

# 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,<sup>1</sup> SERGE POIRAUDEAU,<sup>1</sup> ALICE BEREZNE,<sup>1</sup> THIERRY BAUBET,<sup>2</sup>  
VÉRONIQUE LE-GUERN,<sup>1</sup> JEAN CABANE,<sup>3</sup> LOÏC GUILLEVIN,<sup>1</sup> MICHEL REVEL,<sup>1</sup>  
JACQUES FERMANIAN,<sup>4</sup> AND LUC MOUTHON<sup>1</sup>

**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  $\pm$  SD age and disease duration  $54 \pm 12$  years and  $9 \pm 8$  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 obliterative microvascular phe-

nomena (1). Patients with SSc are classified according to the extent of skin involvement: limited SSc (lSSc), with no detectable skin involvement; limited cutaneous SSc

Presented at the 68th Annual Scientific Meeting of the American College of Rheumatology, San Antonio, TX, October 2004.

<sup>1</sup>François Rannou, MD, PhD, Serge Poiraudau, MD, PhD, Alice Berezne, MD, Véronique Le-Guern, MD, Loïc Guillevin, MD, Michel Revel, MD, Luc Mouthon, MD, PhD: Assistance Publique-Hôpitaux de Paris, Cochin Hospital, Paris 5 University, Paris, France; <sup>2</sup>Thierry Baubet, MD: Assistance Publique-Hôpitaux de Paris and Avicenne Hospital, Paris 13 University, Bobigny, France; <sup>3</sup>Jean Cabane, MD: Assistance Publique-Hôpitaux de Paris, Saint-Antoine Hospital, Paris 6 University, Paris, France; <sup>4</sup>Jacques Fermanian, MD, PhD: Assistance Publique-Hôpitaux

de Paris, Necker Hospital, Paris 5 University, Paris, France.

Dr. Poiraudau has received consulting fees (less than \$10,000 each) from Sanofi-Aventis and Pfizer. Dr. Guillevin has received consulting fees (more than \$10,000) from Wyeth. Dr. Mouthon has received consulting fees (less than \$10,000 each) from Laboratoire Français du Fractionnement Biologique et des Biotechnologies and Pfizer.

Address correspondence to Luc Mouthon, MD, PhD, Department of Internal Medicine, Cochin Hospital, 27 Rue du Faubourg Saint-Jacques, 75679 Paris Cedex 14, France. E-mail: luc.mouthon@cch.ap-hop-paris.fr.

Submitted for publication December 1, 2005; accepted in revised form March 16, 2006.

(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 pa-

tients 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 (lSSc, 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 parameters studied, nonparametric 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

**Table 1. Demographic and clinical characteristics of 50 patients with SSc\***

Characteristic	Value
Age at evaluation, mean $\pm$ SD years	54 $\pm$ 12
Female sex	44 (88)
Age at disease onset, mean $\pm$ SD	46 $\pm$ 12
Disease duration at evaluation, mean $\pm$ SD years	9.1 $\pm$ 8.8
Skin involvement	49 (98)
Limited SSc	1 (2)
Limited cutaneous SSc	26 (52)
Diffuse SSc	23 (46)
Karnofski index score, mean $\pm$ SD (range)	73 $\pm$ 9.5 (60–90)
Raynaud's syndrome	48 (96)
Pitting scars	28 (56)
Digital ulcers	11 (22)
Gastrointestinal tract involvement	43 (86)
Arthralgias	40 (80)
Myalgias	34 (68)
Dyspnea (NYHA), mean $\pm$ SD	2.06 $\pm$ 0.68
Interstitial lung disease	20 (40)
Pulmonary arterial hypertension	5 (10)
Heart involvement	4 (8)
Renal crisis	4 (8)

\* Values are the number (percentage) unless otherwise indicated. SSc = systemic sclerosis; NYHA = New York Heart Association.

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  $\pm$  SD age at the time of evaluation was 54  $\pm$  12 years, and mean  $\pm$  SD disease duration was 9  $\pm$  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  $\pm$  SD HFI and Kapandji scores of 15.08  $\pm$  10.22 and 78.75  $\pm$  19.94, respectively (Table 2). The mean  $\pm$  SD CHFS score was 16.56  $\pm$  16.40 and explained 75% of the variance in HAQ global score. Mean  $\pm$  SD global disability scores assessed by the HAQ and sHAQ were 1.07  $\pm$  0.68 and 0.96  $\pm$  0.50, respectively. Mean  $\pm$  SD SF-36 PCS and MCS scores were 43.75  $\pm$  21.23 and 50.74  $\pm$  18.82, respectively. The mean  $\pm$  SD perceived handicap score was 14.96  $\pm$  5.61. A higher mean anxiety score than depression score was observed, with values of 10.29  $\pm$  4.30 and 6.17  $\pm$  3.68, respectively. Significant

Table 2. Outcome measure scores of patients with SSc\*

Scores	Whole group (n = 50)	lcSSc group (n = 23)	dSSc group (n = 27)	P†
HFI (range 4–42)	15.08 ± 10.22 (4–40)	11.57 ± 8.21 (4–32)	19.35 ± 11.95 (4–40)	0.007
Kapandji (range 0–10)	78.75 ± 19.94 (22–100)	85.69 ± 15.50 (33–100)	70.00 ± 20.70 (22–100)	0.004
CHFS (range 0–90)	16.56 ± 16.40 (0–87)	11.07 ± 11.04 (0–41)	23.48 ± 19.45 (0–87)	0.01
HAQ (range 0–3)	1.07 ± 0.68 (0–3)	0.90 ± 0.57 (0–1.75)	1.28 ± 0.75 (0–3)	0.05
sHAQ (range 0–3)	0.96 ± 0.50 (0.08–2.08)	0.85 ± 0.48 (0.08–1.77)	1.10 ± 0.49 (0.38–2.08)	0.07
SF-36 (range 0–100)				
PCS	43.75 ± 21.23 (10–81)	43.81 ± 19.68 (12–82)	43.67 ± 23.44 (10–81)	0.981
MCS	50.74 ± 18.82 (21–82)	54.14 ± 18.59 (21–82)	47.94 ± 18.87 (26–79)	0.245
Physical functioning	51.97 ± 24.25 (0–90)	47.61 ± 25.49 (0–90)	55.43 ± 23.09 (0–85)	0.258
Physical role	32.35 ± 34.38 (0–100)	40.21 ± 41.79 (0–75)	25.89 ± 25.89 (0–100)	0.161
Emotional role	48.69 ± 44.83 (0–100)	56.52 ± 46.53 (0–100)	42.26 ± 43.15 (0–100)	0.266
Social functioning	59.86 ± 19.70 (25–100)	58.70 ± 20.01 (25–100)	60.78 ± 19.69 (25–100)	0.710
Bodily pain	47.27 ± 23.95 (0–100)	43.17 ± 24.31 (10–100)	50.52 ± 23.57 (0–100)	0.279
Mental health	55.77 ± 18.26 (8–92)	60.70 ± 18.47 (8–88)	51.86 ± 17.42 (32–92)	0.086
Vitality	39.42 ± 15.42 (15–80)	40.65 ± 18.23 (15–65)	38.45 ± 13.03 (15–80)	0.628
General health perception	38.03 ± 16.60 (0–82)	39.70 ± 18.42 (10–82)	36.66 ± 15.14 (0–72)	0.530
Perceived individualized handicap (MACTAR) (range 0–30)	14.96 ± 5.61 (5–30)	14.11 ± 5.26 (7–23)	15.96 ± 5.97 (5–30)	0.260
Anxiety (HADa) (range 0–21)	10.29 ± 4.30 (3–18)	6.59 ± 3.09 (4–18)	5.65 ± 4.33 (3–17)	0.388
Depression (HADd) (range 0–21)	6.17 ± 3.68 (1–21)	10.97 ± 4.22 (1–12)	9.43 ± 4.34 (1–21)	0.207

\* 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; sHAQ = 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 ( $P \leq 0.05$ ).

differences were observed between patients with lcSSc and dSSc for HFI, Kapandji, CHFS, and HAQ scores (Table 2).

**Construct validity of the CHFS.** The CHFS scale had good convergent validity with global disability (HAQ and sHAQ) and global hand mobility (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 varimax rotation on the 2 factors is shown in Table 5.

**Construct validity of the HAQ and aggregate sHAQ.** The HAQ and sHAQ 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 sHAQ 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 aggregate sHAQ (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 sHAQ after varimax 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 sHAQ); 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 subscales 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 3. Convergent and divergent validities of the Cochin Hand Function Scale, HAQ and sHAQ, and PCS and MCS scores of the SF-36 for patients with SSc (correlation with other variables)\***

Scales	Spearman's correlation coefficient
Cochin Hand Function Scale	
Convergent validity	
sHAQ	0.81
HAQ	0.75
Kapandji index	0.63
HFI	0.58
SF-36 physical functioning	0.53
Perceived individualized handicap (MACTAR)	0.48
SF-36 PCS	0.45
Divergent validity	
Anxiety (HADa)	0.16
SF-36 MCS	0.14
Depression (HADd)	0.05
Disease duration	0.15
Age	0.01
HAQ	
Convergent validity	
CHFS	0.75
SF-36 physical functioning	0.58
Perceived individualized handicap (MACTAR)	0.57
SF-36 PCS	0.56
Kapandji index	0.42
HFI	0.40
Divergent validity	
Disease duration	0.32
Age	0.31
Anxiety (HADa)	0.20
SF-36 MCS	0.09
Depression (HADd)	0.04
Aggregate sHAQ	
Convergent validity	
CHFS	0.81
SF-36 physical functioning	0.64
Perceived individualized handicap (MACTAR)	0.48
SF-36 PCS	0.64
Kapandji index	0.49
HFI	0.47
Divergent validity	
Disease duration	0.33
Age	0.31
Anxiety (HADa)	0.29
SF-36 MCS	0.27
Depression (HADd)	0.09
SF-36 PCS	
Convergent validity	
HAQ	0.56
sHAQ	0.64
MACTAR	0.49
CHFS	0.45
Divergent validity	
Disease duration	0.15
Age	0.27
Depression (HADd)	0.38
SF-36 MCS	0.57
Anxiety (HADa)	0.13
SF-36 MCS	
Convergent validity	
Depression (HADd)	0.68
Anxiety (HADa)	0.63
Divergent validity	
HAQ	0.09
sHAQ	0.28
CHFS	0.14
Disease duration	0.18
Age	0.03
SF-36 PCS	0.57
MACTAR	0.06

\* See Table 2 for definitions.

**Table 4. Factor analysis of the CHFS, HAQ and sHAQ, and SF-36 with the 8 subscales\***

	Eigenvalue	Variance, %	Cumulative %
CHFS			
Factor 1	10.92	50.56	50.56
Factor 2	1.67	21.07	71.63
HAQ and sHAQ			
HAQ			
Factor 1	5.04	63.05	63.05
sHAQ			
Factor 1	5.18	39.82	39.82
Factor 2	2.70	20.80	60.62
Factor 3	1.07	8.22	68.84
SF-36 with the 8 subscales			
Factor 1	3.42	42.74	42.74
Factor 2	1.30	16.29	59.03
Factor 3	1.01	12.55	71.58

\* See Table 2 for definitions.

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-

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

	Factor 1	Factor 2	Factor 3
CHFS questions			
1	0.850†	0.209	
2	0.727†	0.391	
3	0.683†	0.504	
4	0.705†	0.405	
5	0.694†	0.322	
6	0.715†	0.542	
7	0.787†	0.151	
8	0.827†	0.239	
9	0.859†	0.138	
10	0.856†	0.216	
11	0.869†	0.031	
12	0.770†	0.174	
13	0.112	0.893†	
14	0.060	0.937†	
15	0.842†	0.285	
16	0.519	0.537†	
17	0.319	0.693†	
18	0.821†	0.278	
HAQ and sHAQ domains			
HAQ			
1	0.810		
2	0.828		
3	0.758		
4	0.802		
5	0.822		
6	0.851		
7	0.687		
8	0.782		
sHAQ			
1	0.806†	-0.019	0.081
2	0.827†	0.072	-0.220
3	0.761†	-0.016	0.046
4	0.795†	0.195	0.144
5	0.828†	0.057	0.058
6	0.847†	0.057	0.082
7	0.688†	0.130	0.225
8	0.777†	0.180	0.033
Raynaud's phenomenon	0.105	0.708†	0.503
Digital tip ulcers	0.073	0.179	0.883†
Lung symptoms	0.145	0.806†	0.154
Gastrointestinal symptoms	0.128	0.761†	0.121
Overall assessment	0.209	0.850†	0.188
SF-36 with 8 subscales			
Physical functioning	0.244	0.185	0.837†
Physical role	0.807†	0.019	0.284
Emotional role	0.876†	0.318	0.069
Social functioning	0.644†	0.252	0.457
Bodily pain	0.439	0.484	0.443
Mental health	0.259	0.739†	0.018
Vitality	0.086	0.719†	0.085
General health perception	0.037	0.674†	0.646

\* See Table 2 for definitions.  
 † Highest loading of each item (except for items that are equally correlated in several factors).

prising 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 6. Spearman's correlation coefficients of the PCS and MCS scores of the SF-36 with the 8 subscales\***

	PCS	MCS
Physical functioning	0.66	0.18
Physical role	0.83	0.47
Emotional role	0.55	0.86
Social functioning	0.67	0.72
Bodily pain	0.69	0.62
Mental health	0.22	0.75
Vitality	0.09	0.40
General health perception	0.40	0.42

\* See Table 2 for definitions.

tween 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 sur-

ing is explained by the occurrence of Raynaud's phenomenon, 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 physical 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 subscales supposedly belonging to the physical components extracted in the same factor as subscales supposedly belonging to the mental components. Moreover, bodily pain had triple-factor content and general health perception had dual-factor content. Finally, the correlation coefficient between the scores of the PCS and MCS components was higher than expected, which suggests that several of the original 8 subscales were equally correlated with the physical and mental components of the questionnaire. This was the case for social functioning, bodily pain, and general health perception subscales. The factorial complexity of the SF-36 has already been pointed out by the analysis of 3 British surveys conducted to obtain population 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). Therefore, the relevance of providing physical and mental component 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 perception ( $r = 0.35$ ). The weak correlation between physical function and pain in SSc has been observed in other illnesses such as RA and knee or hand osteoarthritis (21–23,43).

Our results support the use of 3 distinct scores for 1) physical and emotional roles and social functioning, 2) mental 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 conceptually 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 subscale 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 applied to recruit patients. Because all patients were members 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 limitation 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 population, 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 assignment 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 assessing 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 aggregate 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 patients with SSc. PCS and MCS may not be used to summarize 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 Sclérodermiques 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, Poiraudau, Revel, and Mouthon.

**Acquisition of data.** Drs. Rannou, Baubet, Le-Guern, Cabane, and Mouthon, and Ms Berezne.

**Analysis and interpretation of data.** Drs. Rannou, Poiraudeau, Guillevin, and Mouthon.

**Manuscript preparation.** Drs. Rannou, Poiraudeau, Baubet, Cabane, and Mouthon, and Ms Berezne.

**Statistical analysis.** Drs. Poiraudeau and Fermanian.

## REFERENCES

- Clements PJ, Hurwitz EL, Wong WK, Seibold JR, Mayes M, White B, et al. Skin thickness score as a predictor and correlate of outcome in systemic sclerosis: high-dose versus low-dose penicillamine trial. *Arthritis Rheum* 2000;43:2445–54.
- MacGregor AJ, Canavan R, Knight C, Denton CP, Davar J, Coghlan J, et al. Pulmonary hypertension in systemic sclerosis: risk factors for progression and consequences for survival. *Rheumatology (Oxford)* 2001;40:453–9.
- Steen VD, Medsger TA. Severe organ involvement in systemic sclerosis with diffuse scleroderma. *Arthritis Rheum* 2000;43:2437–44.
- Bouros D, Wells AU, Nicholson AG, Colby TV, Polychronopoulos V, Pantelidis P, et al. Histopathologic subsets of fibrosing alveolitis in patients with systemic sclerosis and their relationship to outcome. *Am J Respir Crit Care Med* 2002;165:1581–6.
- Steen VD, Medsger TA Jr. Long-term outcomes of scleroderma renal crisis. *Ann Intern Med* 2000;133:600–3.
- Casale R, Buonocore M, Matucci-Cerinic M. Systemic sclerosis (scleroderma): an integrated challenge in rehabilitation. *Arch Phys Med Rehabil* 1997;78:767–73.
- Poole JL, Steen VD. The use of the Health Assessment Questionnaire (HAQ) to determine physical disability in systemic sclerosis. *Arthritis Care Res* 1991;4:27–31.
- Ramey DR, Raynauld JP, Fries JF. The health assessment questionnaire 1992: status and review. *Arthritis Care Res* 1992;5:119–29.
- Bruce B, Fries JF. The Stanford Health Assessment Questionnaire: a review of its history, issues, progress, and documentation. *J Rheumatol* 2003;30:167–78.
- Merkel PA, Herlyn K, Martin RW, Anderson JJ, Mayes MD, Bell P, et al. Measuring disease activity and functional status in patients with scleroderma and Raynaud's phenomenon. *Arthritis Rheum* 2002;46:2410–20.
- Poole JL, Williams CA, Bloch DA, Hollak B, Spitz P. Concurrent validity of the Health Assessment Questionnaire Disability Index in scleroderma. *Arthritis Care Res* 1995;8:189–93.
- Steen VD, Medsger TA Jr. The value of the Health Assessment Questionnaire and special patient-generated scales to demonstrate change in systemic sclerosis patients over time. *Arthritis Rheum* 1997;40:1984–91.
- Clements PJ, Wong WK, Hurwitz EL, Furst DE, Mayes M, White B, et al. The Disability Index of the Health Assessment Questionnaire is a predictor and correlate of outcome in the high-dose versus low-dose penicillamine in systemic sclerosis trial. *Arthritis Rheum* 2001;44:653–61.
- Khanna D, Furst DE, Clements PJ, Park GS, Hays RD, Yoon J, et al, and the Relaxin Study Group Scleroderma Clinical Trials Consortium. Responsiveness of the SF-36 and the Health Assessment Questionnaire Disability Index in a systemic sclerosis clinical trial. *J Rheumatol* 2005;32:832–40.
- Georges C, Chassany O, Mouthon L, Tiev K, Toledano C, Meyer O, et al. Validation of French version of the Scleroderma Health Assessment Questionnaire (SSc HAQ). *Clin Rheumatol* 2005;24:3–10.
- Marra CA, Woolcott JC, Kopec JA, Shojania K, Offer R, Brazier JE, et al. A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. *Soc Sci Med* 2005;60:1571–82.
- Brazier JE, Harper R, Munro J, Walters SJ, Snaith L. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. *Rheumatology (Oxford)* 1999;38:870–7.
- Poole JL. Grasp pattern variations seen in the scleroderma hand. *Am J Occup Ther* 1994;48:46–54.
- Sandqvist G, Eklund M. Hand Mobility in Scleroderma (HAMIS) test: the reliability of a novel hand function test. *Arthritis Care Res* 2000;13:369–74.
- Sandqvist G, Eklund M. Validity of HAMIS: a test of hand mobility in scleroderma. *Arthritis Care Res* 2000;13:382–7.
- Duruoz MT, Poiraudeau S, Fermanian J, Menkes CJ, Amor B, Dougados M, et al. Development and validation of a rheumatoid hand functional disability scale that assesses functional handicap. *J Rheumatol* 1996;23:1167–72.
- Poiraudeau S, Lefevre-Colau MM, Fermanian J, Revel M. The ability of the Cochin rheumatoid arthritis hand functional scale to detect change during the course of disease. *Arthritis Care Res* 2000;13:296–303.
- Poiraudeau S, Chevalier X, Conrozier T, Flippo RM, Liote F, Noel E, et al. Reliability, validity, and sensitivity to change of the Cochin hand functional disability scale in hand osteoarthritis. *Osteoarthritis Cartilage* 2001;9:570–7.
- Brower LM, Poole JL. Reliability and validity of the Duruoz Hand Index in persons with systemic sclerosis (scleroderma). *Arthritis Rheum* 2004;51:805–9.
- Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;118:622–9.
- Ware JE Jr, Sherbourne DC. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473–83.
- Strand V. Clinical trial design in systemic lupus erythematosus: lessons learned and future directions. *Lupus* 2004;13:406–11.
- Kosinski MA, Keller SD, Hatoum HT, Kong SX, Ware JE Jr. The SF-36 Health Survey as a generic outcome measure in clinical trials of patients with osteoarthritis and rheumatoid arthritis: tests of data quality, scaling assumptions and score reliability. *Med Care* 1999;37:MS10–22.
- Georges C, Chassany O, Mouthon L, Tiev K, Marjanovic Z, Meyer O, et al. Quality of life assessment with the MOS-SF36 in patients with systemic sclerosis. *Rev Med Interne* 2004;25:16–21. In French.
- Khanna D, Clements PJ, Furst DE, Chon Y, Elashoff R, Roth MD, et al, and the Scleroderma Lung Study Group. Correlation of the degree of dyspnea with health-related quality of life, functional abilities, and diffusing capacity for carbon monoxide in patients with systemic sclerosis and active alveolitis: results from the Scleroderma Lung Study. *Arthritis Rheum* 2005;52:592–600.
- Subcommittee for Scleroderma Criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Preliminary criteria for the classification of systemic sclerosis (scleroderma). *Arthritis Rheum* 1980;23:581–90.
- LeRoy EC, Medsger TA Jr. Criteria for the classification of early systemic sclerosis. *J Rheumatol* 2001;28:1573–6.
- Keitel W, Hoffmann H, Weber G, Krieger U. Evaluation of the percentage of functional decrease of the joints using a motor function test in rheumatology. *Dtsch Gesundheitsw* 1971;26:1901–3. In German.
- Kapandji A. Clinical test of apposition and counter-apposition of the thumb. *Ann Chir Main* 1986;5:67–73. In French.
- Kapandji A. Proposal for a clinical score for flexion-extension of the long fingers. *Ann Chir Main* 1987;6:288–94. In French.
- Lefevre-Colau MM, Poiraudeau S, Oberlin C, Demaille S, Fermanian J, Rannou F, et al. Reliability, validity, and responsiveness of the modified Kapandji index for assessment of functional mobility of the rheumatoid hand. *Arch Phys Med Rehabil* 2003;84:1032–8.
- Tugwell P, Bombardier C, Buchanan WW, Goldsmith CH, Grace E, Hanna B. The MACTAR Patient Preference Disability Questionnaire: an individualized functional priority ap-

- proach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. *J Rheumatol* 1987;14:446–51.
38. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983;67:361–70.
  39. Fermanian J. Mesure de l'accord entre deux juges: cas quantitatif. *Rev Epidém et Santé Publ* 1984;32:408–13.
  40. Bowling A, Bond M, Jenkinson C, Lamping DL. Short form 36 (SF-36) Health Survey questionnaire: which normative data should be used? Comparisons between norms provided by the Omnibus survey in Britain, the Health survey for England and the Oxford Healthy Life Survey. *J Public Health Med* 1999;21:255–70.
  41. Leplege A, Ecosse E, Verdier A, Perneger TV. The French SF-36 Health Survey: translation, cultural adaptation and preliminary psychometric evaluation. *J Clin Epidemiol* 1998;51:1013–23.
  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.
  43. Faucher M, Poiraudéau S, Lefevre-Colau MM, Rannou F, Fermanian J, Revel M. Algo-functional assessment of knee osteoarthritis: comparison of the test-retest reliability and construct validity of the WOMAC and Lequesne indexes. *Osteoarthritis Cartilage* 2002;10:602–10.
  44. Comrey AL. *A first course in factor analysis*. New York: New York Academic Press; 1973.
  45. Kline P. *The hand book of psychological testing*. London: Rouledge; 1993.
  46. Nunnally JO. *Psychometric theory*. New York: McGraw Hill; 1978.
  47. Guadagnoli E, Velicer WF. Relation of sample size to the stability of component patterns. *Psychol Bull* 1988;103:265–75.
  48. Tabachnik BF, Fidell LS. *Using multivariable statistics*. 2nd ed. New York: Harper and Collins; 1989.