
Reliability and validity of the delta finger-to-palm (FTP), a new measure of finger range of motion in systemic sclerosis

K.S. Torok¹, N.A. Baker², M. Lucas³, R.T. Domsic³, R. Boudreau⁴, T.A. Medsger Jr³

¹Children's Hospital of Pittsburgh, Department of Pediatrics, Division of Rheumatology; ²School of Health and Rehabilitation Sciences, Department of Occupational Therapy; ³Department of Medicine, Division of Rheumatology and Clinical Immunology; ⁴Department of Epidemiology, Center for Aging and Population Health; University of Pittsburgh, Pittsburgh, Pennsylvania, USA.

Kathryn S. Torok, MD
Nancy A. Baker, ScD
Mary Lucas, RN, MPH
Robyn T. Domsic MD, MPH
Robert Boudreau, PhD
Thomas A. Medsger Jr, MD

Please address correspondence to:

Thomas A. Medsger Jr, MD,
Division of Rheumatology and
Clinical Immunology,
3500 Terrace Street, Room 720,
Pittsburgh, PA 15261, USA.
E-mail: medsger@dom.pitt.edu

Received on November 17, 2009; accepted
in revised form on April 20, 2010.

Clin Exp Rheumatol 2010; 28 (Suppl. 58):
S28-S36.

© Copyright CLINICAL AND
EXPERIMENTAL RHEUMATOLOGY 2010.

Key words: Systemic sclerosis, hand,
outcome measures, physical function.

ABSTRACT

Objectives. To determine the reliability and validity of a new measure of finger motion in patients with systemic sclerosis (SSc), the 'delta finger-to-palm' (delta FTP) and compare its psychometric properties to the traditional measure of finger motion, the finger-to-palm (FTP).

Methods. Phase 1: The reliability of the delta FTP and FTP were examined in 39 patients with SSc. Phase 2: Criterion and convergent construct validity of both measures were examined in 17 patients with SSc by comparing them to other clinical measures: Total Active Range of Motion (TAROM), Hand Mobility in Scleroderma (HAMIS), the Duruoz Hand Index (DHI), Health Assessment Questionnaire (HAQ), and modified Rodnan skin score (mRSS). Phase 3: Sensitivity to change of the delta FTP was investigated in 24 patients with early diffuse cutaneous SSc.

Results. Both measures had excellent intra-rater and inter-rater reliability (ICC 0.92 to 0.99). Fair to strong correlations ($r_s=0.49-0.94$) were observed between the delta FTP and TAROM, HAMIS, and DHI. Fair to moderate correlations were observed between delta FTP and HAQ components related to hand function and upper extremity mRSS. Correlations of the traditional FTP with these measures were fair to strong, but most often the delta FTP outperformed the FTP. The effect size and standardised response mean for the mean delta FTP were 0.50 and 1.10 respectively, over a 2–8 month period.

Conclusion. The delta FTP is a valid and reliable measure of finger motion in patients with SSc which outperforms the FTP.

Introduction

Systemic sclerosis (SSc) is an autoimmune disease which involves multiple organ systems. In addition to skin, SSc affects subcutaneous tissue, tendons,

and synovium of the hands, resulting in finger joint contractures (1). These contractures cause limitation in finger range of motion (ROM), leading to functional limitations in activities of daily living (ADLs) (2-7) and are associated with disease severity and progression (8). An outcome measure which easily and accurately determines finger ROM in patients with SSc would be desirable, both for patient management and use in clinical trials.

The time-honoured method for measuring finger ROM in SSc is the finger-to-palm (FTP) distance (distance from the tip of the third finger to the distal palmar crease in maximal active flexion), sometimes referred to as fist closure (9). Although recommended as a secondary outcome measure for clinical trials in SSc, the FTP has been validated in only one study, in which the authors reported an interobserver variation of 24% and a coefficient of variation of 0.52 (10). To date, the FTP has been shown to be only a fair outcome measure (11). Experts agree that the FTP possesses face (credibility) and partial construct (biological sense) validity, but has not been assessed for criterion (accuracy) or discriminate (sensitivity to change) validity (11). To address this issue and improve on the FTP, we propose a new measure of finger ROM, the "delta FTP", which combines both finger joint flexion and extension. The delta FTP is the difference of the distance measured between the 3rd fingertip and the distal palmar crease with fingers in full extension and the distance with fingers in full flexion. Validity is the extent to which an instrument measures what it is intended to measure (12). There are several different types of validity. Face validity or credibility considers whether a measure reflects what it is intended to measure (13). The FTP provides a summation of flexion of all three finger joints (MCP, PIP, DIP). However, the FTP does not represent full finger motion because

Competing interests: none declared.

limitations in finger extension have not been taken into account. Some patients with SSc have fingertips which are “fixed” in palmar flexion without the ability to extend. These individuals have severe hand dysfunction, but paradoxically have a “falsely normal” traditional FTP measurement (8). The misclassification of this group of patients as having mild or no limitation threatens the face validity of the FTP. Preliminary studies comparing the FTP to the delta FTP in our SSc patient cohort support the delta FTP as a more sensitive indicator of functional limitation. In 25 SSc patients with low delta FTPs (<4cm), indicating severely limited ROM, the mean concurrent HAQ-DI was 1.21, considerably higher (more disability) than the mean values previously published in SSc (HAQ-DI = 0.86–1.09) (4, 14–17). Six (24%) of the individuals in this sample had an FTP measurement of <2.0 cm and would have been misclassified as having minimal or no limitation in finger motion if only the traditional FTP had been recorded (4, 16).

In concurrent criterion validity, the criterion data (gold standard) are collected at the same time as the new measurement (or “predictor” data) for a particular concept. The current ‘gold standard’ for measuring finger ROM is goniometry. Flexion and extension of each individual joint is measured separately and the results are combined to give a cumulative score, the Total Active Range of Motion (TAROM) (2). Using goniometry has limitations as it requires training and is time-consuming. Conversely, the delta FTP is quick and easy to administer, providing a summation of flexion and extension of all three finger joints (MCP, PIP, DIP). Comparing the delta FTP and the FTP to the TAROM would provide evidence to support concurrent criterion validity.

Convergent construct validity examines the association between measures that should be statistically related as they reflect similar concepts. In this study, the correlation between the delta FTP (and the FTP) and the Hand Mobility in Scleroderma (HAMIS) test (18, 19), a performance based test of finger/hand motion, and the Duruoz Hand Index

(Rheumatoid Hand Function Disability Scale or Cochin Hand Scale) (20, 21), a self report of hand function, were examined to evaluate the convergent construct validity of the delta FTP. In addition, since disease severity markers in SSc, such as the Health Assessment Questionnaire Disability Index (HAQ-DI) and modified Rodnan skin score (mRSS), have correlated with the FTP in previous studies (15), we examined these and other disease severity markers in relation to the delta FTP.

In addition to truth (face, criterion and construct validity), a clinical outcome measure should demonstrate discrimination (reliability and sensitivity to change) and feasibility (practicality) (22). To evaluate reliability, both intra-rater (within rater) and inter-rater (between raters) reliability of the delta FTP were examined. To investigate the delta FTP’s sensitivity to change, the effect size and standardised response mean were evaluated in patients with serial delta FTP measurements and change in clinical status assessed by mRSS.

The goals of the present study were to evaluate the reliability and validity of the delta FTP in a cohort with SSc by studying its (i) intra-rater and inter-rater reliability, (ii) concurrent criterion and convergent construct validity, and to (iii) compare these psychometric properties with those of the traditional FTP.

Participants and methods

Participants

This study was completed in three phases. Intra-rater and inter-rater reliability were examined in Phase 1, concurrent criterion and convergent construct validity in Phase 2, and sensitivity to change in Phase 3. Subjects for Phases 1 and 2 were SSc patients with either diffuse cutaneous (dc) or limited cutaneous (lc) disease, while subjects for Phase 3 were all dcSSc patients with early disease (defined as <2 years from onset of skin thickening), recruited from the University of Pittsburgh Scleroderma Clinic during 2007–2008. The studies were approved by the University of Pittsburgh Institutional Review Board.

Patients were classified according their clinical phases of skin thickening as edematous, fibrotic or atrophic (23).

Arbitrary definitions for these phases were used, based on duration of skin thickening at the time of determination of the delta FTP: edematous <2 years; fibrotic 2–4 years; and atrophic >4 years.

Measures

FTP and Delta FTP measurements

The standard finger-to-palm (FTP) measurement is obtained using a ruler to measure the distance (in centimetres) between the tip of the pulp on the 3rd finger and the distal palmar crease while the patient attempts to make a full fist (maximal finger flexion at all 3 finger joints: MCP, PIP, and DIP). The delta FTP is finger extension, the distance between the 3rd fingertip and the distal palmar crease while the patient attempts full finger extension, minus the FTP. A more detailed description of the method used to measure FTP and delta FTP is included in Appendix 1.

The Duruoz Hand Index (DHI)

(also known as the Rheumatoid Hand Function Disability Scale or Cochin Hand Scale)

The DHI is a self-report questionnaire which measures hand functional ability (20) and is both reliable and valid in patients with SSc (5). The DHI asks how much difficulty the patient has performing 18 daily tasks of living in reference to hand function. Each item is scored from 0 (performed without difficulty) to 5 (impossible to do). A total score is obtained by adding the scores of all questions (range 0–90). A higher score indicates greater disability.

Health Assessment Questionnaire (HAQ)

The HAQ is a self-report questionnaire which focuses on 8 domains of function in ADLs: dressing/grooming, arising, eating, walking, hygiene, reach, grip and activity (24). Each domain is assigned a rank from 0 (can do without difficulty) to 3 (unable to do) and a composite disability index (DI) is calculated, with a higher score indicating greater disability. The HAQ has been shown to be a valid and reliable instrument measuring the impact of disease on functional status (14) and has been

validated in patients with SSc (25). Improvement of the HAQ-DI in SSc has been shown to reflect clinical improvement of disease status, correlating with physician global assessment (17).

Total Active Range of Motion (TAROM)
TAROM is a composite goniometric measurement of finger range of motion. The TAROM is calculated as the sum of simultaneous active MCP, PIP and DIP flexion in a fist position minus the sum of any extension deficits noted while all finger joints are in maximum extension. This measurement is commonly employed by occupational therapists (OTs) to evaluate deficits in overall hand range of motion and to document response to therapy (2). The TAROM of the 3rd finger was assessed by a single experienced OT (TT) during Phase 2 using a Roylean finger goniometer (Sammons Preston, USA), which permits measurement from 30 degrees of hyperextension to 120 degrees of flexion. The 3rd fingers of both hands were measured to permit a direct comparison with the delta FTP. The reliability of finger and hand goniometric measurements is extensively reported in the literature (2, 26-31). Therefore, we collected this measure only once to limit subject pain and fatigue.

Hand Mobility in Scleroderma (HAMIS)
The HAMIS is a 9-item performance-based test designed to examine the effects of SSc on hand mobility and hand function, focusing on detection of limitation of motion in performing tasks related to finger, wrist and forearm motions (18, 19). Each task is scored from 0 (no impairment) to 3 (cannot do), yielding a total possible score of 27 for each hand and 54 overall. Higher scores reflect poorer hand mobility. The HAMIS is a reliable and valid tool in assessing hand mobility in patients with SSc, and serves as an indicator of hand function (18, 19). The HAMIS was performed by a trained OT (MP) during Phase 2.

Modified Rodnan Skin Score (mRSS)
The mRSS is a clinical assessment of skin thickness by palpation at 17 surface anatomical sites, graded from 0

(normal) to 3 (severe thickening) with a maximum of 51 points (32). In dcSSc the mRSS has demonstrated convergent validity, correlating with the absolute value of, and changes in joint tenderness, oral aperture, handspread, FTP distance, disability and survival (11, 15, 33). An experienced physician (TAM) performed finger, hand, forearm and upper arm mRSS measurements during Phase 2.

Other variables

Visual Analogue Scales (VASs) were included to assess disease severity. Patients completed the following VASs to assess the impact of SSc during the previous week: pain, Raynaud phenomenon, finger ulcers, and overall disease severity (15). In previous SSc studies these scales have performed well and have correlated positively with the appropriate clinical measures (1, 15). VAS scales were scored continuously from 0 to 3, with 3 indicating greater pain or disability.

Methods

Phase 1. Intra-rater and inter-rater reliability

Two delta FTP and FTP measurements were performed by 2 physician examiners on 39 consecutive outpatients seen between November 2007 and January 2008. The measurements were taken approximately 30 minutes apart to minimize the effect of patient fatigue from repetitive finger motion.

Phase 2. Concurrent criterion and convergent construct validity

Seventeen SSc patients were recruited to participate in a one-day study. Patients with a broad range of delta FTP (0.9–8.9) measured at recent clinical visits were specifically recruited in order to test the validity of this measure across the spectrum of finger range of motion and disease severity. Patients completed DHI and HAQ questionnaires first to minimize the effect of using their hands during performance tests or their reports of pain. The order of the remaining tests was randomised to avoid possible fatigue or pain effects. The TAROM, HAMIS, mRSS, and delta FTP/FTP were each conduct-

ed in a separate room by a particular physician or OT.

Phase 3. Sensitivity to change

A retrospective analysis was performed on 24 patients with early dcSSc. Early dcSSc was defined as disease duration at time of initial visit <2 years from onset of skin thickening. This subgroup was targeted because these patients are the most likely to be entered into clinical trials of potential disease modifying treatments and show changes in skin score. Subjects were seen twice, and changes in disease status between clinical visits were documented by mRSS and concurrent delta FTP measurements.

Statistical analysis

Phase 1

Intraclass correlation coefficients (ICCs) with 95% confidence intervals were calculated to examine the intra-rater and inter-rater reliability of the delta FTP and FTP. Intraclass correlation coefficients of >0.80 were considered strong (34).

Phase 2

Spearman rho correlation coefficients were used to examine the concurrent criterion validity of the delta FTP and FTP with the TAROM, and to examine the convergent construct validity of the delta FTP and FTP with the HAMIS, DHI, and measures of disease severity – mRSS, HAQ DI and subsets, and patient-reported VAS. When appropriate, the left-hand or right-hand delta FTP and FTP measurements were used in comparison to measures that had separate right and left hand measurements or scores. A *p*-value of <0.05 was considered significant. Correlation coefficient values >0.70 were considered strong, 0.70–0.50 moderate, 0.49–0.30 fair, and <0.30 poor (35).

Phase 3

Responsiveness to change was evaluated using 2 statistical approaches, standardised response mean (SRM) and effect size (ES). The SRM is defined as the mean change in score between the baseline and follow-up visit divided by the standard deviation of the individual change in scores. A higher SRM indicates greater responsiveness.

A SRM >0.80 is considered large (36). ES is the ratio of observed change to a measure of variance. We used Cohen's d, in which the numerator is the mean change in the delta FTP from baseline visit to follow-up and the denominator is the standard deviation of delta FTP at baseline. Using Cohen's d, an ES of 0.20–0.49 represents a small change, 0.50–0.79 a medium change, and >0.80 a large change (37). In general an ES of 0.20 to 0.50 is considered relevant for detecting the minimally important difference (MID), the smallest change in a score which may influence clinical management (38).

All statistical procedures were performed with Stata software v.10 (State Corp., College Station, TX, USA) and SPSS v.16 (SPSS Inc., Chicago, IL, USA).

Results

Patients

Table I shows the baseline demographic and clinical characteristics of the subjects in Phases 1, 2 and 3. Overall, the groups are representative of SSc, with mean ages in the 50s, predominantly Caucasian and female. Phase 2 had a disproportionate number of males (42%) compared to clinical practice. Diffuse cutaneous (dcSSc) disease was more frequent in both Phase 1 and 2 groups (70% and 82% respectively), in part due to referral bias. Phase 3 specifically targeted dcSSc patients in the early phase of disease (<2 years from onset of skin thickening); therefore, all patients were in the edematous clinical phase of skin thickening. Other physical examination findings consistent with early disease were more notable in the Phase 3 subjects, such as swelling of fingers (63%), tendon friction rubs (46%), and higher mRSS (22.5+10.5). Phase 1, the reliability study, and Phase 2, the validity study, included patients from all three cutaneous phases, with 44% and 47% of the patients in the atrophic phase, respectively.

Concomitant medication use in each phase demonstrated less than 30% being on either NSAIDs or prednisone. The maximum prednisone dose was 10 mg per day at the time data were collected. DMARDs were used signifi-

Table I. Demographic and clinical characteristics of systemic sclerosis patients in study Phases 1, 2 and 3.

Characteristic	Phase 1 Reliability (n=39)	Phase 2 Validity (n=17)	Phase 3 Responsiveness to change (n=24)
Age (mean ± SD years)	50.0 ± 13.0	52.4 ± 11.2	53.05 ± 11.53
Sex (female)	77%	58%	80%
Ethnicity (Caucasian)	85%	94%	100%
SSc subtype			
diffuse	70%	82%	100%
limited	30%	18%	
Disease duration from onset of skin thickening (mean ± SD years) (median and IQR)	5.8 ± 5.6 2.9 (1.8 – 6.9)	8.5 ± 8.1 3.7 (2.2 – 13.1)	0.8 ± 0.4 0.8 (0.6 – 1.2)
Cutaneous phases*			
Edematous	11 (28%)	3 (18%)	24 (100%)
Fibrotic	11 (28%)	6 (35%)	0
Atrophic	17 (44%)	8 (47%)	0
Physical examination findings			
Finger joint swelling	0	0	0
Swelling of fingers	7 (18%)	7 (42%)	15 (63%)
Tendon friction rubs	3 (8%)	1 (6%)	11 (46%)
mRSS (total)	15.31 ± 12.6	12.7 ± 8.9	22.5 ± 10.48
Right FTP	2.9 ± 1.6	3.7 ± 1.4	3.2 ± 1.3
Right delta FTP	5.8 ± 2.8	5.7 ± 2.6	6.9 ± 2.3
Medications			
NSAIDs	5 (13%)	1 (6%)	5 (21%)
Prednisone [†]	8 (21%)	4 (24%)	7 (29%)
DMARDs [‡]	8 (21%)	11 (65%)	22 (91%)

*Clinical phases based on duration of skin thickening: edematous <2 years; fibrotic 2–4 years; atrophic >4 years.

[†]prednisone dose ranged from 2–10 mg per day for all 3 Phases.

[‡]DMARDs include d-penicillamine (85%), mycophenolate mofetil (13%), and cyclophosphamide, methotrexate or hydroxychloroquine (<5%).

Table II. Intra- and inter-rater reliability* of delta FTP and FTP[†] in 39 SSc patients in Phase 1.

	Delta FTP [†] ICC	FTP [†] ICC
Intra-rater agreement		
Left hand, rater 1	0.99 (0.98, 0.99)	0.99 (0.98, 0.99)
Left hand, rater 2	0.99 (0.99, 0.99)	0.99 (0.99, 0.99)
Right hand, rater 1	0.99 (0.98, 0.99)	0.98 (0.97, 0.99)
Right hand, rater 2	0.99 (0.99, 0.99)	0.99 (0.99, 0.99)
Inter-rater agreement		
Left hand, 1 st measurement	0.96 (0.91, 0.98)	0.92 (0.83, 0.96)
Left hand, 2 nd measurement	0.97 (0.93, 0.98)	0.95 (0.88, 0.97)
Right hand, 1 st measurement	0.97 (0.94, 0.98)	0.94 (0.89, 0.97)
Right hand, 2 nd measurement	0.97 (0.95, 0.99)	0.94 (0.90, 0.97)

*Intra-rater and inter-rater reliability assessed by intraclass correlation coefficient (ICC) and reported as ICC (95% confidence interval).

[†]FTP: finger-to-palm distance.

cantly more often in the Phase 3 group (early dcSSc), with 91% of the subjects on a DMARD compared to 21% of the subjects in Phase 1 and 65% in Phase 2. The most common DMARD prescribed in all patients combined was

d-penicillamine (85%), followed by mycophenolate mofetil (13%). The remainder of DMARDs were used in less than 5% of the patients and included cyclophosphamide, methotrexate and hydroxychloroquine.

Phase 1. Intra-rater and inter-rater reliability

The range of FTP was 0.1 to 7.9 cm, with a mean of 2.9 cm (± 1.6). The delta FTP ranged from 0.3 cm to 10.1 cm, with a mean of 5.8 cm (± 2.8). The intra-rater and inter-rater agreement was excellent for both the delta FTP and FTP measurements. ICCs ranged from 0.92 to 0.99 (Table II).

Phase 2. Validity

Table III shows the mean values \pm SD of the clinical variables analysed. Spearman correlation coefficients are summarised in Table IV. The delta FTP is expected to be inversely related to most outcome measures, resulting in a negative correlation coefficient. As the disease progresses, the third fingertip motion (extension minus flexion) is reduced and a smaller delta FTP results. In contrast, the HAQ DI, patient VAS, mRSS, DHI and HAMIS scores increase as finger impairment worsens and functional limitation increases. The TAROM is the only variable which should be positively correlated with the delta FTP. Conversely, the FTP should have a positive correlation with all variables except TAROM because the FTP typically increases as finger ROM decreases. The delta FTP demonstrated strong correlations with the TAROM and HAMIS, moderate correlations with the upper extremity mRSS, and fair correlations with the DHI and HAQ subsets of grip, dressing, eating, and hygiene (Table IV). Patient VAS variables showed little or no correlation with the delta FTP, with correlation coefficients ranging from 0.02 to 0.19 (data not shown). Correlations between the traditional FTP and the above outcome measures varied from little to strong correlations, paralleling the delta FTP.

Phase 3. Sensitivity to change

Over a mean period of 4.8 ± 3.4 months between office visits, the mean change in the mRSS was 5.25 ± 4.25 in these early dcSSc patients. As expected, mRSS increased with disease progression as the delta FTP decreased. The mean difference of the delta FTP between these visits was 1.0 cm for the left hand, 1.3 cm for the right hand, and

Table III. Mean values and SD for clinical variables in 17 SSc patients in Phase 2.

Variable	Overall (mean left and right) [†]	Left hand	Right hand
FTP (cm)	3.38 ± 1.23	3.05 ± 1.10	3.66 ± 1.44
Delta FTP (cm)	5.96 ± 2.38	6.23 ± 2.28	5.66 ± 2.58
TAROM (degrees)	172.18 ± 66.30	181.24 ± 73.94	163.12 ± 63.20
HAMIS (0 – 54; 0 – 27 R/L)	15.53 ± 11.20	7.65 ± 5.42	7.88 ± 5.88
mRSS upper extremity (0 – 30; 0 – 15 R/L)	14.82 ± 5.68	7.29 ± 2.49	7.53 ± 2.79
DHI total (0 – 90)	17.18 ± 16.65		
HAQ DI (0 – 3)	0.89 ± 0.69		
HAQ dressing (0 – 3)	0.88 ± 0.86		
HAQ eating (0 – 3)	0.76 ± 0.83		
HAQ hygiene (0 – 3)	1.06 ± 1.03		
HAQ grip (0 – 3)	1.29 ± 0.99		
Pain VAS (0 – 3)	0.81 ± 0.88		
Hand Pain VAS (0 – 3)	0.95 ± 0.91		
Raynaud VAS (0 – 3)	0.93 ± 0.71		
Digital Ulcer VAS (0 – 3)	0.64 ± 0.80		
Disease Severity VAS (0 – 3)	0.90 ± 0.79		

SD: Standard deviation; FTP: Finger-To-Palm distance; TAROM: Total Active Range of Motion; HAMIS: Hand Mobility in Scleroderma; DHI: Duruoz Hand Index; mRSS: modified Rodnan Skin Score; HAQ DI: Health Assessment Questionnaire Disability Index; VAS: Visual Analogue Scale.

[†]mean value of left hand and right hand for given variables.

Table IV. Correlation of variables with delta FTP and FTP in 17 SSc patients in Phase 2[†].

Variable	Delta FTP [‡] r_s (p-value)	FTP [‡] r_s (p-value)
TAROM	0.94*	-0.89*
HAMIS	-0.82*	0.70*
mRSS (upper extremity)	-0.68*	0.56 [‡]
DHI Total	-0.49 [‡]	0.47
HAQ DI (0 – 3)	-0.20	0.12
HAQ dressing (0 – 3)	-0.40	0.39
HAQ eating (0 – 3)	-0.49 [‡]	0.51 [‡]
HAQ hygiene (0 – 3)	-0.42	0.40
HAQ grip (0 – 3)	-0.30	0.20

[†] FTP: finger-to-palm distance.

[‡] Correlations of variables with the mean values of delta FTP and FTP between left and right hands are represented in the Table, individual hand correlations were similar.

TAROM: Total Active Range of Motion; HAMIS: Hand Mobility in Scleroderma; DHI: Duruoz Hand Index; mRSS: modified Rodnan Skin Score; HAQ DI: Health Assessment Questionnaire Disability Index.

* $p < 0.01$, [‡] $p < 0.05$.

1.1 cm for both hands combined. The effect size was calculated to be 0.47 for the left hand, 0.57 for the right hand, and 0.51 for mean delta FTP between hands. The SRMs were 1.02, 1.23 and 1.07 respectively.

Discussion

This study is the first to specifically analyse the psychometric properties of the FTP, a method commonly used to assess finger ROM in SSc. The FTP demonstrated sound reliability, concurrent criterion and convergent construct validity, making it a relatively valid

outcome measure. However, the FTP has the potential to misclassify a subgroup of patients with SSc whose fingers are “fixed” in flexion and have extremely limited finger ROM. This was demonstrated both in our preliminary data (see Introduction) and in some of the current study subjects. For example, of the 39 Phase 1 subjects, 10 had a delta FTP less than 4cm, implying severe impairment of motion, 2 of whom had a FTP < 2 cm, which is interpreted as minimal impairment (8). If judged by FTP alone, 2 (20%) of these 10 patients with the greatest impairment by

delta FTP would have been categorised as having minimal impairments. The misclassification of this subgroup of patients challenges the face validity of the FTP as a tool intended to measure finger motion in patients with SSc, threatening the overall validity of the FTP. Use of the delta FTP avoids such misclassification.

In addition to the face validity superiority, other lines of evidence support the integrity of the delta FTP as an outcome measure. It demonstrated excellent intra- and inter-rater reliability, with ICC values of 0.96 or greater, exceeding the reported reliability of the goniometer, the 'gold standard' method to measure finger ROM which has been reported to range from 0.58 for the DIP joint to 0.83 for the MCP joint in healthy subjects (39) and from 0.61 to 0.96 in patients with RA (40). These results reflect an average variation of finger joint measurement of 3 to 5 degrees for a single observer and 5 to 7 degrees between observers using a goniometer (2, 29-31, 41). The delta FTP appears to provide a more reliable means of measuring finger ROM than goniometric measurement. The reliability of the delta FTP was not influenced by the clinical phase of skin thickening. Repeatability (both intra- and inter-rater) of the delta FTP was similar in those with very atrophic or 'claw like' hands compared to those with edematous or fibrotic hands.

Strong correlations between the TAR-OM measured by goniometry and the delta FTP support concurrent criterion validity of the delta FTP in subjects in all three cutaneous phases of SSc (edematous, fibrotic, and atrophic). Although it was unnecessary to test construct validity of the delta FTP given the presence of a gold standard, we investigated the delta FTP's relationship to both a hand performance test (HAMIS) and a self-report questionnaire of hand function (DHI). We hypothesised that if the delta FTP had good convergent construct validity, there would be a large correlation between the HAMIS and the delta FTP, because they both include measures of finger ROM, and a smaller correlation between the DHI and the delta FTP, because the DHI

serves as an overall assessment of hand functional ability. The data supported these hypotheses and the convergent construct validity of the delta FTP. The correlation between the HAMIS and the delta FTP was strong, while the correlation between the DHI and delta FTP was fair, supporting the notion that the DHI partially indicates the concept of finger ROM but is influenced by other factors, such as dexterity and strength of fingers/hands and the patient's adaptations to the impairments. Similar results were seen for the HAQ subcomponents associated with hand function (dressing, eating, hygiene and grip), also supporting convergent construct validity.

To evaluate the delta FTP's ability to assess disease severity, it was compared to both the mRSS and patient VASs. The mRSS is a more objective measurement of disease severity in dcSSc, with certain total skin scores associated with mild, moderate or severe disease (8), while the VAS is a subjective patient-reported measure of how patients consider that their disease has affected their function over the past week. This may explain the delta FTP's moderate correlations with the mRSS and poor correlations with the patient-reported VAS. The relationship between the delta FTP and the mRSS support the delta FTP as a measure reflecting disease severity.

This relationship was further evaluated in Phase 3, in which early dcSSc patients with change in disease status, reflected by change in mRSS, were evaluated for change in delta FTP. The delta FTP demonstrated both a moderate effect size and large standardised response mean in this group of subjects with early dcSSc (<2 years from onset of skin thickening), supporting it as a sensitive measure to detect change in clinical status, which is an important characteristic of an outcome measure. High responsiveness indicates that fewer subjects are needed to demonstrate a significant difference.

An additional criterion to justify an outcome measure as valid is feasibility. The delta FTP is a highly feasible instrument both clinically and as a research measure. This is apparent in the

speed with which the measurement can be taken (less than 2 minutes), the simplicity of the equipment and training, and the ease of interpretation.

The possible contributors to finger joint contractures in SSc are (1) skin and subcutaneous tissue edema and thickening; (2) tendonitis with fibrosis leading to tendon shortening; and (3) arthropathy with synovitis and joint damage. Skin thickening is definitely related to contractures (see Phase 2, negative correlation between delta FTP and finger skin thickening [-0.68, $p < 0.01$]). Based on the high frequency of palpable tendon friction rubs in early dcSSc (540/909 or 59% in our Scleroderma Databank), we believe that tendon shortening also plays an important role (42). None of the study patients had finger joint swelling on physical examination (Table I). We do not have adequate data on hand radiographs performed simultaneously with delta FTP measurements. However, among 1264 Databank SSc patients with hand radiographs available, only 208 (16%) had either joint space narrowing or erosions. For these reasons, we do not believe that arthritis is the primary cause of SSc joint contractures.

Limitations

In Phase 1 repeated delta FTP and FTP measurements were taken 30 minutes apart to minimise patient fatigue at the expense of recall bias, which has the potential to inflate intra-rater ICC. However, if intra-rater bias is present, it likely has had only a minimal effect given the very strong inter-rater ICC. In Phase 2 the order of the performance tests was randomized to prevent fatigue/pain effects on subjects. However, the participant sample was too small to statistically analyse the data to see if this method was effective. In Phase 2 repeated measurements to evaluate intra-rater ICC for goniometry could have been performed. We omitted this additional step to limit patient burden and because the reliability of goniometry is well established in a trained user. Phase 2 may have selection bias because in general, healthier patients (less severe disease) are more willing to participate in studies compared with the

sequentially examined clinic patients in Phase 1. However, it is notable that 5 of the 17 subjects (29%) had a delta FTP less than 4, indicating severely limited ROM which is usually associated with more severe disease. Phase 3 is a retrospective chart review, and thus only provides preliminary data supporting sensitivity to change. A large prospective study in early dcSSc will be required to confirm these findings.

In conclusion, the delta FTP is a valid outcome measure assessing finger ROM in SSc patients of all cutaneous disease phases. It satisfies the criteria of a good outcome measure, including reliability (intra- and inter-rater), validity (face, concurrent criterion, and convergent construct), sensitivity to change (responsiveness) and feasibility. Although the traditional FTP is a partially validated outcome measure of finger ROM in SSc, its tendency to misclassify impairment is a major concern, challenging its face validity. The delta FTP does not carry this potential to misclassify patients, and with its other psychometric properties determined in this study, is a better outcome measure than the FTP. Sensitivity to change was studied and demonstrated only in early SSc patients. This is precisely the subset which is considered ideal for enrolment in clinical trials of potential disease-modifying drugs (43).

Acknowledgements

We thank Marie Pace, OTR/L, CHT and Terry Taylor OTR/L, CHT for their assistance with the goniometric and other occupational therapy maneuvers, and to several staff members of the University of Pittsburgh Division of Rheumatology and Clinical Immunology for volunteering their time to help conduct the one-day study (Phase 2).

References

- MEDSGER TA JR: Natural history of systemic sclerosis and the assessment of disease activity, severity, functional status, and psychologic well-being. *Rheum Dis Clin North Am* 2003; 29: 255-73, vi.
- HUNTER JM, MACKIN E, CALLAHAN AD: *Rehabilitation of the hand: surgery and therapy*. 4th ed. St. Louis: Mosby; 1995.
- POOLE JL, GALLEGOS M, O'LINC S: Reliability and validity of the Arthritis Hand Function Test in adults with systemic sclerosis (scleroderma). *Arthritis Care Res* 2000; 13: 69-73.
- 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.
- BROWER LM, POOLE JL: Reliability and validity of the Duruoz Hand Index in persons with systemic sclerosis (scleroderma). *Arthritis Rheum* 2004; 51: 805-9.
- GUCCIONE AA: Arthritis and the process of disablement. *Phys Ther* 1994; 74: 408-14.
- RANNOU F, POIRAUDEAU S, BEREZNE A *et al.*: 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. *Arthritis Rheum* 2007; 57: 94-102.
- MEDSGER TA JR, BOMBARDIERI S, CZIRJAK L, SCORZA R, DELLA ROSSA A, BENCIVELLI W: Assessment of disease severity and prognosis. *Clin Exp Rheumatol* 2003; 21 (Suppl. 29): S42-6.
- CLEMENTS PJ, FURST DE, WONG WK *et al.*: High-dose versus low-dose D-penicillamine in early diffuse systemic sclerosis: analysis of a two-year, double-blind, randomized, controlled clinical trial. *Arthritis Rheum* 1999; 42: 1194-203.
- FURST DE, CLEMENTS PJ, HARRIS R, ROSS M, LEVY J, PAULUS HE: Measurement of clinical change in progressive systemic sclerosis: a 1 year double-blind placebo-controlled trial of N-acetylcysteine. *Ann Rheum Dis* 1979; 38: 356-61.
- MERKEL PA, CLEMENTS PJ, REVEILLE JD, SUAREZ-ALMAZOR ME, VALENTINI G, FURST DE: Current status of outcome measure development for clinical trials in systemic sclerosis. Report from OMERACT 6. *J Rheumatol* 2003; 30: 1630-47.
- CARMINES EG ZR: *Reliability and Validity Assessment*. Beverly Hills, CA: Sage; 1979.
- MCDOWELL I, NEWELL C: *Measuring health: a guide to rating scales and questionnaires*. 2nd ed. New York: Oxford University Press; 1996.
- MERKEL PA, HERLYN K, MARTIN RW *et al.*: Measuring disease activity and functional status in patients with scleroderma and Raynaud's phenomenon. *Arthritis Rheum* 2002; 46: 2410-20.
- 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 *et al.*: Correlates of the disability index of the health assessment questionnaire: a measure of functional impairment in systemic sclerosis. *Arthritis Rheum* 1999; 42: 2372-80.
- LAWRENCE E, POPE J, AL ZAHRALY Z, LALANI S, BARON M: The relationship between changes in self-reported disability (measured by the Health Assessment Questionnaire - HAQ) in scleroderma and improvement of disease status in clinical practice. *Clin Exp Rheumatol* 2009; 27 (Suppl. 54): 32-7.
- 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 *et al.*: Development and validation of a rheumatoid hand functional disability scale that assesses functional handicap. *J Rheumatol* 1996; 23: 1167-72.
- POIRAUDEAU S, CHEVALIER X, CONROZIER T *et al.*: Reliability, validity, and sensitivity to change of the Cochin hand functional disability scale in hand osteoarthritis. *Osteoarthritis Cartilage* 2001; 9: 570-7.
- BOERS M, BROOKS P, STRAND CV, TUGWELL P: The OMERACT filter for Outcome Measures in Rheumatology. *J Rheumatol* 1998; 25: 198-9.
- CLEMENTS PJ, FURST DE: *Systemic sclerosis*. 2nd ed. Philadelphia: Lipincott Williams & Wilkins; 2004.
- FRIES JF, SPITZ P, KRAINES RG, HOLMAN HR: Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980; 23: 137-45.
- 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.
- HAMILTON GF, LACHENBRUCH PA: Reliability of goniometers in assessing finger joint angle. *Phys Ther* 1969; 49: 465-9.
- BOONE DC, AZEN SP, LIN CM, SPENCE C, BARON C, LEE L: Reliability of goniometric measurements. *Phys Ther* 1978; 58: 1355-90.
- GAJDOSIK RL, BOHANNON RW: Clinical measurement of range of motion. Review of goniometry emphasizing reliability and validity. *Phys Ther* 1987; 67: 1867-72.
- ELLIS B, BRUTON A: A study to compare the reliability of composite finger flexion with goniometry for measurement of range of motion in the hand. *Clin Rehabil* 2002; 16: 562-70.
- COOK JR, BAKER NA, CHAM R, HALE E, REDFERN MS: Measurements of wrist and finger postures: a comparison of goniometric and motion capture techniques. *J Appl Biomech* 2007; 23: 70-8.
- MAYERSON NH, MILANO RA: Goniometric measurement reliability in physical medicine. *Arch Phys Med Rehabil* 1984; 65: 92-4.
- CLEMENTS PJ, LACHENBRUCH P, SIEBOLD J *et al.*: Inter and intraobserver variability of total skin thickness score (modified Rodnan TSS) in systemic sclerosis. *J Rheumatol* 1995; 22: 1281-5.
- STEEN VD, MEDSGER TA JR: Improvement in skin thickening in systemic sclerosis associated with improved survival. *Arthritis Rheum* 2001; 44: 2828-35.
- SHROUT PE: Intraclass Correlations. *Psychological Bulletin* 1979.
- FERMANIAN J: [Measurement of agreement between 2 judges. Qualitative cases]. *Rev Epidemiol Sante Publique* 1984; 32: 140-7.
- HUSTED JA, COOK RJ, FAREWELL VT, GLADMAN DD: Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol* 2000; 53: 459-68.

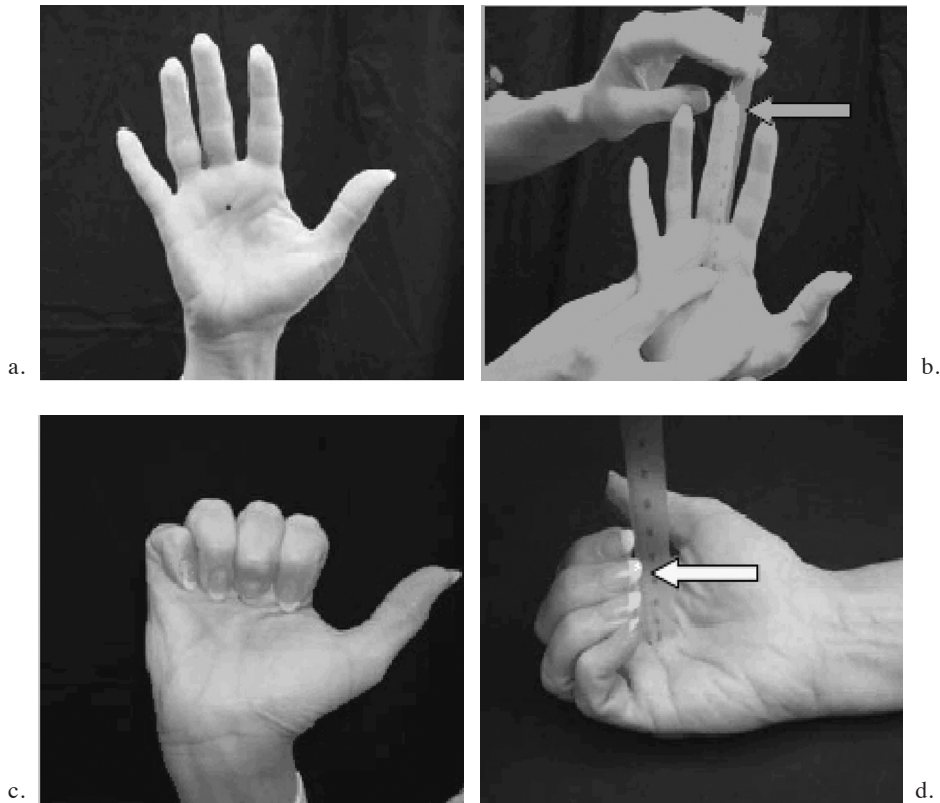
37. COHEN J: A Power Primer. *Psychological Bulletin* 1992; 112: 155-9.
38. SLOAN JA, CELLA D, HAYS RD: Clinical significance of patient-reported questionnaire data: another step toward consensus. *J Clin Epidemiol* 2005; 58: 1217-9.
39. STAM HJ, ARDON MS, DEN OUDEN AC, SCHREUDERS TA, ROEBROECK ME: The compangle: a new goniometer for joint angle measurements of the hand. A technical note. *Eura Medicophys* 2006; 42: 37-40.
40. LEFEVRE-COLAU MM: Reliability of two goniometers in assessing rheumatoid finger mobility. *Eur Med Phys* 2001; 37: 7.
41. HELLEBRANDT FA, DUVAL EN *et al.*: The measurement of joint motion. Part II. Reliability of goniometry. *Phys Ther Rev* 1949; 29: 302.
42. STEEN VD, MEDSGER TA JR: The palpable tendon friction rub: an important physical examination finding in patients with systemic sclerosis. *Arthritis Rheum* 1997; 40: 1146-51.
43. WHITE B, BAUER EA, GOLDSMITH LA *et al.*: Guidelines for clinical trials in systemic sclerosis (scleroderma). I. Disease-modifying interventions. *Arthritis Rheum* 1995; 38: 351-60.

Appendix

Measurement of fingertip to palm (FTP) distances in systemic sclerosis

PREPARATION:

1. Use a narrow (20cm) centimeter plastic ruler.
2. Place a dot with a ballpoint pen on the distal palmar crease proximal to the third finger.
3. Ask the patient to 'loosen up' by flexing and extending his/her fingers several times.



MEASUREMENTS:

FTP in Extension (a and b)

1. Ask the patient to completely extend his/her fingers.
2. Measure and record to the nearest millimeter the distance from the distal palmar crease dot to the end of the third finger (skin, not end of nail).
3. Measure and record both right and left hands.

FTP in Flexion (c and d)

1. Ask the patient to completely flex his/her fingers by trying to touch the tip of the third finger to the ink dot. The goal is to bend the fingers at all three joints (MCP, PIP, DIP).
2. Measure and record to the nearest millimeter the distance from the distal palmar crease dot to the tip of the third finger.
3. Measure and record both right and left hands.

Extension Minus Flexion, the 'Delta FTP' Measurement:

To obtain the difference between FTP Extension and Flexion, subtract the flexion measurement from the extension measurement and record in millimeters.