires. As our study does not include stroke survivors with aphasia and/or cognitive impairment, it is also not possible to generalize our findings to all patients after stroke.

In conclusion, our results show that SF-36 functions well as a measure of HRQoL in Swedish patients after stroke. The internal consistency reliability was high for all scales. The scaling success ratings were high for 6 scales and somewhat lower for 2 scales; this supports item discriminant validity. Three scales exhibited substantial floor and/or ceiling effects, which is to be expected for a group with serious health problems. The hypothesized association of 7 scales with the 2 summary scales was confirmed. Compared with a normal Swedish population, stroke has a negative effect on HRQoL, especially for patients aged 45–54 years. Our findings indicate that the SF-36 has good divergent validity for discriminating the HRQoL of patients after stroke and normal populations.

ACKNOWLEDGEMENTS

REFERENCES

1. Socialstyrelsen. Nationella riktlinjer för stroke sjukvård 2005 [National Guidelines for Stroke Care 2005]. Stockholm: National Board of Health and Welfare; 2006 (in Swedish).
2. Sullivan M, Karlsson J, Taft C. Hälsoenkät: Svensk Manual och

Tolkningsguide, 2:a upplagan [Swedish interpretation guide, 2nd edition]. Gothenburg: Sahlgrenska University Hospital; 2002 (in Swedish).

3. Ware JE, CD Sherbourne. The MOS 36-item short-form health survey (SF-36). I conceptual framework and item selection. Med Care 1992; 30: 473–483.

4. Hays RD, Hahn H, Marshall G. Use of the SF-36 and other health-related quality of life measures to assess persons with disabilities. Arch Phys Med Rehabil 2002; 83: 4–9.

5. Sullivan M, Karlsson J, Ware JEJ. The Swedish SF-36 Health Survey – I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden. Soc Sci Med 1995; 41: 1349–1358.

6. Anderson C, Laubscher S, Burns R. Validation of the Short Form 36 (SF-36) health survey questionnaire among stroke patients. Stroke 1996; 27: 1812–1816.

7. Hobart JC, Williams LS, Moran K, Thompson AJ. Quality of life measurement after stroke: uses and abuses of the SF-36. Stroke 2002; 33: 1348–1356.

8. Hagen S, Bugge C, Alexander H. Psychometric properties of the SF-36 in the early post-stroke phase. J Adv Nurs 2003; 44: 461–468.

9. Dorman P, Slattery J, Farrell B, Dennis M, Sandercock P. Qualitative comparison of the reliability of health status assessments with the EuroQol and SF-36 questionnaires after stroke. Stroke 1998; 29: 63–68.

10. O'Mahony PG, Rodgers H, Thomson RG, Dobson R, James OFW. Is the SF-36 suitable for assessing health status of older stroke patients? Age Ageing 1998; 27: 19–22.

11. Folstein MF, Folstein SE, McHugh PR. Mini-mental state. A practical method for grading the cognitive state of patients for clinician. J Psychiatr Res 1975; 12: 189–198.

12. Steiner DL, Norman GR, editors. Health measurement scales a practical guide to their development and use. Oxford: Oxford University Press; 2003.

13. McHorney CA, Ware JEJ, Lu JFR, Sherbourne CD. The MOS 36-ltem Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care 1994; 32: 40–66.

14. Cohen J. Statistical power analysis for the behavioral sciences (2nd edn). Hillsdale, NJ: Lawrence Earlbaum Associates; 1988.

15. Ware Jr JE, Gandek B. Methods for testing data quality, scaling assumptions, and reliability: The IQOLA project approach. J Clin Epidemiol 1998; 51: 945–952.

16. Kazdin A, editor. Research design in clinical psychology. Boston: Allyn and Bacon Boston; 2003.

17. Taft C, Karlsson J, Sullivan M. Do SF-36 summary component scores accurately summarize subscale scores? Qual Life Res 2001; 10: 395–404.

18. Bugge C, Hagen S, Alexander H. Measuring stroke patients' health status in the early post-stroke phase using the SF36. Int J Nurs Stud 2001; 38: 319–327.

19. Mayo NE, Wood-Dauphinee S, Robert C, Durcan L, Carlton J. Activity, participation, and quality of life 6 months poststroke. Arch Phys Med Rehabil 2002; 83: 1035–1042.

20. Sveen U, Thommessen B, Bautz-Holter E, Wyller TB, Laake K. Well-being and instrumental activities of daily living after stroke. Clin Rehab 2004; 18: 267–274.

21. Kim P, Warren S, Madill H, Hadley M. Quality of life of stroke survivors. Qual Life Res 1999; 8: 293–301.

22. Naess H, Waje-Andreassen U, Thomassen L, Nyland H, Myhr K-M. Health-related quality of life among young adults with ischemic stroke on long-term follow-up. Stroke 2006; 37: 1232–1236.

23. Jaracz K, Kozubski W. Quality of life in stroke patients. Act Neurol Scand 2003; 107: 324–329.