Evaluation of the Satisfaction with Appearance Scale and Its Short Form in Systemic Sclerosis: Analysis from the UCLA Scleroderma Quality of Life Study

Sarah D. Mills, Rina S. Fox, Erin L. Merz, Philip J. Clements, Suzanne Kafaja, Vanessa L. Malcarne, Daniel E. Furst, and Dinesh Khanna

ABSTRACT. ***Objective.*** Changes in appearance are common in patients with systemic sclerosis (SSc) and can significantly affect well-being. The Satisfaction with Appearance Scale (SWAP) measures body image dissatisfaction in persons with visible disfigurement; the Brief-Satisfaction with Appearance Scale (Brief-SWAP) is its short form. The present study evaluated the reliability and validity of SWAP and Brief-SWAP scores in SSc.

Methods. A sample of 207 patients with SSc participating in the University of California, Los Angeles Scleroderma Quality of Life Study completed the SWAP. Brief-SWAP scores were derived from the SWAP. The structural validity of both measures was investigated using confirmatory factor analysis. Internal consistency reliability of total and subscale scores was assessed with Cronbach's alpha coefficients. Convergent and divergent validity was evaluated using the Center for Epidemiological Studies Depression Scale, the Health Assessment Questionnaire-Disability Index, and the Medical Outcomes Study Short Form-36 questionnaire.

Results. SWAP and Brief-SWAP total scores were highly correlated (r = 0.97). The 4-factor structure of the SWAP fit well descriptively; the 2-factor structure of the Brief-SWAP fit well descriptively and statistically. Internal consistencies for total and subscale scores were good, and results supported convergent and divergent validity.

Conclusion. Both versions are suitable for use in patients with SSc. The Brief-SWAP is most efficient; the full SWAP yields additional subscales that may be informative in understanding body image issues in patients with SSc. (First Release June 1 2015; J Rheumatol 2015;42:1624–30; doi:10.3899/jrheum.141482)

Key Indexing Terms:
QUALITY OF LIFE SYSTEMIC SCLEROSIS SCLERODERMA

Systemic sclerosis (SSc) is a chronic, multisystem, connective tissue disease that attacks healthy body tissue[1]. SSc may be divided into 2 subtypes: (1) limited cutaneous SSc is characterized by skin involvement limited to the fingers, hands, lower arms, lower legs, and face; (2) and diffuse cutaneous SSc includes more widespread skin and organ involvement[1,2,3]. Visible changes in appearance are common in patients with SSc. These changes can harm self-image and quality of life[2,4,5], and may result in body image dissatisfaction (BID; also called appearance dissatisfaction) and psychological distress[6,7]. To date, there has been limited research on BID in SSc[4,5].

More research on BID in SSc, and its relationship to quality of life, is needed. However, research efforts have been hampered by the lack of measures appropriate to and validated for patients with SSc. Although a variety of BID measures are available, most were intended for use in other settings or with other populations (e.g., eating disorders) and require modifications, and/or have not been evaluated for use in SSc. Identifying measures of BID that are reliable and valid for use in SSc is critical to understanding how disease-related physical changes affect quality of life in this population.

From the San Diego State University (SDSU)/University of California, San Diego (UC San Diego) Joint Doctoral Program in Clinical Psychology, and Department of Psychology, SDSU, San Diego; Department of Psychology, California State University, Dominguez Hills, Carson; Department of Medicine, University of California, Los Angeles, Los Angeles, California; Department of Medicine, University of Michigan, Ann Arbor, Michigan, USA.

Supported by the grant Evaluation of Health-related Quality of Life in Systemic Sclerosis from the Scleroderma Foundation Inc. Dr. Khanna has been funded by the US National Institutes of Health (NIH)/National Institute of Arthritis and Musculoskeletal and Skin Diseases K24 AR063120 and K23 AR053858. Ms. Mills was supported by the UC San Diego Cota-Robles Fellowship. Dr. Furst has served as a consultant or on speaker bureaus for the NIH.

S.D. Mills, MS, MPH, Doctoral Student; R.S. Fox, MS, MPH, Doctoral Student, SDSU/UC San Diego Joint Doctoral Program in Clinical Psychology; E.L. Merz, PhD, MPH, Assistant Professor, Department of Psychology, California State University, Dominguez Hills; P.J. Clements, MD, MPH, Professor Emeritus; S. Kafaja, MD, Clinical Instructor, Department of Medicine, University of California, Los Angeles; V.L. Malcarne, PhD, Professor, Department of Psychology, SDSU/UC San Diego Joint Doctoral Program in Clinical Psychology, and Department of Psychology, SDSU; D.E. Furst, MD, Professor, Department of Medicine, University of California, Los Angeles; D. Khanna, MD, MSc, Associate Professor, Department of Medicine, University of Michigan.

Address correspondence to Dr. V.L. Malcarne, Doctoral Training Facility, 6363 Alvarado Court, Suite 103, San Diego, California 92120-4913, USA. E-mail: vmalcarne@mail.sdsu.edu

Accepted for publication April 21, 2015.

The Satisfaction with Appearance Scale (SWAP). The SWAP[8] is a 14-item measure of BID that was originally developed for use with individuals with physical disfigurements as a result of burn injuries, but has since been adapted and used in research on SSc[9,10]. The SWAP was designed to measure 2 central aspects of body image: subjective satisfaction with appearance, and the social-behavioral effect of disfigurement, and a 2-factor structure was hypothesized[11]. In the original validation sample of patients with burn injuries, unexpectedly, each of the SWAP's 14 items loaded onto 1 of 4 factors (subscales) labeled Social Distress, Facial Features, Non-Facial Features, and Perceived Social Impact[8].

The SWAP has since been adapted for use in SSc with the word "burn" replaced with the word "illness" or "scleroderma." Few studies, however, have examined the psychometric properties of the measure in SSc. Benrud-Larson, *et al*[12] used the SWAP to examine the relationship between BID and psychosocial functioning in 129 female, predominantly white patients with SSc. Internal consistency reliability was excellent ($\alpha = 0.90$) and SWAP total scores significantly correlated with measures of depressive symptoms, disability, psychosocial functioning, and pain in the expected directions and magnitudes, providing evidence of convergent validity. The factor structure of the measure, however, was not evaluated.

Jewett, *et al*[10] examined the psychometric properties of the SWAP in a sample of 217 women with SSc from the Johns Hopkins Scleroderma Center (JHSC) and 654 women with SSc from the Canadian Scleroderma Research Group (CSRG) registry. Patients were predominantly white and diagnosed with the limited disease subtype (70% and 72.2%, respectively). Internal consistency reliability was excellent in both samples ($\alpha = 0.90$ and 0.91). Confirmatory factor analysis (CFA) was used to examine a 2-factor structure (Subjective Dissatisfaction and Perceived Social Impact). After 2 pairs of item error covariances were freed, the 2-factor structure fit well in both samples based on descriptive fit indices. Evidence of convergent validity for the total score of the measure was provided by significant correlations in expected directions with measures of depressive symptoms, pain, and quality of life.

Heinberg, *et al*[9] analyzed the factor structure of the SWAP using a sample (n = 254) drawn from the same Johns Hopkins dataset as Jewett, *et al*[10], but including a 15th item that had been administered ("My appearance makes others feel uncomfortable"). Patients completed the 15-item version of the SWAP at baseline and 18 months later. The sample was predominantly female, white, and diagnosed with limited disease. Principal components analysis at each timepoint resulted in the extraction of 2 factors, Subjective Dissatisfaction and Perceived Social Impact, with the new item added to the latter scale. Internal consistency reliability was good for both subscales ($\alpha \geq 0.88$). Based on this, the authors suggested that the 4-factor structure reported for patients with burn injuries was not suitable for persons with SSc[9]. However, the authors used exploratory methods and did not statistically compare a 4-factor model with a 2-factor model.

The Brief-SWAP. Jewett, *et al*[10] derived a 6-item Brief-SWAP from the more commonly used 14-item SWAP, attempting to retain the 2 subscales (Subjective Dissatisfaction and Perceived Social Impact) previously identified for SSc. Jewett, *et al* argued that many of the items of the SWAP were superfluous, and chose 3 items to represent each of the 2 subscales based on theoretical and psychometric considerations. As described above, Jewett, *et al* analyzed data from samples of female patients with SSc from the JHSC and the CSRG. CFA showed support for the hypothesized 2-factor model of the Brief-SWAP. Two 3-item subscales were supported and named Subjective Dissatisfaction and Perceived Social Impact. Internal consistency reliability (Cronbach's alpha coefficient) for the Brief-SWAP total score was 0.82 in both samples.

A second study by the same research team, also using a Canadian sample drawn from the CSRG registry, evaluated the psychometric properties of the Brief-SWAP in 489 women and men with SSc[13]. The 2-factor structure was replicated using CFA, and the same two 3-item subscales were derived, renamed as Dissatisfaction with Appearance (replacing Subjective Dissatisfaction) and Social Discomfort (replacing Perceived Social Impact). Internal consistency reliability was good for both subscales ($\alpha = 0.82$ and 0.83, respectively).

To date, the structural validity of the SWAP and Brief-SWAP has only been examined in an all-female sample from the JHSC, and a female and male sample from the CSRG. An additional study examined the structural validity of the 15-item version of the SWAP in a predominantly female sample, also drawn from the JHSC. Further examination of the reliability and validity of the SWAP and Brief-SWAP is needed in distinct populations of patients with SSc to further establish the generalizability of these measures' psychometric properties.

The present study contributes to the literature by attempting to replicate previously reported factor structure and psychometric findings from previous studies[8,9,10,13] in a diverse sample in terms of sex, ethnicity, and disease subtype. The aims of this study were to (1) examine and compare the structural validities of the SWAP and Brief-SWAP, (2) examine and compare internal consistency reliability coefficients for the SWAP and Brief-SWAP, and (3) examine and compare convergent and divergent validity for the 2 measures.

MATERIALS AND METHODS

Patients. The sample consisted of 207 patients with SSc (confirmed by study rheumatologists) who were participating in a single-center, longitudinal study. Disease subtype classification was made according to American College of Rheumatology criteria[14]. The study was approved by the University of California, Los Angeles Institutional Review Board.

SWAP[8]. The SWAP is a 14-item measure of BID. Table 1 gives individual items and corresponding subscales. Respondents rate the extent to which each item reflects their feelings about their appearance on a scale ranging from 1 (strongly disagree) to 7 (strongly agree). Items 4 to 11 are reverse-scored. Total scores, as well as 4 subscale scores (Social Distress, Facial Features, Non-Facial Features, and Perceived Social Impact), can be calculated. To calculate SWAP scores, 1 is subtracted from each item to anchor all items at 0, and then item scores are summed. Scores for the Facial Features and Non-Facial Features subscales can range from 0 to 24, and scores for the Social Distress and Perceived Social Impact subscales can range from 0 to 18. Total scores can range from 0 to 84. Higher scores indicate greater BID. Completion time is estimated at 5 min.

Brief-SWAP[10]. The Brief-SWAP is a 6-item short form derived from the SWAP[8]. Table 1 shows individual items and corresponding subscales. Total scores, as well as 2 subscale scores (Dissatisfaction with Appearance, Social Discomfort), can be calculated. Scores are calculated by subtracting 1 from each item to anchor items at 0. Items for the Dissatisfaction with Appearance subscale are reverse-scored, and then item scores are totaled. Subscale scores can range from 0 to 18, and total scores can range from 0 to 36. Higher scores indicate greater BID. The Brief-SWAP was not given in the present study; rather, Brief-SWAP scores were derived from the SWAP. Completion time is estimated at 2 min.

Center for Epidemiologic Studies Depression Scale-Short Form (CES-D Short Form)[15]. The CES-D Short Form is a 10-item version of the widely used CES-D[16], a screening measure of depressive symptoms. Scores can range from 0 to 30, with higher scores indicating more frequent depressive symptoms. Internal consistency reliability was good in the present sample ($\alpha = 0.83$).

Health Assessment Questionnaire-Disability Index (HAQ-DI)[17]. The HAQ-DI is a 20-item measure of functional ability that has been validated for SSc[18,19]. Responses are rated on a scale ranging from 0 (no disability) to 3 (completely disabled). A total score is calculated by averaging the 8 category scores (i.e., dressing, rising, walking, eating, hygiene, reach, grip, and usual activities). The HAQ-DI demonstrated strong internal consistency reliability in the present sample ($\alpha = 0.93$).

Modified Rodnan Skin Score (mRSS)[20]. The mRSS is a physician-administered measure of skin disease severity validated for patients with SSc[21,22]. The mRSS total score is determined by measuring the scope and severity of skin thickening in 17 body areas by palpitation on a scale ranging from 0 (uninvolved) to 3 (severe thickening). Scores can range from 0 to 51, with higher scores indicating greater severity.

Medical Outcomes Study Short Form-36 questionnaire (SF-36)[23]. The SF-36 measures quality of life in 8 domains. Physical component summary (PCS) and mental component summary (MCS) scores are derived from the domain scores, with higher scores indicating better quality of life. The SF-36 has previously demonstrated good reliability and validity in patients with SSc[24]. The standard 4-week recall version of the SF-36 version 2.0 was used.

Statistical analysis. Descriptive statistics for demographic and medical variables, and all measures, were calculated for the total sample. Pearson correlations were calculated to demonstrate overlapping variance between the SWAP and Brief-SWAP.

CFA was used to determine the best fitting factor structures of the SWAP and Brief-SWAP in patients with SSc. The goodness of fit of the previously established 4-factor structure (Social Distress, Facial Features, Non-Facial Features, and Perceived Social Impact) of the 14-item SWAP and the 2-factor structure (Dissatisfaction with Appearance and Social Discomfort) of the 6-item Brief-SWAP were initially examined. Interfactor correlations were specified among the latent variables. As recommended by Bentler, overall model fit was determined by consulting 3 fit indices[25]: (1) the root mean square error of approximation (RMSEA)[26], an absolute index of overall model fit; (2) the standardized root mean residual (SRMR)[27]; and (3) the robust comparative fit index (CFI)[28]. For RMSEA and SRMR indices, values less than 0.08 were considered acceptable fit and values less than 0.05 were considered good fit. For CFI, values greater than 0.90 were considered acceptable fit and values greater than 0.95 were considered good fit. Models were determined to fit well if values for at least 2 of the descriptive fit indices

Table 1. Factor loadings of the 4-factor SWAP and 2-factor Brief-SWAP. All factor loadings are significant ($p < 0.01$) for both the SWAP and Brief-SWAP. For the Brief-SWAP, only factor loadings for the 6 items are presented. Values are factor loadings.

Items	4-factor SWAP	2-factor Brief-SWAP
Facial features		
I am satisfied with my overall appearance.	0.82	—
I am satisfied with the appearance of my scalp.	0.64	—
I am satisfied with the appearance of my face.	0.87	0.71*
I am satisfied with the appearance of my neck.	0.77	—
Non-facial features		
I am satisfied with the appearance of my hands.	0.67	0.74*
I am satisfied with the appearance of my arms.	0.85	0.80*
I am satisfied with the appearance of my legs.	0.78	—
I am satisfied with the appearance of my chest.	0.82	—
Social distress		
Because of changes in my appearance caused by my scleroderma, I am uncomfortable in the presence of my family.	0.78	—
Because of changes in my appearance caused by my scleroderma, I am uncomfortable in the presence of my friends.	0.85	—
Because of changes in my appearance caused by my scleroderma, I am uncomfortable in the presence of strangers.	0.82	0.72**
Perceived social impact		
Changes in my appearance have interfered with my relationships.	0.77	—
I feel that my scleroderma is unattractive to others.	0.85	0.89**
I don't think people would want to touch me.	0.82	0.77**

* Subjective Dissatisfaction subscale items. ** Perceived Social Impact subscale. SWAP: Satisfaction with Appearance Scale; Brief-SWAP: Brief-Satisfaction with Appearance Scale.

indicated at least acceptable model fit. The likelihood ratio chi-square was also reported for completeness; however, it was not used as the primary indicator of model fit because it is highly influenced by sample size and almost always statistically significant, and thus not a good index of degree of fit[29].

Next, the best fitting factor structures for the SWAP and Brief-SWAP were compared. Because likelihood-ratio tests cannot be used to compare non-nested models[30], the Akaike information criterion (AIC)[31] and the sample size-adjusted Bayesian information criterion (sBIC)[32] were used to evaluate comparative model fit. For both criteria, smaller values indicate better model fit. Both AIC and sBIC criteria reward parsimony. Thus, model comparison using AIC and sBIC indices were considered in conjunction with other model fit and psychometric validation results.

Internal consistency reliability was examined for the SWAP, Brief-SWAP, and all subscales using Cronbach's alpha coefficient. Convergent validity constructs were selected to replicate previous research[8,9,10] using constructs known to be associated with BID in patients with SSc. The factors for each form of the SWAP and Brief-SWAP were expected to be moderately positively associated with measures of depressive symptoms (CES-D), physical function (HAQ-DI), and disease severity (mRSS), and moderately negatively associated with a quality of life measure (SF-36 PCS and MCS). For divergent validity, based on previous research[8], the SWAP and Brief-SWAP were expected to have little to no correlation with bodily pain (SF-36 Bodily Pain Scale), after controlling for depression.

RESULTS

Descriptive statistics. Table 2 gives sample characteristics and means and SD for all measures. The sample (n = 207) was predominantly female (83.1%), white (71.5%), married (57%), and had some college or higher education (81.6%). Mean age of the sample was 54.1 years (SD 15.4). About half of the sample had limited SSc (50.2%), followed by diffuse SSc (40.1%). Time since diagnosis of SSc was 7.57 years (SD 7.9) and the mean mRSS, a widely used measure of disease severity, was 8.70 (SD 8.5). The mean percent predicted forced vital capacity for the total sample was 78.98% (SD 21.76). Only 4.8% of patients reported renal crisis. The correlation between SWAP and Brief-SWAP total scores was significant and very strong ($r = 0.97$, $p < 0.01$).

Table 2. Sociodemographic and disease variables for patients with SSc from the UCLA Scleroderma Quality of Life Study (n = 207). Values are mean (SD) unless otherwise specified.

Variable	Value
Demographic variables	
Age, yrs	54.1 (15.4)
White, n (%)	148 (71.5)
Highest level of education, n (%)	
Some college or higher	169 (81.6)
Annual income, n (%)	
≥ US$75,000	71 (34.3)
Female, n (%)	172 (83.1)
Married, n (%)	118 (57.0)
Medical variables	
Time since diagnosis of SSc, yrs	7.57 (7.9)
Disease type, n (%)	
Limited SSc	104 (50.2)
Diffuse SSc	83 (40.1)
Sine SSc	4 (1.9)
Overlap	9 (4.3)
Missing	7 (3.4)
mRSS score	8.70 (8.5) [max = 51]
Predicted FVC, %	78.98 (21.76)
Renal crisis, n (%)	10 (4.8)
Self-report questionnaire scores	
SWAP	32.6 (20.3) [max = 84]
Brief-SWAP	16.3 (9.6) [max = 36]
CES-D	8.3 (5.8) [max = 30]
SF-36 MCS	48.8 (12.2) [max = 100]
SF-36 PCS	38.7 (10.0) [max = 100]
HAQ-DI	0.8 (0.7) [max = 3]

SSc: systemic sclerosis; UCLA: University of California, Los Angeles; mRSS: Modified Rodnan skin score; FVC: forced vital capacity; SWAP: Satisfaction with Appearance Scale; Brief-SWAP: Brief-Satisfaction with Appearance Scale; CES-D: Center for Epidemiologic Studies Depression Scale-Short Form; SF-36: Medical Outcomes Study Short Form-36; MCS: mental component summary; PCS: physical component summary; HAQ-DI: Health Assessment Questionnaire-Disability Index.

SWAP. First, a 4-factor model for the 14-item SWAP was examined using CFA (Table 1). Interfactor correlations were specified among the 4 latent variables. This 4-factor model did not fit well statistically [chi-square (71) = 149.01, $p < 0.01$], but it did fit well descriptively (RMSEA = 0.07, SRMR = 0.04; CFI = 0.96). Correlations among the 4 factors were all statistically significant (Table 3). Next, a 2-factor model for the SWAP was examined. The Dissatisfaction with Appearance factor was identified by 8 variables (combining the Facial Features and Non-Facial Features subscales) while the Social Discomfort factor was identified by 6 variables (combining the Social Distress and Perceived Social Impact subscales). This 2-factor model did not fit well statistically [chi-square (76) = 274.23, $p < 0.01$], but it did fit well descriptively (RMSEA = 0.11, SRMR = 0.06, CFI = 0.90). The interfactor correlation was large and statistically significant ($r = 0.71$, $p < 0.01$). A chi-square difference test was used to statistically compare the 4-factor model to the 2-factor model. The 2 models were statistically significantly different [Δchi-square (5) = 125.22, $p < 0.01$], indicating that the 4-factor model fit the observed data better than the 2-factor model.

Brief-SWAP. A 2-factor model for the 6-item Brief-SWAP was tested using CFA (Table 1 gives all standardized factor loadings for this model). An interfactor correlation was specified between the 2 latent variables. This 2-factor model fit well statistically [chi-square (8) = 14.24, $p = 0.08$], and descriptively (RMSEA = 0.06, SRMR = 0.03; CFI = 0.99). The interfactor correlation was large and statistically significant ($r = 0.79$, $p < 0.01$). Given the high interfactor correlation, a 1-factor model was also tested with a single latent variable indicated by 6 observed variables. This 1-factor model did not fit well statistically [chi-square (9) = 52.08, $p < 0.01$], but it did fit well descriptively (RMSEA = 0.15, SRMR = 0.05; CFI = 0.92). The 2 models were then statistically compared to determine the superior fit to the data. A chi-square difference test demonstrated that the 2 models fit

Table 3. Intercorrelations of SWAP subscales from the 4-factor confirmatory factor analysis. Values are r.

Variables	Facial Features	Non-facial Features	Social Distress	Perceived Social Impact
Facial features	1.00	0.84*	0.63*	0.69*
Non-facial features	—	1.00	0.57*	0.72*
Social distress	—	—	1.00	0.82*
Perceived social impact	—	—	—	1.00

* $p < 0.05$. SWAP: Satisfaction with Appearance Scale.

differently [Δchi-square (1) = 37.85, $p < 0.01$], indicating that the 2-factor model fit the observed data better than the 1-factor model.

The 2 best fitting models, the 4-factor model for the SWAP and the 2-factor model for the Brief-SWAP, were then compared. The AIC and sBIC values were lower for the 2-factor Brief-SWAP than for the 4-factor SWAP (AIC = 4790.32 vs 10,334.99. sBIC = 4798.44 vs 10,342.87), suggesting that the 2-factor Brief-SWAP provided better model fit to the observed data.

Internal consistency reliability. Internal consistency reliability was excellent for the SWAP ($\alpha = 0.93$) and good for the Brief-SWAP ($\alpha = 0.87$). All hypothesized subscales of the SWAP and Brief-SWAP also had good reliability (SWAP: Facial Features: $\alpha = 0.86$, Non-Facial Features: $\alpha = 0.86$, Social Distress: $\alpha = 0.89$, Perceived Social Impact: $\alpha = 0.85$; Brief-SWAP: Dissatisfaction with Appearance: $\alpha = 0.79$, Social Discomfort: $\alpha = 0.83$).

Convergent and divergent validity. As anticipated, significant positive moderate correlations with depression, level of physical functioning, and disease severity were found for both the SWAP and Brief-SWAP (Table 4 and Table 5). Also, as expected, better mental and physical health-related quality of life was associated with greater satisfaction with appearance. Providing evidence of divergent validity, after controlling for depression, the relationships of bodily pain to the SWAP and Brief-SWAP scores were nonsignificant. For the subscales of the SWAP and Brief-SWAP, all correlations were significant, of expected magnitudes, and in expected directions.

DISCUSSION

Our study examined the psychometric properties of the SWAP and Brief-SWAP in a sample of patients with SSc in the United States. Total scores on the SWAP and Brief-SWAP were similar to those reported for other SSc samples[10,12,13]. In addition, replicating previous studies[10,12], mean SWAP total scores were higher than those from the original sample of hospitalized patients with burn injuries[8].

A primary aim was to identify and compare the best-fitting factor structures for the SWAP and Brief-SWAP. In the present analysis, a 4-factor model best fit the data for the SWAP, supporting the use of the 4 subscales in SSc (Facial Features, Non-Facial Features, Social Discomfort, Perceived Social Impact). For the Brief-SWAP, the 2-factor model best fit the data, supporting the use of the 2 Brief-SWAP subscales, Dissatisfaction with Appearance and Social Discomfort. The 2-factor structure of the Brief-SWAP demonstrated better fit to the sample data than did the 4-factor structure of the longer SWAP. Alpha coefficients for all total scores and subscales demonstrated good reliability. Therefore, with 8 fewer items, the Brief-SWAP more parsimoniously measures BID. Jewett, *et al*[10] suggested that the 2-factor Brief-SWAP provided better fit because the Brief-SWAP contains items that focused on body parts relevant in SSc, and items were removed from the 14-item SWAP that were endorsed by a few patients with SSc. In the current sample, it is also not surprising that the Brief-SWAP demonstrated better comparative model fit to the SWAP, given that the AIC and sBIC indicators reward parsimony[33].

Table 4. Convergent and discriminant validity results for the SWAP and four SWAP subscales. Values are r (95% CI).

Variables	SWAP, Total Score	SWAP, Facial Features	SWAP, Non-facial Features	SWAP, Social Distress	SWAP, Perceived Social Impact
CES-D	0.39 (0.26–0.51)**	0.25 (0.11–0.40)**	0.33 (0.19–0.46)**	0.49 (0.26–0.51)**	0.39 (0.25–0.51)**
HAQ-DI	0.33 (0.18–0.45)**	0.21 (0.07–0.35)**	0.38 (0.23–0.40)**	0.24 (0.10–0.38)**	0.26 (0.12–0.40)**
mRSS	0.25 (0.12–0.36)**	0.16 (0.01–0.29)*	0.24 (0.12–0.35)**	0.22 (0.09–0.35)**	0.25 (0.10–0.38)**
SF-36 PCS	–0.24 (–0.38 – –0.09)**	–0.17 (–0.33 – –0.03)*	–0.29 (–0.43 – –0.14)**	–0.16 (–0.30 – –0.02)*	–0.19 (–0.32 – –0.05)*
SF-36 MCS	–0.35 (–0.48 – –0.22)**	–0.24 (–0.39 – –0.11)**	–0.21 (–0.35 – –0.06)**	–0.38 (–0.50 – –0.26)**	–0.39 (–0.50 – –0.26)**
SF-36 bodily pain†	–0.01(–0.15–0.13)	0.10 (–0.05–0.24)	–0.11 (–0.25–0.04)	0.01 (–0.14–0.14)	–0.03 (–0.17–0.11)

* $p < 0.05$ (2-tailed). **$p < 0.01$ (2-tailed). † Partial correlations controlling for depression using the CES-D. SWAP: Satisfaction with Appearance Scale; CES-D: Center for Epidemiologic Studies Depression Scale-Short Form; HAQ-DI: Health Assessment Questionnaire-Disability Index; mRSS: Modified Rodnan skin score; SF-36: Medical Outcomes Study Short Form-36; PCS: physical component summary; MCS: mental component summary.

Table 5. Convergent and discriminant validity results for the Brief-SWAP and 2 Brief-SWAP subscales*. Values are r (95% CI).

Variables	Brief-SWAP, Total Score	Brief-SWAP, Dissatisfaction with Appearance	Brief-SWAP, Social Discomfort
CES-D	0.33 (0.20–0.46)**	0.24 (0.10–0.38)**	0.36 (0.23–0.49)**
HAQ-DI	0.34 (0.21–0.45)**	0.32 (0.20–0.44)**	0.28 (0.14–0.41)**
mRSS	0.28 (0.17–0.40)**	0.25 (0.12–0.36)**	0.27 (0.14–0.40)**
SF-36 PCS	–0.22 (–0.36 – –0.08)**	–0.22 (–0.36 – –0.07)**	–0.19 (–0.33 – –0.05)**
SF-36 MCS	–0.30 (–0.43 – –0.17)**	–0.18 (–0.31 – –0.05)*	–0.36 (–0.47 – –0.23)**
SF-36 bodily pain†	–0.04 (–0.18–0.10)	–0.03 (–0.18–0.10)	–0.04 (–0.18–0.10)

* $p < 0.05$ (2-tailed). **$p < 0.01$ (2-tailed). † Partial correlations controlling for depression using the CES-D. Brief-SWAP: Brief-Satisfaction with Appearance Scale; CES-D: Center for Epidemiologic Studies Depression Scale-Short Form; HAQ-DI: Health Assessment Questionnaire-Disability Index; mRSS: Modified Rodnan skin score; SF-36: Medical Outcomes Study Short Form-36; PCS: physical component summary; MCS: mental component summary.

However, both models had good overall fit and convergent validity, suggesting that decision making regarding which measure to use should not be based purely on this comparison. Rather, either measure may be useful, depending on the type of information a researcher or clinician is seeking.

The present sample differs from previous validation samples on several key demographic characteristics. First, both men and women are included, unlike the original Canadian study validating the Brief-SWAP that had an all-female sample[10]. Additionally, the present sample had a higher percentage of patients with diffuse disease in comparison with the CSRG and JHSC samples. The proportion of patients with diffuse versus limited disease varies greatly depending on geographic region and ethnicity, with some epidemiological studies reporting diffuse disease in more than 70% of the disease population[34]. In addition, the present study sample had a lower percentage of white patients compared with previous samples. Data from multiethnic cohorts suggest that non-white patients are at increased risk for more severe SSc, in particular regarding diffuse skin involvement[34]. Also, patients with diffuse disease often report higher levels of BID.

There are limitations to the current study. Only the original SWAP was completed; Brief-SWAP scores were derived from the original measure, and item order and context have been shown to influence responses[35]. Because there are no other measures of BID that have been validated for use in SSc, convergent validity analyses focused on measures of constructs previously found to be associated with BID in patients with SSc.

The present findings support the use of the SWAP and Brief-SWAP in patients with SSc. Previous studies using the SWAP or Brief-SWAP have reported both total and subscale scores. In the present study, correlations among subscales for the SWAP and Brief-SWAP were large and statistically significant, suggesting that use of a total score to provide an overall measure of BID is appropriate for both measures. In addition, the factor analyses suggested that subscale scores can be used to assess particular aspects of BID. The SWAP may be preferred in research because it includes 4 subscales that measure specific aspects of BID. However, the Brief-SWAP's 2 subscales yield information on both subjective dissatisfaction with appearance and appearance-related social concerns while reducing administration time. The Brief-SWAP may be a useful screening measure, aiding in the identification of individuals in need of additional assessment and support.

REFERENCES

1. Seibold J. Scleroderma. In: Harris ED Jr, Budd RC, Genovese MC, Firestein GS, Sergent JS, Sledge C, eds. Kelley's textbook of rheumatology, 7th edition. Philadelphia: Elsevier Saunders; 2005:1279-308.
2. Clements PJ, Furst DE. Systemic sclerosis, 2nd edition. Baltimore: Lippincott Williams & Wilkins; 2003.
3. Schieir O, Thombs BD, Hudson M, Boivin JF, Steele R, Bernatsky S, et al; Canadian Scleroderma Research Group. Prevalence, severity, and clinical correlates of pain in patients with systemic sclerosis. Arthritis Care Res 2010;62:409-17.
4. Malcarne VL, Fox RS, Mills SD, Gholizadeh S. Psychosocial aspects of systemic sclerosis. Curr Opin Rheumatol 2013;25:707-13.
5. Thombs BD, van Lankveld W, Bassel M, Baron M, Buzza R, Haslam S, et al. Psychological health and well-being in systemic sclerosis: state of the science and consensus research agenda. Arthritis Care Res 2010;62:1181-9.
6. Malcarne VL, Hansdottir I, Greenbergs HL, Clements PJ, Weisman MH. Appearance self-esteem in systemic sclerosis. Cognit Ther Res 1999;23:197-208.
7. van Lankveld WG, Vonk MC, Teunissen H, van den Hoogen FH. Appearance self-esteem in systemic sclerosis—subjective experience of skin deformity and its relationship with physician-assessed skin involvement, disease status and psychological variables. Rheumatology 2007;46:872-6.
8. Lawrence JW, Heinberg LJ, Roca R, Munster A, Spence R, Fauerbach JA. Development and validation of the Satisfaction With Appearance Scale: assessing body image among burn-injured patients. Psychol Assess 1998;10:64-70.
9. Heinberg LJ, Kudel I, White B, Kwan A, Medley K, Wigley F, et al. Assessing body image in patients with systemic sclerosis (scleroderma): validation of the Adapted Satisfaction with Appearance Scale. Body Image 2007;4:79-86.
10. Jewett LR, Hudson M, Haythornthwaite JA, Heinberg L, Wigley

FM, Baron M, et al; Canadian Scleroderma Research Group. Development and validation of the brief-satisfaction with appearance scale for systemic sclerosis. Arthritis Care Res 2010;62:1779-86.
11. Heinberg LJ. Theories of body image disturbance: perceptual, developmental, and sociocultural factors. In: Thompson JK, ed. Body image, eating disorders, and obesity: an integrative guide for assessment and treatment. Washington: American Psychological Association; 1996:27-47.
12. Benrud-Larson LM, Heinberg LJ, Boling C, Reed J, White B, Wigley FM, et al. Body image dissatisfaction among women with scleroderma: extent and relationship to psychosocial function. Health Psychol 2003;22:130-9.
13. Jewett LR, Hudson M, Malcarne VL, Baron M, Thombs BD; Canadian Scleroderma Research Group. Sociodemographic and disease correlates of body image distress among patients with systemic sclerosis. PLoS One 2012;7:e33281.
14. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Arthritis Rheum 1980;23:581-90.
15. Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). Am J Prev Med 1994;10:77-84.
16. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977; 1:385-401.
17. Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales. J Rheumatol 1982;9:789-93.
18. Cole JC, Khanna D, Clements PJ, Seibold JR, Tashkin DP, Paulus HE, et al. Single-factor scoring validation for the Health Assessment Questionnaire-Disability Index (HAQ-DI) in patients with systemic sclerosis and comparison with early rheumatoid arthritis patients. Qual Life Res 2006;15:1383-94.
19. 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.
20. 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.
21. Clements P, Lachenbruch P, Siebold J, White B, Weiner S, Martin R, et al. Inter and intraobserver variability of total skin thickness score (modified Rodnan TSS) in systemic sclerosis. J Rheumatol 1995;22:1281-5.
22. Clements PJ, Lachenbruch PA, Seibold JR, Zee B, Steen VD, Brennan P, et al. Skin thickness score in systemic sclerosis: an assessment of interobserver variability in 3 independent studies. J Rheumatol 1993;20:1892-6.
23. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473-83.
24. Khanna D, Furst DE, Clements PJ, Park GS, Hays RD, Yoon J, et al; 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.
25. Bentler P. On tests and indices for evaluating structural models. Pers Individ Dif 2007;42:825-9.
26. Steiger JH. Structural model evaluation and modification: an interval estimation approach. Multivariate Behav Res 1990; 25:173-80.
27. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999;6:1-55.
28. Bentler PM. Comparative fit indexes in structural models. Psychol Bull 1990;107:238-46.
29. Beauducel A, Wittmann WW. Simulation study on fit indexes in CFA based on data with slightly distorted simple structure. Struct Equ Modeling 2005;12:41-75.
30. Harrington D. Chapter 4: Assessing confirmatory factor analysis model fit and model revision. In: Harrington D, ed. Confirmatory factor analysis. New York: Oxford University Press; 2009:50-77.
31. Gerbing DW, Anderson JC. Monte Carlo evaluations of goodness-of-fit indices for structural equation models. In: Bollen KA, Long JS, eds. Testing structural equation models. California: Sage; 1993:40-65.
32. Akaike H. A new look at the statistical model identification. IEEE Trans Automat Control 1974;19:716-23.
33. Kline RB. Principles and practice of structural equation modeling, 2nd edition. New York: The Guilford Press; 2005.
34. Reveille JD. Ethnicity and race and systemic sclerosis: how it affects susceptibility, severity, antibody genetics, and clinical manifestations. Curr Rheumatol Rep 2003;5:160-7.
35. Schwarz N. Self reports: how the questions shape the answers. Am Psychol 1999;54:93-105.