Published in final edited form as:

Spec Care Dentist. 2014 January ; 34(1): . doi:10.1111/scd.12036.

Test-Retest Reliability of Oral Health Impact Profile (OHIP-49) in Adults with Systemic Sclerosis

Hon K. Yuen, PhD, OTR/L [Professor and and Director of Research] and

Department of Occupational Therapy, School of Health Professions, University of Alabama at Birmingham, 1530 3rd Avenue South, Birmingham, AL 35294. Tel: 205-934-6301; Fax: 205-975-7787

Sybil L. Nelson, MS [Graduate student]

Division of Biostatistics and Epidemiology, Department of Medicine, College of Medicine, Medical University of South Carolina, SC 29425

Hon K. Yuen: yuen@uab.edu; Sybil L. Nelson: princene@musc.edu

Abstract

Purpose—To assess the test-retest reliability of the Oral Health Impact Profile (OHIP) in 3-month intervals among people with systemic sclerosis (SSc).

Methods—Thirty-nine adults with SSc completed the OHIP-49 at baseline, 3 months and 6 months. Intraclass correlation (ICC), standard error of measurement (SEM), SEM%, coefficient of repeatability (CoR), and 95% limits of agreement (LoA) were used to assess test-retest reliability and measurement variability.

Results—The ICC of the OHIP for the first time point (baseline to 3-month) was excellent (>.80), and the second time point (3-month to 6-month) demonstrated good reproducibility (>.60). However, the large SEM, SEM%, CoR and 95% LoA indicated that the OHIP was neither a precise measurement nor sensitive to change in response to interventions.

Conclusion—The values of the ICC indicated that the OHIP scores demonstrated acceptable stability across each of the two 3 months' retest duration, supporting the reliability of the OHIP for group-level comparisons. However, using the OHIP as a measure of an individual's oral health-related quality of life in adults with SSc is not recommended.

Introduction

Systemic sclerosis (SSc) belongs to a group of autoimmune connective tissue diseases and is characterized by inflammation, vascular dysfunction, and excessive fibrosis of connective tissue supporting the skin and visceral organs.¹ SSc is classified into two main types: limited and diffuse cutaneous SSc. Limited cutaneous SSc is characterized by restricted skin involvement (often just the hands and face), prominent vascular manifestations, exhibiting features of CREST (calcinosis, Raynaud's phenomenon, esophageal dysmotility, sclerodactyly, telangiectasias) syndrome, and slow progression.^{2,3} Diffuse cutaneous SSc is characterized by rapidly progressing symmetrical widespread skin involvement, and visceral organ involvement.^{2,3}

Orofacial complications, which include xerostomia (dry mouth) and microstomia (small mouth), are major clinical manifestations that contribute to oral health problems in people with SSc.⁴ Oral distress, such as a burning sensation resulting from xerostomia, has been

Conflict of Interest: The author declares that they have no conflict of interest.

shown to have a negative impact on both general and oral health-related quality of life (OHRQoL).⁷ Whereas, small mouth and mouth furrows are the two top orofacial concerns reported by adults with SSc.⁸ In addition to the psychological distress secondary to facial disfigurement, microstomia can result in multiple functional debilitating sequelae, such as difficulty with mastication and speech.^{4,9} As a result, microstomia can have a profound negative impact on the relationships of patients with others (e.g., social interactions) and on their quality of life.⁹

The Oral Health Impact Profile (OHIP) is the most widely used instrument to measure OHRQoL in adults with oral disease conditions.¹⁰ OHIP, a self-administered questionnaire, is designed to measure people's perception on the frequency of problems associated with the mouth, teeth, or dentures on their well-being.¹⁰ OHIP-49 consists of 49 questions covering seven dimensions (functional limitations, physical pain, psychological discomfort, physical disability, psychological disability, social disability, and handicap). Each question is rated on a 5-point Likert scale (0 = never, 1 = hardly ever, 2 = occasionally, 3 = fairly often, 4 = very often) according to how often the person has felt the severity of oral distress during the past twelve months. The unweighted sum of all questions (i.e., additive method) is recommended to compute the total score.¹¹ As a result, the total score ranges from 0 to 196 with higher scores indicating poor OHRQoL. OHIP-14, a shortened version of the OHIP-49, consists of two questions in each of the seven above-mentioned dimensions.¹² Studies showed the degree of reliability and validity of the OHIP-14, as a measure of OHRQoL in community living older adults, are comparable to that of the OHIP-49.¹² The psychometric properties of OHIP (-49 and -14) have been evaluated in a variety of oral disease conditions; however, few studies included patients exhibiting systemic diseases such as SSc.

Stability of an instrument as applied to repeated measurement in clinical trials is crucial. Without a sound knowledge on the test-retest reliability of a measure, it would not be accurate to assess the degree to which an instrument is sensitive enough to detect a change in responses of a particular patient group to targeted interventions.¹³ As recommended, test-retest is most informative when the time interval chosen between the test and retest is appropriate for identifying stability in reference to the clinical trial protocol.¹⁴ Given that the follow-up duration on the majority of clinical trials in oral health intervention were beyond 1 month, with 3 and 6 month intervals were common,¹⁵⁻¹⁷ it is important to consider these time intervals for retesting when investigating the stability of the OHIP. Given the lack of reliability information of the OHIP as it applies to measurement of OHRQoL in people with SSc, the purpose of this study is to assess the test-retest reliability of OHIP (-49 and -14) in 3-month intervals among people with SSc.

Methods

Participants

Patients with SSc from the scleroderma clinic at the Medical University of South Carolina (MUSC) with records in the local connective tissue disease database were invited to participate. Participants eligible for the study were adults (aged > 18 years old) diagnosed with SSc at least 1 year prior to study baseline assessment, and who fulfilled the American College of Rheumatology preliminary classification criteria for SSc.¹⁸ Exclusion criteria were localized scleroderma (e.g. morphea, linear scleroderma, and en coup de sabre), less than 10 natural teeth, an upper and/or lower full denture, or complaint of any major jaw joint problems (e.g., severe pain or dislocation). Details of the subject selection criteria have been reported elsewhere.¹⁹

Procedures

This study was part of a larger trial evaluating the effect of a multi-faceted intervention program on oral health status in people with SSc which employed a single-blinded, randomized, controlled design.¹⁹ Following the baseline mouth assessment, eligible participants were randomly assigned to one of two groups: a multi-faceted oral health intervention or a usual dental care control group. Participants in the multi-faceted oral health intervention group each received a rechargeable, powered Oral-B® oscillating-rotating-pulsating toothbrush and a Reach® Access™ Flosser, and taught how to use each device to properly clean their teeth. Participants in the control group each received a manual toothbrush (Oral-B® Complete Advantage Deep Clean toothbrush) and dental floss (Crest® glide shred guard floss), and taught how to floss their teeth. Participants assigned to the intervention group who had an oral aperture of less than 40 mm, were taught to perform manual mouth-stretching and oral augmentation exercises. At 3-month and 6-month post baseline assessments, participants received another clinical mouth assessment. In addition, participants completed the OHIP-49 at each assessment time point. To compensate for health illiteracy and/or hand function impairments of the participants, the research coordinator read each question to them and recorded responses. The protocol was approved by the MUSC Institutional Review Board.

Data analysis

To compute the composite score of the OHIP-14 and OHIP-49, we used the additive method, which is the summation of the scores of all 14 or 49 questions, respectively. Results of the clinical trial indicated that the multi-faceted intervention group had a statistically significant positive effect on the reduction of gingival inflammation at the 6-month assessment,¹⁹ but no clinically significant improvement (~10%) on the size of oral aperture after the orofacial stretching exercise and no statistically significant difference from the control group on the size of oral aperture at the 6-month assessment.²⁰ Since only a small increase in the size of the oral aperture and a moderate improvement in the inflammatory status of gingival tissue were seen across the study period, the multi-faceted intervention should have little impact on participants' OHRQoL. To confirm that, several t tests were performed to assure that no significant differences in the OHIP-49 scores between the intervention and the control groups at each assessment occasion (baseline, 3-month, and 6-month) were observed, with all p-values >.05. In addition, the associations between changes in the size of oral aperture and changes in the OHIP-49 scores across any two successive assessment occasions (i.e., baseline to 3-month and 3-month to 6-month) were not significant (p-values >.05), with Pearson's correlation coefficients ranging from .02 to .19. Results of the same analyses on each group (intervention and control) were consistent with that of the whole group. As a result, data from both groups were combined to evaluate the test-retest reliability of the two OHIP versions (i.e., OHIP-49 and OHIP-14).

To assess the test-retest reliability (reproducibility or stability) between any two successive measurement time points of the OHIP-49 and OHIP-14, the following statistical methods were used:

1. The 95% confidence interval (CI) for the mean differences between any two successive measurement time points of the OHIP scores was used to assess systematic change between assessment occasions.²¹ If the CI did not include zero, a statistically significant difference (i.e., systematic change) in the OHIP scores between assessment occasions was indicated.²¹
2. Association between successive assessment occasions was assessed by the intraclass correlation coefficient (ICC), which is a measure of relative reliability.²¹ The ICC with 95% CI provided an estimate of variance due to measurement error

of the OHIP. The ICC[2,1] model used to assess test-retest reliability is derived from a 2-way random-effect model in which both time and participant are modeled as random, and which addresses both systematic and random errors (i.e., absolute agreement).^{22,23} Values of ICC range from 0 to 1 with 0 indicating no reproducibility and 1 indicating perfect reproducibility. ICC values above .75 are considered to have excellent reliability, values between .6 and .75 are good reliability, values between .4 and .6 are fair reliability, with values below .4 suggestive of poor reliability.^{24,25}

3. The standard error of measurement (SEM) quantified the actual size of the OHIP measurement variability (in the unit of measurement) between two assessment time points and gave the measurement error in absolute values (absolute reliability).²¹ SEM provided an estimate of how reliably the OHIP estimated an individual's true score. SEM is equal to (MS_E) or $SD * (1-ICC)$, where MS_E is the within-subject mean square error term, and SD is the mean standard deviation of two time points.²¹ SEM represents the limit for the smallest change that indicates a real (clinical) improvement for a group of individuals.²¹ In other words, in order for the OHIP to indicate a real change for a group of participants, its scores needed to be outside the range of measurement error. SEM% was also computed which is equal to $(SEM / \text{mean of all measurements from the two occasions}) * 100\%$.²⁶
4. The coefficient of repeatability (CoR) of a measure for reliability estimation is the minimal detectable difference (i.e. beyond measurement error between two time points) that must be observed before concluding that a real (clinical) change in an individual's score on a measure has occurred.²⁷ Any change in score that is less than the CoR value reflects the measurement error of the instrument.²⁷ CoR is defined as the value below which the absolute difference between test-retest scores is expected to lie within 95% CI.²⁷ CoR as a 95% CI for the difference between two measurement occasions equals to $1.96 * SEM * \sqrt{2}$.²⁷
5. As the ICC does not measure agreement which results in ignoring any systematic bias between the two measurement occasions,²⁸ the Bland-Altman method on limits of agreement (LoA), which assesses the agreement between measurement occasions, was used to provide an indication of the systematic random error of the OHIP data.²⁹ The 95% LoA around the mean difference were calculated as 1.96 times the standard deviation of the difference scores.²⁹ The 95% LoA provided an interval where 95% of the differences in the OHIP scores between two time points were expected to lie, which suggests that there is no systematic bias or instrument 'learning'.²⁹ This statistic represented the test-retest differences expected for 95% of the participants.²⁹ Values exceeding 95% LoA indicated a real (clinical) change in an individual participant.²⁹

Results

There were 39 participants (22 in the intervention group and 17 in the control group) who completed the OHIP-49 at baseline, 3- and 6-month assessments. Table 1 shows the socio-demographics and clinical characteristics of the 39 participants. The mean $\pm$ SD and median time interval between baseline and 3-month assessment were 3.5 ± 1.1 and 3.2 months, respectively. The mean $\pm$ SD and median time interval between 3- and 6-month assessments were 3.7 ± 1.5 and 3.4 months, respectively. There was no significant difference in retest duration between these two periods with $P = .2$. The mean $\pm$ SD of the OHIP-49 and OHIP-14 scores at three assessment time points for the whole ($N=39$), intervention ($n=22$), and control ($n=17$) groups are shown in Table 2.

The 95% CI for the mean difference of the OHIP scores between baseline and 3-month assessment did not include zero, which indicated a significant systematic bias, with scores systematically lower (better) at 3-month (see Table 3). Further analyses of each group (intervention and control) revealed that 14 of 17 participants in the control group indicated an improvement in the OHIP scores at 3-month, however, no significant systematic change in the OHIP scores from baseline to 3-month was observed in the intervention group (see Table 4). As a result, we also assessed the test-retest reliability of each study group.

The ICC of the OHIP-49 for the first two assessment time points (baseline to 3-month) and the second two time points (3-month to 6-month) was .84 and .69 with CIs of (.66, .92) and (.48, .82) respectively (see Table 3). Whereas, for the intervention group, the ICCs of the OHIP-49 for the first and second two time points were .84 and .86, with a much narrower CI (.70, .94) for the second two time points (i.e., 3-month to 6-month) (see Table 4). Even a significant systematic bias was observed at the 3-month retest for the control group, the ICC of the OHIP-49 was .85, which is excellent; however, the CI was very wide (.40 to .95) (see Table 5).

The SEM of the OHIP-49 for the first two time points was 12.1, and 18.3 for the second two time point (see Table 3). The mean score for the first two time points was 33, and the SEM % was 36.6%; whereas, the mean for the second two time points was 28.7, and the SEM % was 63.9%. The CoR values demonstrated a change of at least 33.4 for the first two time points and 50.8 for the second two time points suggesting a very large score change. The 95% LoAs used for describing the total measurement error for the first and second two time points were -24.4 and 42.4, and -51.3 and 50.4, which means the amount of measurement error varied across participants with a range of 66.8 and 101.7, respectively, suggesting a very large between-subject variability.

Applying the same statistical methods to each group (intervention and control), the indices of test-retest reliability and measurement variability are shown in Tables 4 and 5. Overall, results of the OHIP-14 were congruent with that of the OHIP-49 (see Tables 3–5).

Discussion

Results of the ICC indicated that the OHIP (-49 and -14) scores demonstrated acceptable stability across each of the two 3 months' retest periods in these study participants. The values of the ICC and its 95% CI supported the reliability of the OHIP-49 and -14 for group-level comparisons in people with SSc. However, the large SEMs and SEM%s indicated that the OHIP-49 and -14 demonstrated large measurement variability, and the large CoR also indicated that the OHIP was neither a precise measurement nor sensitive to change among the study participants. The 95% LoA used for describing the total measurement error for each of the two 3 months' retest duration were very large. Therefore, using either version of the OHIP as a measure of an individual's OHRQoL in adults with SSc is not recommended.

The much larger change in the OHIP-49 and -14 scores from baseline to 3-month assessment compared to 3-month to 6-month assessment may be related to regression to the mean. Other factors such as medication and dental treatment prior to completing the OHIP-49 were not controlled, which may also contribute to the instability of the OHIP. The reason for the presence of systematic bias between baseline and 3-month assessment for the OHIP scores in the control group was unknown as participants from each study group had no knowledge on what treatment the other group received.

The high value of ICC with systematic differences between the baseline and 3 month time points indicated that within-subject variability was small, relative to between-subject variability. However, a lower value of ICC and wide 95% CI (especially the lower limits of

the 95% CI) between 3 and 6 month time points, indicating less precision in the estimation of the reliability coefficient, is likely attributable to the poor ICC of the control group. On the other hand, the values of the ICC and its 95% CI of the OHIP for the intervention group are remarkably similar, with both retest durations about the same (~3 months).

It is understandable that the length of time between the two test administrations has an effect on the test-retest reliability. We did not expect any significant health status changes of the participants during the study period (i.e., across 6 months) as the majority of them were under close medical supervision (with regular [i.e., 3-month] follow-up from their rheumatologists) and their symptoms were under control through medication. However, a longer interval, with SSc disease progression, and time without preventive and restorative dental care, may increase the chances that a change in participants' oral health status could occur. Studies of test-retest reliability for OHRQoL instruments usually employed an interval of less than 1 month between test administrations, with the majority of them employing an interval of around 2 weeks.^{30,31} This time frame is generally believed to be a reasonable compromise between recollection bias and unwanted clinical change. If the time is too long, true changes in the construct may be reflected in the scores, as quality of life is a dynamic construct.³² When a time interval of 1–2 weeks was used for OHIP, excellent reliability coefficients (ICC) were observed.^{30,31} For a time interval of 3 months between administrations, coefficients were much lower (around .6),^{33,34} which is regarded as fair reliability. It is possible that if a shorter retest duration was employed in the present study, the width of the 95% CI for the ICC, the CoR, and the 95% LoA may have been reduced, thus making the measure more sensitive to detecting change with better measurement precision.

Limitations

The exclusion criteria such as less than 10 natural teeth and no major jaw joint problems, though having such oral issues are not very common, did exclude a small portion of adults with SSc from participation in the study, which may affect the representativeness of the study participants. Future study should investigate other aspects of the psychometric properties of OHIP as it applied to the SSc population.

Conclusion

Many oral health intervention studies have a 3 to 6 months follow-up time to evaluate the treatment effects on OHRQoL which needs to be measured. The acceptable ICC of OHIP scores at 3-month retest for adults with SSc is a prerequisite for assessment of OHRQoL changes over a relatively long time periods. The values of the ICC and its 95% CI supported the reliability of the OHIP for group-level comparisons in people with SSc. However, due to poor precision measurement, insensitivity to change and large amount of measurement error which varied across participants, using the OHIP as a measure of an individual's oral health-related quality of life in adults with SSc is not recommended.

Acknowledgments

The authors wish to thank the project coordinator, Samantha Mahoney, for subject recruitment and data entry, and Dr. Laura Vogtle for her suggestions.

Source of Funding: This study was supported by the Scleroderma Foundation and the National Institutes of Health, National Center of Research Resources Grant Number P20 RR017696, South Carolina COBRE for Oral Health Research.

References

1. Derk CT, Jimenez SA. Systemic sclerosis: current views of its pathogenesis. *Autoimmun Rev.* 2003; 2:181–191. [PubMed: 12848944]
2. 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; 23:581–590. [PubMed: 7378088]
3. LeRoy EC, Black C, Fleischmajer R, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. *J Rheumatol.* 1988; 15:202–205. [PubMed: 3361530]
4. Scardina GA, Messina P. Systemic sclerosis: description and diagnostic role of the oral phenomena. *Gen Dent.* 2004; 52:42–47. [PubMed: 15055670]
5. Avouac J, Sordet C, Depinay C, et al. Systemic sclerosis-associated Sjogren's syndrome and relationship to the limited cutaneous subtype: results of a prospective study of sicca syndrome in 133 consecutive patients. *Arthritis Rheum.* 2006; 54:2243–2249. [PubMed: 16802363]
6. Swaminathan S, Goldblatt F, Dugar M, et al. Prevalence of sicca symptoms in a South Australian cohort with systemic sclerosis. *Intern Med J.* 2008; 38:897–903. [PubMed: 18422567]
7. Stewart CM, Berg KM, Cha S, et al. Salivary dysfunction and quality of life in Sjogren syndrome: a critical oral-systemic connection. *J Am Dent Assoc.* 2008; 139:291–299. quiz 358–299. [PubMed: 18310733]
8. Paquette DL, Falanga V. Cutaneous concerns of scleroderma patients. *J Dermatol.* 2003; 30:438–443. [PubMed: 12810990]
9. Nagy G, Kovacs J, Zeher M, et al. Analysis of the oral manifestations of systemic sclerosis. *Oral Surg Oral Med Oral Pathol.* 1994; 77:141–146. [PubMed: 8139831]
10. Slade GD, Spencer AJ. Development and evaluation of the Oral Health Impact Profile. *Community Dent Health.* 1994; 11:3–11. [PubMed: 8193981]
11. Allen PF, Locker D. Do item weights matter? An assessment using the oral health impact profile. *Community Dent Health.* 1997; 14:133–138. [PubMed: 9332036]
12. Slade GD. Derivation and validation of a short-form oral health impact profile. *Community Dent Oral Epidemiol.* 1997; 25:284–290. [PubMed: 9332805]
13. Rudner LM, Schafer WD. How to write a scholarly research report. *Practical Assessment, Research & Evaluation.* 1999:6.
14. Health USDo, Human Services FDACfDE, Research et al. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes.* 2006; 4:79. [PubMed: 17034633]
15. Robinson PG, Pankhurst CL, Garrett EJ. Randomized-controlled trial: effect of a reservoir biteguard on quality of life in xerostomia. *J Oral Pathol Med.* 2005; 34:193–197. [PubMed: 15752252]
16. Saito A, Ota K, Hosaka Y, et al. Potential impact of surgical periodontal therapy on oral health-related quality of life in patients with periodontitis: a pilot study. *J Clin Periodontol.* 2011; 38:1115–1121. [PubMed: 22093073]
17. Eitner S, Wichmann M, Schlegel KA, et al. Oral health-related quality of life and implant therapy: an evaluation of preoperative, intermediate, and post-treatment assessments of patients and physicians. *J Craniomaxillofac Surg.* 2012; 40:20–23. [PubMed: 21482130]
18. Masi AT. 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; 23:581–590. [PubMed: 7378088]
19. Yuen HK, Weng Y, Bandyopadhyay D, et al. Effect of a multi-faceted intervention on gingival health among adults with systemic sclerosis. *Clin Exp Rheumatol.* 2011; 29:S26–32. [PubMed: 21586215]
20. Yuen HK, Marlow NM, Reed SG, et al. Effect of orofacial exercises on oral aperture in adults with systemic sclerosis. *Disabil Rehabil.* 2012; 34:84–89. [PubMed: 21951278]
21. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res.* 2005; 19:231–240. [PubMed: 15705040]

22. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979; 86:420–428. [PubMed: 18839484]
23. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods.* 1996; 1:30–46.
24. Fleiss, J. *The Design and Analysis of Clinical Experiments.* New York: Wiley; 1986.
25. Juniper, EF.; Guyatt, GH.; Jaeschke, R. How to develop and validate a new health-related quality of life instrument. In: Spilker, S., editor. *Quality of life and pharmacoeconomics in clinical trials.* Philadelphia: Lippincott-Raven; 1996. p. 49-56.
26. Lexell JE, Downham DY. How to assess the reliability of measurements in rehabilitation. *Am J Phys Med Rehabil.* 2005; 84:719–723. [PubMed: 16141752]
27. Bland, JM. *An Introduction to Medical Statistics.* 3. Oxford: Oxford University Press; 2000.
28. Bland JM, Altman DG. Applying the right statistics: analyses of measurement studies. *Ultrasound Obstet Gynecol.* 2003; 22:85–93. [PubMed: 12858311]
29. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986; 1:307–310. [PubMed: 2868172]
30. Kuo HC, Chen JH, Wu JH, et al. Application of the Oral Health Impact Profile (OHIP) among Taiwanese elderly. *Qual Life Res.* 2011; 20:1707–1713. [PubMed: 21472393]
31. Reissmann DR, John MT, Schierz O. Influence of administration method on oral health-related quality of life assessment using the Oral Health Impact Profile. *Eur J Oral Sci.* 2011; 119:73–78. [PubMed: 21244515]
32. Allison PJ, Locker D, Feine JS. Quality of life: a dynamic construct. *Soc Sci Med.* 1997; 45:221–230. [PubMed: 9225410]
33. Bae KH, Kim HD, Jung SH, et al. Validation of the Korean version of the oral health impact profile among the Korean elderly. *Community Dent Oral Epidemiol.* 2007; 35:73–79. [PubMed: 17244140]
34. Souza RF, Patrocinio L, Pero AC, et al. Reliability and validation of a Brazilian version of the Oral Health Impact Profile for assessing edentulous subjects. *J Oral Rehabil.* 2007; 34:821–826. [PubMed: 17919248]

Table 1

Characteristics of the Participants (N=39)

Characteristics	Mean±SD (range) or n (%)
Diagnosis and co-morbidities	
Diffuse cutaneous SSc	14 (35.9%)
Limited cutaneous SSc	25 (64.1%)
Disease duration (years)	7.8±6.1 (1.0–24.7)
Oral aperture (mm)	36.7±9.6 (10.0–56.7)
Sicca syndrome	11 (28.2%)
Self-report dry mouth symptoms	29 (74.4%)
Number of missing teeth (excluded 3 rd molars)	3.9±4.3 (0–18)
Socio-demographics	
Age (years)	51.9±12.8 (22–76)
Female	31 (79.5%)
African American ^a	21 (53.8%)
Married / live together	18 (46.2%)
Education (less than college)	20 (51.3%)
Employment (FT and PT)	14 (35.9%)
Annual income (< \$20,000)	15 (38.5%)
Annual income (\$55,000)	13 (33.3%)
No dental insurance	17 (43.6%)
Non-smoker (current)	36 (92.3%)

^a Only African Americans and Caucasians; SSc = systemic sclerosis; FT = full time, PT = part time; SD = standard deviation

Table 2

The mean and standard deviation of the OHIP-49 and OHIP-14 scores at three assessment time points for the whole (N=39), intervention (n=22), and control (n=17) groups

	N=39		n=22		n=17	
	OHIP-49	OHIP-14	OHIP-49	OHIP-14	OHIP-49	OHIP-14
Baseline	37.46±36.92	9.97±11.53	33.36±34.94	8.09±10.59	42.76±39.78	12.41±12.54
3-month	28.46±29.26	7.41±8.89	27.27±31.39	7.18±9.65	30.00±27.13	7.71±8.08
6-month	28.95±35.47	7.44±10.43	28.50±32.17	7.32±9.52	29.53±40.36	7.59±11.80

Note. OHIP = Oral Health Impact Profile

Table 3

Indices of test-retest reliability and measurement variability for the whole group (N=39)

	Time	Mean _{diff} ±SD _{diff}	ICC (95% CI)	SEM	SEM%	CoR	BA method on 95% LoA, lower limit	BA method on 95% LoA, upper limit
OHIP-49	Baseline-3 mo	9.0±17.04(3.48, 14.52)	.84 (.66, .92)	12.05	36.56	33.40	-24.40	42.40
	3 mo - 6 mo	-.49±25.94(-8.90, 7.92)	.69 (.48, .82)	18.34	63.89	50.84	-51.33	50.35
OHIP-14	Baseline-3 mo	2.56±5.75 (.70, 4.43)	.82 (.65, .91)	4.07	46.84	11.28	-8.71	13.83
	3 mo - 6 mo	-.03±8.61(-2.82, 2.77)	.61 (.37, .78)	6.09	82.02	16.88	-16.92	16.85

Notes. The negative values in Meandiff indicate that the OHIP scores were higher (i.e., poor oral health-related quality of life) at retest time point.

BA = Bland-Altman, CoR = coefficient of repeatability, ICC = intraclass correlation, LoA = limits of agreement, Meandiff = mean difference, mo = month, OHIP = Oral Health Impact Profile, SD_{diff} = standard deviation of the difference scores, SEM = standard error of measurement, SEM% = (SEM / mean of all measurements from the two occasions)*100%.

Table 4
Indices of test-retest reliability and measurement variability for the intervention group (n=22)

Time	Mean _{diff} ±SD _{diff} (95% CI)	ICC (95% CI)	SEM	SEM%	CoR	BA method on 95% LoA, lower limit	BA method on 95% LoA, upper limit
OHIP-49	Baseline-3 mo	6.09±18.15 (-1.96, 14.14)	.84 (.66, .93)	12.83	42.32	35.56	-29.48
	3 mo - 6 mo	-1.23±16.99 (-8.76, 6.30)	.86 (.70, .94)	12.01	43.07	33.29	-34.53
OHIP-14	Baseline-3 mo	.91±5.54 (-1.55, 3.36)	.85 (.68, .94)	3.92	51.34	10.87	-9.95
	3 mo - 6 mo	-.14±6.06 (-2.82, 2.55)	.81 (.59, .92)	4.28	59.03	11.86	-12.02

Notes. The negative values in Meandiff indicate that the OHIP scores were higher (i.e., poor oral health-related quality of life) at retest time point.

BA = Bland-Altman, CoR = coefficient of repeatability, ICC = intraclass correlation, LoA = limits of agreement, Meandiff = mean difference, mo = month, OHIP = Oral Health Impact Profile, SD_{diff} = standard deviation of the difference scores, SEM = standard error of measurement, SEM% = (SEM / mean of all measurements from the two occasions)*100%.

Table 5

Indices of test-retest reliability and measurement variability for the control group (n=17)

	Time	Mean _{diff} ±SD _{diff} (95% CI)	ICC (95% CI)	SEM	SEM%	CoR	BA method on 95% LoA, lower limit	BA method on 95% LoA, upper limit
OHIP-49	Baseline-3 mo	12.77±15.19 (4.96, 20.57)	.85 (.40, .95)	10.74	29.52	29.77	-17.00	42.54
	3 mo – 6 mo	.47±34.89 (-17.47, 18.41)	.50 (.02, .79)	24.67	82.88	68.39	-67.91	68.85
OHIP-14	Baseline-3 mo	4.71±5.45 (1.90, 7.51)	.79 (.28, .93)	3.85	38.27	10.67	-5.97	15.39
	3 mo – 6 mo	.12±11.31 (-5.70, 5.93)	.39 (-.12, .73)	8.00	104.58	22.18	-22.05	22.29

Notes. The negative values in Meandiff indicate that the OHIP scores were higher (i.e., poor oral health-related quality of life) at retest time point.

BA = Bland-Altman, CoR = coefficient of repeatability, ICC = intraclass correlation, LoA = limits of agreement, Meandiff = mean difference, mo = month, OHIP = Oral Health Impact Profile, SD_{diff} = standard deviation of the difference scores, SEM = standard error of measurement, SEM% = (SEM / mean of all measurements from the two occasions)*100%.