

Reliability and Validity of the Tender and Swollen Joint Counts and the Modified Rodnan Skin Score in Early Diffuse Cutaneous Systemic Sclerosis: Analysis from the Prospective Registry of Early Systemic Sclerosis Cohort

Jessica K. Gordon, Gandikota Girish, Veronica J. Berrocal, Meng Zhang, Christopher Hatzis, Shervin Assassi, Elana J. Bernstein, Robyn T. Domsic, Faye N. Hant, Monique Hinchcliff, Elena Schiopu, Virginia D. Steen, Tracy M. Frech, and Dinesh Khanna

ABSTRACT. Objective. To determine the inter/intraobserver reliability of the tender and swollen joint counts (TJC, SJC) and the modified Rodnan Skin Score (mRSS) in diffuse cutaneous systemic sclerosis (dcSSc) and to assess content validity of the TJC/SJC.

Methods. Ten rheumatologists completed the SJC, TJC, and mRSS on 7 patients. Musculoskeletal ultrasound (MSUS) was performed.

Results. Interobserver and intraobserver reliability for the TJC was 0.97 and 0.99, for the SJC was 0.24 and 0.71, and for the mRSS was 0.81 and 0.94, respectively. MSUS abnormalities did not correspond with SJC/TJC.

Conclusion. We demonstrate excellent inter- and intraobserver reliability for the mRSS and TJC in dcSSc. However, the SJC and TJC did not correspond to MSUS. (J Rheumatol First Release March 15 2017; doi:10.3899/jrheum.160654)

Key Indexing Terms:

SYSTEMIC SCLEROSIS
ULTRASONOGRAPHY

OUTCOME MEASURES
DIAGNOSTIC IMAGING

INFECTIONS AND ARTHRITIS
JOINT COUNT

Arthritis and tendinopathy are manifestations of systemic sclerosis (SSc) that affect up to 97% of patients^{1,2} and contribute to disability and impaired quality of life³. Synovitis, joint contractures, and tendon friction rubs are associated with disease severity and progression in early SSc^{4,5}. Despite the high prevalence and morbidity of arthropathy in SSc, its treatment has not been studied as a primary endpoint in randomized controlled trials, and arthritis-specific outcome measures have not been validated.

The tender joint count (TJC) and swollen joint count (SJC) are core outcome measures used to assess disease activity in rheumatoid arthritis (RA) in the context of clinical trials and observational studies⁶. Adequate interobserver and intraobserver reliability has been demonstrated for the 28-joint count in RA⁷. However, the 28-joint count can miss disease activity in psoriatic arthritis^{8,9} or may overinterpret disease activity when fibromyalgia coexists with RA¹⁰.

The TJC and SJC are used as outcome measures in SSc

From the Department of Rheumatology, and Department of Epidemiology and Biostatistics, Hospital for Special Surgery; Department of Rheumatology, Columbia University, New York, New York; Department of Radiology, and Department of Biostatistics, University of Michigan; University of Michigan Scleroderma Program, Ann Arbor, Michigan; Department of Rheumatology, University of Texas, Houston, Texas; Department of Rheumatology, University of Pittsburgh, Pittsburgh, Pennsylvania; Department of Rheumatology, Medical University of South Carolina, Charleston, South Carolina; Department of Rheumatology, Northwestern University, Chicago, Illinois; Department of Rheumatology, Georgetown University, Washington, DC; Department of Rheumatology, University of Utah, Salt Lake City, Utah, USA.

Supported by the University of Michigan Scleroderma Program. The Prospective Registry of Early Systemic Sclerosis (PRESS) registry is supported by a US National Institutes of Health/National Institute of Arthritis and Musculoskeletal and Skin Diseases K24 grant (AR063120 to D. Khanna), and is funded in part by a Scleroderma Foundation SCORE grant, although the SCORE grant did not fund this substudy of PRESS. D. Khanna has received funding from Genentech/Roche and Bayer.

J.K. Gordon, MD, MSc, Department of Rheumatology, Hospital for Special Surgery; G. Girish, MBBS, Department of Radiology, University of Michigan; V.J. Berrocal, MSc, PhD, Department of Biostatistics, University of Michigan; M. Zhang, PhD, Department of Epidemiology and Biostatistics, Hospital for Special Surgery; C. Hatzis, BA, Department of Rheumatology, Hospital for Special Surgery; S. Assassi, MD, MS, Department of Rheumatology, University of Texas; E.J. Bernstein, MD, MSc, Department of Rheumatology, Columbia University; R.T. Domsic, MD, MPH, Department of Rheumatology, University of Pittsburgh; F.N. Hant, DO, Department of Rheumatology, Medical University of South Carolina; M. Hinchcliff, MD, MS, Department of Rheumatology, Northwestern University; E. Schiopu, MD, University of Michigan Scleroderma Program; V.D. Steen, MD, Department of Rheumatology, Georgetown University; T.M. Frech, MD, MS, Department of Rheumatology, University of Utah; D. Khanna, MD, MS, University of Michigan Scleroderma Program.

Address correspondence to Dr. J.K. Gordon, Hospital for Special Surgery, New York, New York 10021, USA. E-mail: gordonj@hss.edu

Accepted for publication February 3, 2017.

clinical studies because they appear to have high face and content validity^{11,12,13}. Our objectives were to evaluate inter-observer and intraobserver reliability of the 28-joint TJC and SJC, and to assess the content validity of the TJC and SJC through comparison to musculoskeletal ultrasound (MSUS) in patients with early diffuse cutaneous SSc (dcSSc) because this has not been previously studied. We additionally assessed the interobserver and intraobserver reliability for the modified Rodnan Skin Score (mRSS) in this context.

MATERIALS AND METHODS

Study design. A cross-sectional study was conducted at the University of Michigan (UM) on 1 day in December 2014. Seven patients with early dcSSc were included from the Prospective Registry of Early Systemic Sclerosis (PRESS) cohort, a multicenter registry with participation of 12 US SSc centers. The protocol was approved by the Institutional Review Board (IRB) at UM (ID HUM00059408) and patients provided written consent prior to enrollment. Ten rheumatologists from PRESS centers participated in this exercise.

Training and examinations. The rheumatologists underwent a training session by DK to perform the TJC and SJC. Assessment occurred at 2 separate times on the same day for 5 subjects. One patient underwent the 10 initial examinations and 1 repeat examination. One patient underwent 6 initial examinations because of scheduling conflicts. These patients were excluded from intrarater analysis. To prevent bias, each patient was examined separately, underwent MSUS, and then underwent repeat examinations.

Bilateral US. Greyscale and power Doppler MSUS of the bilateral hands and wrists was performed using high-frequency linear array transducers (12–17 MHz) on clinical US machines (LOGIQ 9, GE Healthcare; and iU22, Philips Healthcare). Static and dynamic images of flexor and extensor tendons at the wrist, distal radioulnar, and metacarpophalangeal (MCP) and interphalangeal joints were obtained. Presence of joint effusion, synovitis, tenosynovitis, tendinosis, and erosions were documented. The images were interpreted by the same musculoskeletal radiologist (GG) to reduce variability.

Statistical analysis. Interrater and intrarater reliability were computed treating the TJC, SJC, and mRSS as continuous variables and by fitting a linear mixed model to the examiners' ratings with random effects for patient, rater, and patient by rater. This methodology was chosen over the κ statistic because of the repeated characteristic of the data. We compared the first SJC/TJC to MSUS of the hands and wrists (22 joints) using descriptive statistics. We calculated sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for each rheumatologist's first examination versus MSUS.

RESULTS

The mean (SD) age of the 7 subjects was 41.6 years (19.8) and the mean (SD) disease duration from the first non-Raynaud symptom was 2.7 years (0.8). Three patients were women. Two patients (28.6%) had the RNA polymerase 3 antibody and 1 (14.2%) had the Scl-70 antibody. The mean (SD) mRSS at time of entry into the PRESS Registry was 21.1 (9.6). None had clinical overlap or RA, but rheumatoid factor and cyclic citrullinated peptide antibodies (CCP) were not tested.

TJC scores. The mean (SD) TJC28 for the patients was 4.2 (2.0). Interobserver reliability for the TJC28 was 0.97, while intraobserver reliability was 0.99. A second analysis was

performed limited to the 22 joints of the bilateral hands and wrists assessed by MSUS. The 22-joint TJC showed an interrater reliability of 0.97 and intrarater reliability of 0.99.

SJC scores. The mean (SD) SJC for the patients was 1.3 (0.8). Interobserver reliability for the SJC was 0.24, and intraobserver reliability for the SJC was 0.71. The 22-joint SJC showed an interrater reliability of 0.17 and intrarater reliability of 0.65.

Results of mRSS. The interobserver reliability for the mRSS was 0.81, and the intraobserver reliability for the mRSS was 0.94. Additionally, the interobserver mean for the mRSS was 14.67 with a within-patient SD of 4.04. The intraobserver mean for the mRSS was 15.04 while the within-patient SD was 2.30.

Comparison of the TJC and SJC versus MSUS. Seven patients underwent MSUS of the hands and wrists (22 joints) resulting in an US evaluation of 154 joints. Five patients (71.4%) had at least 1 joint with synovial thickening or synovitis on MSUS. Only 9.7% (15/154) of joints on MSUS showed synovitis or synovial thickening. No erosions were observed and only wrists and MCP showed any abnormality. There were 1452 individual assessments of each joint for swelling and tenderness from the first round of examination limited to hands and wrists (Figure 1). Only 4.1% (60/1452) of examinations were positive for swelling and 20.4% (296/1452) were positive for tenderness; 2% (3/150) of investigators' examinations of joints with synovial thickening or synovitis on MSUS identified swelling, and 9.3% (14/150) identified tenderness (Table 1). The mean (SD) sensitivity for the identification of a joint as swollen using the MSUS as a gold standard was 0.02 (0.05) and the mean specificity was 0.96 (0.03). The mean (SD) PPV was 0.04 (0.08) and the NPV was 0.89 (0.01) for swelling. The mean sensitivity for the identification of a joint as tender was 0.09 (0.03) and the specificity for the TJC was 0.77 (0.03). The mean (SD) PPV was 0.05 (0.01) and the NPV was 0.88 (0.01) for tenderness. For individual patient joints noted to be tender, the median proportion of investigators noting tenderness was 70%. For swollen joints, the median proportion of investigators noting swelling was just 10%. The median proportion of assessors noting tenderness and swelling concurrently in a joint was 0 (range 0%–30%). MSUS was used to assess for tendinopathy, osteoarthritis, and skin thickness, but did not show alternative explanations for swelling or tenderness.

DISCUSSION

Improved treatment of musculoskeletal manifestations of SSc is an important unmet need in the care of patients with SSc. Presently, to our knowledge, there are no validated outcome measures that objectively assess arthritis in this group. Some clinical trials (abatacept: NCT02161406, tocilizumab: NCT01532869) are studying the effect of biologics approved for RA in SSc, and the TJC and SJC are used as analytical outcomes. However, SSc arthritis is distinct from RA, and

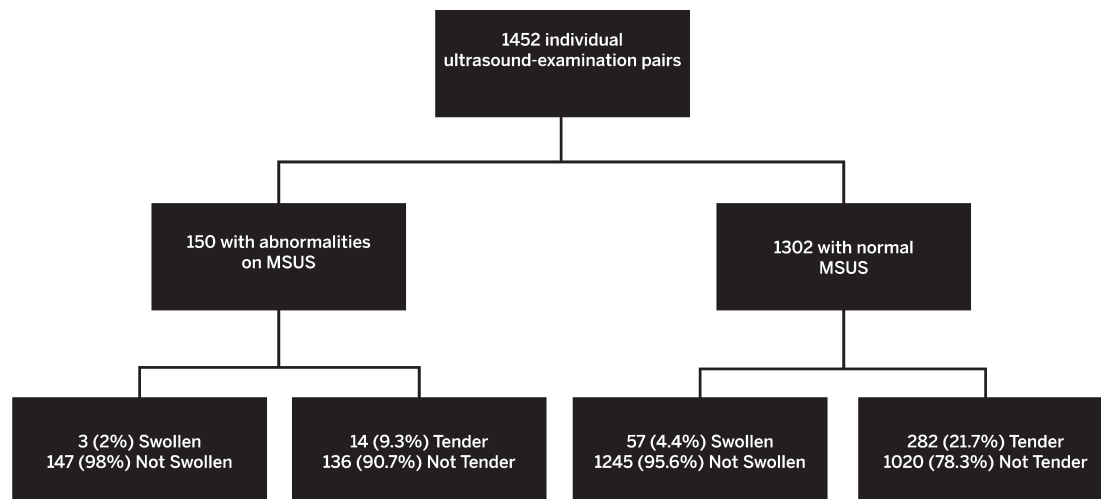


Figure 1. Flow chart of comparison of US and physical examination findings. Seven patients underwent MSUS of the bilateral hands and wrists (22 joints), resulting in 154 discrete US evaluations, which are treated as a gold standard in this exercise. Ten examiners rated each joint for swelling and tenderness, resulting in 1540 potential US examination pairs. Of this, 1452 US physical examination pairs occurred. The number and percentage of those that were noted as tender or swollen are shown. US: ultrasound; MSUS: musculoskeletal ultrasound.

Table 1. Summary of abnormalities noted on US and on clinical examination with proportions of agreement between investigators and US. A joint-by-joint comparison of US versus clinical examination is presented. Values are n/N (%) unless otherwise specified.

Joint Group	Wrist	MCP	PIP	Total
Total no. examined by MSUS	14	70	70	154
Total no. clinical examinations	132	660	660	1452
Joints with synovitis or synovial thickening on MSUS	6/14 (42.8)	9/70 (12.9)	0/70 (0)	15/154 (9.7)
Clinical examinations identifying swelling	6/132 (4.6)	14/660 (2.2)	40/660 (6.1)	60/1452 (4.1)
Clinical examinations identifying tenderness	29/132 (22.0)	117/660 (17.8)	150/660 (22.8)	296/1452 (20.4)
Proportion of clinical examinations of the abnormal joints on MSUS identifying the joint as swollen	3/60 (5.0)	0/90 (0)	—	3/150 (2.0)
Proportion of clinical examinations of the abnormal joints on MSUS identifying the joint as tender	14/60 (23.4)	0/90 (0)	—	14/150 (9.3)
Proportion of clinical examinations of the normal joints on MSUS identifying the joint as not swollen	69/72 (95.9)	556/570 (97.6)	620/660 (94.0)	1245/1302 (95.6)
Proportion of clinical examinations of the normal joints on MSUS identifying the joint as not tender	57/72 (79.2)	453/570 (79.5)	510/660 (77.3)	1020/1302 (78.3)

MSUS was performed on the 22 joints of the bilateral hands and wrists of 7 patients. Ten physician investigators examined each of these joints, with some exclusions for missing data or for ulcers overlying a joint. US: ultrasound; MSUS: musculoskeletal US; MCP: metacarpophalangeal joints; PIP: proximal interphalangeal joints.

potential outcome measures for SSc-related arthritis must be validated specifically for SSc. To our knowledge, ours is the first study to assess the reliability and validity of the 28-joint SJC and TJC in early dcSSc, with patients represented in clinical trials primarily assessing skin disease¹⁴.

In our study, we found near-perfect interrater and intrarater reliability for the TJC, demonstrating good accuracy for this measurement in early dcSSc. The SJC, on the other hand, had only a fair interrater reliability, but substantial intrarater reliability. To provide content validation of the TJC and SJC, we performed a joint-by-joint assessment using MSUS and noted poor agreement between the physical examination findings of individual joints with the MSUS findings. There was a low

sensitivity and PPV, but high specificity and NPV. Discordance between clinical examination and US findings has also been observed in RA¹⁵, and US can add additional information to the physical examination¹⁶. Reliability of the mRSS was confirmatory of historical results¹⁷.

Our study had both strengths and weaknesses. All subjects were examined on the same day, minimizing variability. We specifically recruited patients with early dcSSc not enriched for those with arthritis to assess the TJC and SJC in the very population that would be recruited for dcSSc clinical trials. However, this resulted in subjects with low mean TJC, SJC, and MSUS abnormalities, which contributed to the low PPV. Agreement was high regarding negative findings. A similar

study should be undertaken in patients with SSc with clinical evidence of arthritis and should include assessment of inflammatory markers and anti-CCP status, which were limitations of our present study. MSUS was used as a gold standard; however, magnetic resonance imaging may be better suited for defining joint disease in this population¹⁸. Future study should also include incorporation of multiple MSK radiologists to better ascertain agreement in this population.

Current evidence in our pilot study does not support the use of the TJC and SJC as outcome measures in an early dcSSc population not selected for arthritis. Further attention should focus on the development of outcome measures for the MSK manifestations of SSc and validation of SJC/TJC in patients with SSc and arthritis.

REFERENCES

1. Avouac J, Clements PJ, Khanna D, Furst DE, Allanore Y. Articular involvement in systemic sclerosis. *Rheumatology* 2012;51:1347-56.
2. Baron M, Lee P, Keystone EC. The articular manifestations of progressive systemic sclerosis (scleroderma). *Ann Rheum Dis* 1982;41:147-52.
3. 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.
4. Avouac J, Walker U, Tyndall A, Kahan A, Matucci-Cerinic M, Allanore Y; EUSTAR. Characteristics of joint involvement and relationships with systemic inflammation in systemic sclerosis: results from the EULAR Scleroderma Trial and Research Group (EUSTAR) database. *J Rheumatol* 2010;37:1488-501.
5. Avouac J, Walker UA, Hachulla E, Riemekasten G, Cuomo G, Carreira PE, et al; EUSTAR collaborators; EUSTAR collaborators. Joint and tendon involvement predict disease progression in systemic sclerosis: a EUSTAR prospective study. *Ann Rheum Dis* 2016;75:103-9.
6. Scott IC, Scott DL. Joint counts in inflammatory arthritis. *Clin Exp Rheumatol* 2014;32 (5 Suppl 85):S-7-12.
7. Smolen JS, Breedveld FC, Eberl G, Jones I, Leeming M, Wylie GL, et al. Validity and reliability of the twenty-eight-joint count for the assessment of rheumatoid arthritis activity. *Arthritis Rheum* 1995;38:38-43.
8. Leeb BF, Andel I, Sautner J, Fassl C, Nothnagl T, Rintelen B. The Disease Activity Score in 28 joints in rheumatoid arthritis and psoriatic arthritis patients. *Arthritis Rheumatol* 2007;57:256-60.
9. Coates LC, FitzGerald O, Gladman DD, McHugh N, Mease P, Strand V, et al. Reduced joint counts misclassify patients with oligoarticular psoriatic arthritis and miss significant numbers of patients with active disease. *Arthritis Rheum* 2013;65:1504-9.
10. Pollard LC, Kingsley GH, Choy EH, Scott DL. Fibromyalgic rheumatoid arthritis and disease assessment. *Rheumatology* 2010;49:924-8.
11. Clements PJ, Allanore Y, Khanna D, Singh M, Furst DE. Arthritis in systemic sclerosis: systematic review of the literature and suggestions for the performance of future clinical trials in systemic sclerosis arthritis. *Semin Arthritis Rheum* 2012;41:801-14.
12. Khanna D, Lovell DJ, Giannini E, Clements PJ, Merkel PA, Seibold JR, et al; Scleroderma Clinical Trials Consortium co-authors. Development of a provisional core set of response measures for clinical trials of systemic sclerosis. *Ann Rheum Dis* 2008;67:703-9.
13. Khanna PP, Furst DE, Clements PJ, Maranian P, Indulkar L, Khanna D; D-Penicillamine Investigators. Tendon friction rubs in early diffuse systemic sclerosis: prevalence, characteristics and longitudinal changes in a randomized controlled trial. *Rheumatology* 2010;49:955-9.
14. Khanna D, Denton CP, Jähreis A, van Laar JM, Frech TM, Anderson ME, et al. Safety and efficacy of subcutaneous tocilizumab in adults with systemic sclerosis (faSScinate): a phase 2, randomised, controlled trial. *Lancet* 2016;387:2630-40.
15. Garrigues F, Jousse-Joulin S, Bouttier R, Nonent M, Bressollette L, Sarau A. Concordance between clinical and ultrasound findings in rheumatoid arthritis. *Joint Bone Spine* 2013;80:597-603.
16. Murayama G, Ogasawara M, Nemoto T, Yamada Y, Ando S, Minowa K, et al. Clinical miscount of involved joints denotes the need for ultrasound complementation in usual practice for patients with rheumatoid arthritis. *Clin Exp Rheumatol* 2013;31:506-14.
17. 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.
18. Abdel-Magied RA, Lotfi A, AbdelGawad EA. Magnetic resonance imaging versus musculoskeletal ultrasonography in detecting inflammatory arthropathy in systemic sclerosis patients with hand arthralgia. *Rheumatol Int* 2013;33:1961-6.