Assessing Disability and Quality of Life in Systemic Sclerosis: Construct Validities of the Cochin Hand Function Scale, Health Assessment Questionnaire (HAQ), Systemic Sclerosis HAQ, and Medical Outcomes Study 36-Item Short Form Health Survey

FRANÇOIS RANNOU,1 SERGE POIRAudeau,1 ALICE BEREZNÉ,1 THIERRY BAUBET,2 VÉRONIQUE LE-GUERN,1 JEAN CABANE,3 LOÏC GUILLEVIN,1 MICHEL REVEL,1 JACQUES FERMANIAN,4 AND LUC MOUTHON1

Objective. To assess the construct validity of the Cochin Hand Function Scale (CHFS) and the relevance of using aggregate scores for the scleroderma Health Assessment Questionnaire (sHAQ) and Medical Outcomes Study 36-Item Short Form Health Survey (SF-36) in systemic sclerosis (SSc).

Methods. We evaluated 50 patients with SSc (mean ± SD age and disease duration 54 ± 12 years and 9 ± 8 years, respectively), of which 26 had limited cutaneous SSc (lcSSc) and 23 diffuse SSc (dSSc). Quality of life was assessed by the SF-36, global disability by the Health Assessment Questionnaire (HAQ) and sHAQ, and hand disability by the CHFS. Construct validity was assessed by convergent and divergent validity (Spearman’s rank correlation coefficient) and factor analysis.

Results. The CHFS had good construct validity and its total score explained 75% of the variance of the HAQ. The HAQ had better construct validity than the aggregate sHAQ and their scores correlated well (r = 0.88). The aggregate sHAQ was no better than the HAQ in discriminating between lcSSc and dSSc. SF-36 physical and mental components had acceptable convergent and divergent validity. Factor analysis of the 8 subscales extracted 3 factors explaining 72% of the variance, which differed from the a priori stratification with physical and mental subscales extracted in the same factor.

Conclusion. In patients with SSc, the CHFS has good construct validity, the HAQ should be preferred over the aggregate sHAQ for assessing physical functioning, and use of SF-36 physical and mental components aggregate scores is questionable.

KEY WORDS. Systemic sclerosis; Quality of life; Outcome measures; Validity; Health assessment; Disability.

INTRODUCTION

Systemic sclerosis (SSc) is a connective tissue disease characterized by excessive collagen deposition and by vascular hyperreactivity and obliterator microvascular phenomena (1). Patients with SSc are classified according to the extent of skin involvement: limited SSc (ISSc), with no detectable skin involvement; limited cutaneous SSc...
(lcSSc) (2), with skin involvement essentially limited to the hands and face; and diffuse SSC (dSSc), with proximal skin involvement. In patients with lcSSc, visceral involvement is rare and the prognosis is good, with the exception of the 10–15% of patients who eventually develop pulmonary arterial hypertension (3). Patients with dSSc experience visceral involvement, which is responsible for reduced life expectancy (4–6). In addition to diminishing life expectancy, SSC is responsible for skin, tendon, joint, and vessel damage, which leads to disability, handicap, and worsening of quality of life (7). Therefore, outcome measures with good metric properties assessing disability and health-related quality of life are needed to assess disease evolution and treatment efficacy in SSC.

The Health Assessment Questionnaire (HAQ) has become one of the main instruments to assess disability in musculoskeletal disorders (8,9). In patients with SSC, global disability can be measured with the HAQ (7,10,11), and an acceptable sensitivity to change has been suggested (12–14). Steen and Medsger proposed the use of the scleroderma HAQ (sHAQ), a more disease-specific disability scale (12). Five patient-generated visual analog scales have been added to the original HAQ, assessing Raynaud’s phenomenon, digital tip ulcers, gastrointestinal and lung symptoms, and overall disease severity from the patient’s perspective (12). A French version of the sHAQ has been recently proposed (15). Unlike Steen and Medsger, the authors suggested the use of an aggregate score of the sHAQ scale (aggregate sHAQ). In general, measurement tools that are more disease specific are considered superior to generic instruments in assessing patients (16,17). However, to avoid multiplication of outcome measures and to standardize patients’ assessment over different countries, these specific outcome measures should demonstrate their metric superiority to widely used tools.

Because SSC frequently involves the hands (18,19), a specific tool assessing hand disability is needed. Although the Hand Mobility in Scleroderma index has recently been proposed to specifically assess hand global mobility in patients with SSC (20), it does not evaluate hand disability for activities of daily living. The Cochin Hand Function Scale (CHFS) is a functional disability questionnaire about daily activities (21). This scale has been validated in rheumatoid arthritis (RA) and hand osteoarthritis (22,23) and the reliability and concurrent validity have recently been tested in SSC (24). However, construct validity has not been fully investigated.

Several generic or specific instruments have been proposed to assess quality of life in patients with chronic illnesses (25). The Medical Outcomes Study 36-Item Short Form Health Survey (SF-36) has become one of the most widely used instruments to assess quality of life (26). This questionnaire is composed of 8 subscales that can be summarized in 2 aggregate scores: the physical component score (PCS) and the mental component score (MCS). The questionnaire has been used in systemic illnesses (27,28). In patients with SSC, use of the PCS has been shown to discriminate between lcSSc and dSSc (29). In patients with dSSc, use of the scale has demonstrated a large magnitude of responsiveness (14) and the ability to discriminate between less and more severe breathlessness in patients with lung disease (30). However, the relevance of using the 2 aggregate scores, PCS and MCS, in SSC has not been demonstrated.

Our objective was to evaluate the relevance of using 3 outcome measures (CHFS score, sHAQ aggregate score, and SF-36 MCS and PCS) in the assessment of disability and quality of life in patients with SSC. For this purpose, we assessed the construct validity of these scales by their convergent and divergent validity and factor analysis.

**PATIENTS AND METHODS**

**Study design.** A cross-sectional survey was proposed to patient members of the Association des Sclérodermiques de France, the French SSC patient association.

**Patients.** To be eligible for the study, patients had to fulfill the American College of Rheumatology (formerly the American Rheumatism Association) criteria (31) and/or the LeRoy and Medsger criteria (32) for SSC. All patients were assessed by the same operator within 48 hours (during spring 2004, temperature 20°C) during a meeting of the association. Parameters recorded were age; sex; ethnicity; occupation; national health service status; sick leave; year of onset of Raynaud’s syndrome; age at diagnosis; year of onset of the first non-Raynaud’s phenomenon; disease duration; disease form (ISSc, lcSSc, or dSSc); body mass index; Karnovski index score; dyspnea (assessed by the New York Heart Association 4-point scale); pitting scars; digital ulcers; calcinosis; esophagus, joint, and/or muscle involvement; heart involvement and interstitial lung disease; pulmonary arterial hypertension; renal crisis; physical therapy; and use of hand orthoses. Evidence of esophagus, joint, and/or muscle involvement, heart involvement, interstitial lung disease, pulmonary arterial hypertension, and scleroderma renal crisis was based on patients’ oral reports.

**Quality of life assessment.** The French version of the SF-36 (26) is a self-administered questionnaire covering 8 areas: physical function, physical role, bodily pain, general health, vitality, social function, emotional role, and mental health. For each area, the score ranges from 0 (poorer health status) to 100 (better health status). Scores can also be summarized in 2 global scores: the PCS and MCS.

**Disability assessment.** Global disability was assessed using the HAQ (8) and sHAQ (12), with the scale ranging from 0 (no disability) to 3 (maximal disability). The HAQ and sHAQ both comprise 20 items divided into 8 domains, and the sHAQ has 5 additional domains (scored on a visual analog scale [VAS]) assessing disability induced by SSC symptoms. The aggregate score of the sHAQ is calculated as follows: the 5 VAS scores are converted to a continuous scale from 0 to 3, and the aggregate score is obtained by adding the 8 HAQ domain scores with the 5 VAS converted scores and the sum is divided by 13 (15).

Hand disability was evaluated using the CHFS (21), a
questionnaire with 18 items concerning daily activities. Each question is scored on a scale from 0 (performed without difficulty) to 5 (impossible to do), which is administered by the physician. The total score was obtained by adding the scores from all items (range 0–90).

Global hand and wrist mobility assessment. Global hand and wrist mobility were evaluated using the hand functional index (HFI; first 9 questions of the Keitel functional index developed by Keitel et al [33]) and Kapandji index (34–36). The HFI score ranges from 4 (best mobility) to 42 (worst mobility) and the Kapandji score ranges from 0 (worst mobility) to 100 (best mobility).

Patients’ perceived disability. Patients’ perceived disability was assessed with the McMaster Toronto Arthritis Patient Preference Disability Questionnaire (MACTAR) (37). Patients were asked to select the 3 situations among activities of daily living that caused them maximal trouble. Each item is scored on an 11-point semiquantitative scale (range 0–10). The global score ranges from 0 (no disability) to 30 (maximal disability).

Anxiety and depression assessments. The hospital anxiety and depression scale (HAD) was used to assess anxiety (HADa) and depression (HADd) (38). This scale has 7 questions for each dimension and ranges from 0 (no depression, no anxiety) to 21 (maximal depression, maximal anxiety). We used the MACTAR and HAD to better define the convergent and divergent validity of the CHFS, sHAQ, and SF-36 PCS and MCS (see below in statistical analysis); we did not use the MACTAR and HAD to assess the construct validity of these questionnaires.

Statistical analysis. For each questionnaire (CHFS, HAQ, sHAQ, and SF-36 PCS and MCS), construct validity was investigated in 3 ways. Convergent validity was assessed by correlating the questionnaire scores with scores on variables supposedly assessing similar dimensions or concepts. Divergent validity was assessed by correlating the scores with scores on variables known to assess dimensions or concepts differing from those assessed by the questionnaire tested. Therefore, for example, we hypothesized that HAQ, sHAQ, MACTAR, HFI, Kapandji, and CHFS scores would correlate better with the SF-36 PCS score than HADa, HADd, and SF-36 MCS scores. Because a normal distribution could not be demonstrated for all quantitative variables, Spearman’s rank coefficient was used to assess the correlation between 2 quantitative variables. Spearman’s coefficient values were interpreted as excellent (>0.91), good (0.90–0.71), moderate (0.70–0.51), fair (0.50–0.31), or little or none (<0.30) (39). Principal component analysis was used to extract factors. This analysis was performed for the items of the CHFS, HAQ, and sHAQ and the 8 subscales of the SF-36. Retained factors had eigenvalues >1. Eigenvalues were obtained by matrix algebra and represent the part of the entire variation of the data that can be attributed to each factor. Then, independent factors were obtained using the varimax rotation method, an orthogonal rotation method applied to the initial factorial solution, in an attempt to minimize the number of variables that have high loading in each factor. Finally, analysis of variance was used to assess the weight of hand disability assessed by the CHFS on global disability assessed by the HAD. A P value less than 0.05 was considered significant.

RESULTS

Demographic and clinical data. Eighty patients were asked to participate, and 50 (44 women and 6 men) accepted. All patients were white. Mean ± SD age at the time of evaluation was 54 ± 12 years, and mean ± SD disease duration was 9 ± 8 years. Twelve (24%) patients were currently working, 19 (38%) were retired, 6 (12%) were officially disabled, and 10 (20%) had stopped working for >3 years because of SSc. Twenty-three patients (46%) had dSSc and 26 (52%) had lcSSc (Table 1).

Outcome measure scores. Patients’ hand global mobility was reduced, with mean ± SD HAQ and Kapandji scores of 15.08 ± 10.22 and 78.75 ± 19.94, respectively (Table 2). The mean ± SD CHFS score was 16.56 ± 16.40 and explained 75% of the variance in HAQ global score. Mean ± SD global disability scores assessed by the HAQ and sHAQ were 1.07 ± 0.68 and 0.96 ± 0.50, respectively. Mean ± SD SF-36 PCS and MCS scores were 43.75 ± 21.23 and 50.74 ± 18.82, respectively. The mean ± SD perceived handicap score was 14.96 ± 5.61. A higher mean anxiety score than depression score was observed, with values of 10.29 ± 4.30 and 6.17 ± 3.68, respectively. Significant

| Table 1. Demographic and clinical characteristics of 50 patients with SSc* |
|----------------|----------------|
| Characteristic | Value |
| Age at evaluation, mean ± SD years | 54 ± 12 |
| Female sex | 44 (88) |
| Age at disease onset, mean ± SD years | 46 ± 12 |
| Disease duration at evaluation, mean ± SD years | 9.1 ± 8.8 |
| Skin involvement | 49 (98) |
| Limited SSc | 1 (2) |
| Limited cutaneous SSc | 26 (52) |
| Diffuse SSc | 23 (46) |
| Karnofski index score, mean ± SD (%) | 73 ± 9.5 (60–90) |

* Values are the number (percentage) unless otherwise indicated. SSc = systemic sclerosis; NYHA = New York Heart Association.
between the HAQ and aggregate $s$HAQ scores was high
MCS, and depression (HADd) (Table 3). The correlation
and age; and no correlation with anxiety (HADa), SF-36
mobility (HFI and Kapandji indexes), disease duration,
function and PCS); a weak correlation with global hand
handicap (MACTAR) and disability (SF-36 for physical

differences were observed between patients with lcSSc
and dSSc for HFI, Kapandji, CHFS, and HAQ scores (Table

Construct validity of the CHFS. The CHFS scale had
good convergent validity with global disability (HAQ and
$\text{s}$HAQ) and global hand disability (HFI and Kapandji
indexes), a weaker correlation with patient’s perceived
handicap (MACTAR) and disability (SF-36 for physical
function and PCS), and no correlation with anxiety
(HADa), SF-36 MCS, depression (HADD), disease duration,
and age (Table 3). Factor analysis (Table 4) extracted 2
factors that accounted for 71.63% of the total variance. The
first factor represents mainly activities requiring grip and
pinch strength and the second represents activities requiring
pinch dexterity. The loading of each question after vari-
max rotation is shown in Table 5.

Construct validity of the HAQ and aggregate $s$HAQ.
The HAQ and $\text{s}$HAQ had similar convergent and divergent
validity and good correlation with hand disability (CHFS);
a weaker but fair correlation with patients’ perceived
handicap (MACTAR) and disability (SF-36 for physical
function and PCS); a weak correlation with global hand
mobility (HFI and Kapandji indexes), disease duration,
and age; and no correlation with anxiety (HADa), SF-36
MCS, and depression (HADD) (Table 3). The correlation
between the HAQ and aggregate $s$HAQ scores was high
($r = 0.88$).

For HAQ, factor analysis (Table 4) extracted 1 factor that
accounted for 63.1% of the total variance. For the aggrega-
tate $s$HAQ (Table 4), factor analysis extracted 3 factors that
accounted for 68.8% of the total variance. These factors
differed from the a priori stratification. The first factor
comprised the 8 items of the original HAQ, the second
comprised 4 of the 5 patient-generated items (Raynaud’s
phenomenon, gastrointestinal and lung symptoms, and
overall disease severity from the patient’s perspective),
and the third comprised digital tip ulcers. The loading of
each question of the HAQ and aggregate $s$HAQ after vari-
max rotation is shown in Table 5.

Construct validity of the SF-36. Acceptable convergent
and divergent validities were observed for the PCS and
MCS scores of the SF-36, but correlation between the 2
scores was higher than expected ($r = 0.57$). PCS was fairly
correlated with global disability (HAQ and $s$HAQ); weakly
 correlated with patients’ perceived handicap (MACTAR),
hand disability (CHFS), and depression (HADD); and not
correlated with age, disease duration, and anxiety (HADa)
(Table 3). MCS was fairly correlated with depression and
anxiety and not correlated with global and hand disability,
disease duration, handicap, and age. Scores from 3 sub-
scapes of the SF-36 (social functioning, bodily pain, and
general health perception) were correlated equally with
PCS and MCS (Table 6).

Factor analysis of the 8 subscales (Table 4) extracted 3
factors that accounted for 71.6% of the total variance. These
factors differed from the a priori stratification in that
subscales assessing physical and mental components were
in the same factor. With the a priori stratification, results
are more often presented with the 2 factors PCS and MCS,
PCS being the mean of physical functioning, physical role,
bodily pain, and general health perception scores and

<table>
<thead>
<tr>
<th>Scores</th>
<th>Whole group (n = 50)</th>
<th>lcSSc group (n = 23)</th>
<th>dSSc group (n = 27)</th>
<th>$P^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HFI (range 4–42)</td>
<td>15.08 ± 10.22 (4–40)</td>
<td>11.57 ± 8.21 (4–22)</td>
<td>19.35 ± 11.95 (4–40)</td>
<td>0.007</td>
</tr>
<tr>
<td>Kapandji (range 0–10)</td>
<td>78.75 ± 19.94 (22–100)</td>
<td>65.69 ± 15.50 (33–100)</td>
<td>70.00 ± 20.70 (22–100)</td>
<td>0.004</td>
</tr>
<tr>
<td>CHFS (range 0–90)</td>
<td>16.56 ± 16.40 (0–87)</td>
<td>11.07 ± 11.04 (0–41)</td>
<td>23.48 ± 19.45 (0–87)</td>
<td>0.01</td>
</tr>
<tr>
<td>HAQ (range 0–3)</td>
<td>1.07 ± 0.68 (0–3)</td>
<td>0.90 ± 0.57 (0–1.75)</td>
<td>1.28 ± 0.75 (0–3)</td>
<td>0.05</td>
</tr>
<tr>
<td>$s$HAQ (range 0–3)</td>
<td>0.96 ± 0.50 (0.08–2.08)</td>
<td>0.85 ± 0.48 (0.08–1.77)</td>
<td>1.10 ± 0.49 (0.38–2.08)</td>
<td>0.07</td>
</tr>
<tr>
<td>SF-36 (range 0–100)</td>
<td>43.75 ± 21.23 (10–81)</td>
<td>43.81 ± 19.68 (12–82)</td>
<td>43.67 ± 23.44 (10–81)</td>
<td>0.981</td>
</tr>
<tr>
<td>PCS</td>
<td>50.74 ± 18.82 (21–82)</td>
<td>54.14 ± 18.59 (21–82)</td>
<td>47.94 ± 18.87 (26–79)</td>
<td>0.245</td>
</tr>
<tr>
<td>Physical functioning</td>
<td>51.97 ± 24.55 (0–90)</td>
<td>47.61 ± 25.49 (0–90)</td>
<td>55.43 ± 23.09 (0–85)</td>
<td>0.258</td>
</tr>
<tr>
<td>Physical role</td>
<td>32.35 ± 34.38 (0–100)</td>
<td>40.21 ± 41.79 (0–75)</td>
<td>25.89 ± 25.89 (0–100)</td>
<td>0.161</td>
</tr>
<tr>
<td>Emotional role</td>
<td>48.69 ± 44.83 (0–100)</td>
<td>56.52 ± 46.53 (0–100)</td>
<td>42.26 ± 43.15 (0–100)</td>
<td>0.266</td>
</tr>
<tr>
<td>Social functioning</td>
<td>59.86 ± 19.70 (25–100)</td>
<td>58.70 ± 20.01 (25–100)</td>
<td>60.78 ± 19.69 (25–100)</td>
<td>0.710</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>47.27 ± 23.95 (0–100)</td>
<td>43.17 ± 24.31 (10–100)</td>
<td>50.52 ± 23.57 (0–100)</td>
<td>0.279</td>
</tr>
<tr>
<td>Mental health</td>
<td>55.72 ± 18.26 (8–92)</td>
<td>60.70 ± 18.47 (8–88)</td>
<td>51.86 ± 17.42 (32–92)</td>
<td>0.086</td>
</tr>
<tr>
<td>Vitality</td>
<td>39.42 ± 15.42 (15–80)</td>
<td>40.65 ± 18.23 (15–65)</td>
<td>38.45 ± 13.03 (15–80)</td>
<td>0.628</td>
</tr>
<tr>
<td>General health perception</td>
<td>38.03 ± 16.60 (0–82)</td>
<td>39.70 ± 18.42 (10–82)</td>
<td>36.66 ± 15.14 (0–72)</td>
<td>0.530</td>
</tr>
<tr>
<td>Perceived individualized handi</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(MACTAR) (range 0–50)</td>
<td>14.96 ± 5.61 (5–30)</td>
<td>14.11 ± 5.26 (7–23)</td>
<td>15.96 ± 5.97 (5–30)</td>
<td>0.260</td>
</tr>
<tr>
<td>Anxiety (HADa) (range 0–21)</td>
<td>10.29 ± 4.30 (3–18)</td>
<td>6.59 ± 3.09 (4–18)</td>
<td>5.65 ± 4.33 (3–17)</td>
<td>0.388</td>
</tr>
<tr>
<td>Depression (HADD) (range 0–21)</td>
<td>6.17 ± 6.38 (1–21)</td>
<td>10.97 ± 4.22 (1–12)</td>
<td>9.43 ± 4.34 (1–21)</td>
<td>0.207</td>
</tr>
</tbody>
</table>

* Values are the mean ± SD (range) unless otherwise indicated. SSc = systemic sclerosis; lcSSc = limited cutaneous SSc; dSSc = diffuse SSc; HFI = hand functional index (first 9 items of the Keitel index); CHFS = Cochin Hand Function Scale; HAQ = Health Assessment Questionnaire; $s$HAQ = Health Assessment Questionnaire for scleroderma; SF-36 = Medical Outcomes Study 36-Item Short Form Health Survey; PCS = physical component score; MCS = mental component score; MACTAR = McMaster Toronto Arthritis Patient Preference Disability Questionnaire; HADa = hospital anxiety and depression scale for anxiety; HADD = hospital anxiety and depression scale for depression.

† lcSSc versus dSSc ($P \leq 0.05$).
MCS the mean of mental health, vitality, emotional role, and social functioning scores. The first factor comprised 3 subscales: physical and emotional roles and social functioning. The second factor comprised 2 subscales: mental health and vitality. The third factor was physical functioning. Bodily pain was shared between the 3 factors, and as expected, general health perception was shared between factors 2 and 3. The loading of each subscale and item after varimax rotation is depicted in Table 5.

**DISCUSSION**

Our results provide evidence that the CHFS has a good construct validity in SSc and that hand involvement has a significant impact on global disability. CHFS scores contribute to 75% of the HAQ variance, highlighting the need to specifically assess hand disability in patients with SSc when evaluating treatment. The mean CHFS score for the entire group of patients with SSc was equal to that previously reported for patients with RA (21), and the mean score for patients with dSSc was higher. This result suggests that the disability caused by hand involvement in the patients with SSc who had no Raynaud’s phenomenon at the time of evaluation was as severe as that observed in a population of patients with RA with a comparable disease duration of 10 years (22). Finally, the CHFS score significantly differed between patients with lcSSc and dSSc, which suggests that patients with severe disease have more severe hand disability.

Taken together, our convergent and divergent validity results in a group of 50 patients with SSC and those of other authors, showing significant correlations between CHFS, HAQ, and HFI scores in a group of 40 patients with SSC (24), reflect the good construct validity of the CHFS in SSC as was observed in previous studies of patients with RA (21,22). This impression is reinforced by our results of factor analysis, which extracted 2 easily characterized factors, except for 1 item (item 16), which was shared be-
between both factors. However, these factor analysis results of the CHFS should be interpreted with caution given the small sample size and the number of items. It is not surprising that these factors differ from those extracted in RA and hand osteoarthritis, because joint damage differs in these diseases and because other factors such as skin fibrosis and retraction contribute to hand disability in patients with SSC. For example, in RA the first factor extracted in factor analysis was activities requiring stability and mobility of the wrist (21).

Our results do not support the use of an aggregate score of the sHAQ to assess global disability in patients with SSC. Instruments that are more disease specific are thought to provide more accurate information. However, the usefulness of the HAQ in assessing patients with SSC has been extensively demonstrated (7,12,14) and no data suggest the metric pitfalls of this scale (10). Steen and Medsger proposed the sHAQ for measuring SSC status and changes in disease status (12). They demonstrated that VAS scores for digital ulcers, gastrointestinal symptoms, and lung symptoms correlated with findings for these organ systems and their changes over time (12) but did not propose the use of an aggregate score. Recently, authors of the French adaptation of the sHAQ proposed the use of an aggregate sHAQ global score to assess the multisystemic characteristics of SSC more accurately, but did not provide data suggesting a superiority of the aggregate sHAQ score over the HAQ score to assess disability (15). In the present work, factor analysis suggested a more robust factorial structure of the HAQ because of the unidimensional character of this scale. Factor analysis of the aggregate sHAQ extracted 3 factors, which differed from the a priori stratification and cannot be clinically characterized, especially for the second and third factors. We did not expect a single factor structure for the sHAQ or for the 5 added items, but it is difficult to justify why assessments of Raynaud’s phenomenon and digital tip ulcers were not grouped in the same factor and why patient overall assessment was in the second factor together with organ-specific assessment. In addition, factors extracted in our study differed from those extracted in a previous study conducted in a tertiary care setting with a higher proportion of patients with dSSc (15). Therefore, the factorial structure of the aggregate sHAQ is questionable. Moreover, because the aggregate sHAQ has 5 disease-specific items, this questionnaire could have discriminated lcSSc from dSSc better than the HAQ but did not in our study, and we and others (15) failed to demonstrate differences between both questionnaires. This find-

Table 5. Varimax rotated factor matrix of the CHFS, HAQ and sHAQ, and SF-36 with the 8 subscales*  

<table>
<thead>
<tr>
<th>CHFS questions</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.850†</td>
<td>0.209</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.727†</td>
<td>0.391</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.683†</td>
<td>0.504</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.705†</td>
<td>0.405</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.694†</td>
<td>0.322</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.715†</td>
<td>0.542</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.787†</td>
<td>0.151</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.827†</td>
<td>0.239</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.859†</td>
<td>0.138</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.856†</td>
<td>0.216</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.869†</td>
<td>0.031</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.770†</td>
<td>0.174</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.112</td>
<td>0.893†</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.060</td>
<td>0.937†</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.842†</td>
<td>0.285</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.519</td>
<td>0.537†</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.319</td>
<td>0.693†</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.821†</td>
<td>0.278</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HAQ and sHAQ domains</th>
<th>HAQ</th>
<th>sHAQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0.810</td>
<td>0.806†</td>
</tr>
<tr>
<td>2</td>
<td>0.828</td>
<td>0.827†</td>
</tr>
<tr>
<td>3</td>
<td>0.758</td>
<td>0.761†</td>
</tr>
<tr>
<td>4</td>
<td>0.802</td>
<td>0.795†</td>
</tr>
<tr>
<td>5</td>
<td>0.822</td>
<td>0.828†</td>
</tr>
<tr>
<td>6</td>
<td>0.851</td>
<td>0.852†</td>
</tr>
<tr>
<td>7</td>
<td>0.687</td>
<td>0.777†</td>
</tr>
<tr>
<td>8</td>
<td>0.782</td>
<td>0.708†</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SF-36 with 8 subscales</th>
<th>Physical functioning</th>
<th>Physical role</th>
<th>Emotional role</th>
<th>Social functioning</th>
<th>Bodily pain</th>
<th>Mental health</th>
<th>Vitality</th>
<th>General health perception</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.244</td>
<td>0.807†</td>
<td>0.876†</td>
<td>0.644†</td>
<td>0.439</td>
<td>0.259</td>
<td>0.086</td>
<td>0.037</td>
</tr>
<tr>
<td>2</td>
<td>0.385</td>
<td>0.243</td>
<td>0.318</td>
<td>0.252</td>
<td>0.439</td>
<td>0.739†</td>
<td>0.719†</td>
<td>0.674†</td>
</tr>
</tbody>
</table>

* See Table 2 for definitions.  
† Highest loading of each item (except for items that are equally correlated in several factors).
ing is explained by the occurrence of Raynaud’s phenomen-
on, digital ulcers, gastrointestinal symptoms, and lung
symptoms in both subsets of the disease. Finally, the high
correlation observed between the HAQ and aggregate
sHAQ scores (r = 0.88) suggests that the information
added by the sHAQ may not be essential when assessing
global disability in patients with SSc.
Finally, our results do not support the use of the phys-
ical and mental component aggregate scores of the SF-36
because they are currently defined to assess quality of life
in patients with SSc. Their factorial structure was not
confirmed in this study. Factor analysis extracted 3 factors
that were distinct from the a priori stratification, with
subscases supposedly belonging to the physical compo-
nents extracted in the same factor as subscases supposedly
belonging to the mental components. Moreover, bodily
pain had triple-factor content and general health percep-
tion had dual-factor content. Finally, the correlation co-
efficient between the scores of the PCS and MCS compo-
nents was higher than expected, which suggests that
several of the original 8 subscases were equally correlated
with the physical and mental components of the question-
aire. This was the case for social functioning, bodily pain,
and general health perception subscases. The factorial
complexity of the SF-36 has already been pointed out by
the analysis of 3 British surveys conducted to obtain pop-
ulation norms for adults living at home in the UK (40).
Moreover, the general health perception score has been
shown to be correlated equally with physical and mental
component scores in previous studies of French and North
American populations (41,42), and social functioning and
general health perception had a dual-factor content in a
sample of patients with osteoarthritis and RA (28). There-
fore, the relevance of providing physical and mental com-
ponent scores is debatable in the case of patients with SSc.
In the present study, of the 4 PCS components, physical
functioning was poorly correlated with bodily pain (r =
0.33), physical role (r = 0.34), and general health percep-
tion (r = 0.35). The weak correlation between physical
function and pain in SSc has been observed in other ill-
nesses such as RA and knee or hand osteoarthritis (21–
23,43).
Our results support the use of 3 distinct scores for 1) ph-
ysical and emotional roles and social functioning, 2) men-
tal components restricted to mental health and vitality,
and 3) physical component(s) restricted to physical
functioning. This stratification seems logical because for
the first factor, physical and emotional roles are concept-
tually close to social functioning and the phrasing of the 2
questions evaluating social functioning refers to physical
health and emotional problems. The second factor is also
logical because mental health is closely related to mood
status and depression scores, which in turn are linked to
vitality. For the third factor, the physical functioning sub-
scale of the SF-36 is a disability scale by itself, evaluating
limitations in activities of daily living. Finally, it is not
surprising that concepts such as pain and general health
perception, which are multifactorial, are equally shared
between factors.
We failed to demonstrate significant differences in SF-36
subscale scores between patients with lcSSc and those
with dSSc. These results differ from those reported by
others (29), who found significant differences for physical
functioning, physical role, bodily pain, and general health
perception scores. This difference could be explained by
differences in samples of patients, with a ratio of 1:15
between patients with dSSc and lcSSc in our study versus
2:9 in the other study (15).
One limitation of our study may be the procedure ap-
plied to recruit patients. Because all patients were mem-
bers of the French association of patients, they may not be
representative of the entire French SSc population. In fact,
outcome measure scores observed in our study were
slightly better than those reported in a previous study
conducted in a tertiary care setting (29). Another limita-
tion could be the relatively small number of patients
tested. No consensus exists on the minimum number of
patients needed to perform principal component analysis.
A minimum of 100–300 patients has been proposed as
necessary (44,45) or 5–10 times the number of variables
has been reported as adequate (46). However, because the
required sample size also depends on the magnitude of
real correlations and real number of factors in the popula-
tion, if there are strong correlations (i.e., >0.7) and few
distinct factors, a smaller sample size is adequate (47) and
50 patients should be sufficient (48).
Regarding our results, principal component analysis can
be applied to HAQ (1 factor, high loading of each item in
this factor), sHAQ (for each item, high loading in 1 factor
and low loading in the 2 other factors), and SF-36 (items
retained in each factor have a high loading in 1 factor and
weak loading in others, and 2 items are shared and cannot
be attributed to 1 factor). In the case of CHFS, the assign-
ment of item 16 is problematic and items 3 and 6 assigned
in factor 1 have a substantial load in factor 2 (>0.5).
In conclusion, the CHFS is a valid instrument for assess-
hand disability in patients with SSc. Hand functional
disability is the major component of global disability and
therefore should be systematically evaluated. The HAQ
seems to have better construct validity than the aggregate
sHAQ. Unless further studies demonstrate that the aggre-
gate sHAQ has better sensitivity to change than the HAQ,
the HAQ mean score should be preferred to the aggregate
sHAQ mean score to assess physical functioning in pa-
tients with SSc. PCS and MCS may not be used to sum-
marize quality of life in these patients. We propose the
individualization of 3 subscores evaluating role and social
functioning, mental components restricted to mental
health and vitality, and physical functioning, which
should be tested in further studies.

ACKNOWLEDGMENT
We wish to thank patients from the Association des Scle-
rodermiques de France for their participation in this study.

AUTHOR CONTRIBUTIONS
Dr. Mouthon had full access to all of the data in the study and
takes responsibility for the integrity of the data and the accuracy
of the data analysis.
Study design. Drs. Rannou, Poiraudou, Revel, and Mouthon.
REFERENCES


42. 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;31:247–63.


