SARCOIDOSIS VASCULITIS AND DIFFUSE LUNG DISEASES 2017; 34; 165-169 DOI: 10.36141/svdld.v34i2.5434

Inter-rater reliability of cutaneous sarcoidosis assessment tools via remote photographic assessment

Sara A. Berg[1], Howa Yeung[2], Joseph C. English[3], Emily L. Keimig[4], Ellen J. Kim[1], Robert G. Micheletti[1], Karolyn A. Wanat[5], Marc A. Judson[6], Robert P. Baughman[7], Misha Rosenbach[1]

[1] University of Pennsylvania, Department of Dermatology, Philadelphia, PA; [2] Emory University School of Medicine, Department of Dermatology, Atlanta, GA; [3] University of Pittsburgh, Department of Dermatology, Pittsburgh, PA; [4]Northwestern University, Department of Dermatology, Chicago, IL; [5] University of Iowa Carver College of Medicine, Department of Dermatology and Pathology, Iowa City, IA; [6] Albany Medical College, Albany, NY; [7] University of Cincinnati Medical Center, Cincinnati, OH

Abstract. *Background:* Recently two outcome instruments have been developed and validated for assessing cutaneous sarcoidosis in a live, in-person setting. Teledermatology is a rapidly growing field; yet, to date, no instrument has been validated for use in a remote setting, which could ultimately impact clinical trial design. *Objective:* To assess the interrater reliability of these outcome instruments for store-and-forward teledermatology. *Methods:* Seven sarcoidosis experts, including both pulmonologists and dermatologists, scored photographs of cutaneous sarcoidosis lesions in 13 patients utilizing the Cutaneous Sarcoidosis Activity and Morphology Index (CSAMI), the Sarcoidosis Activity and Severity Index (SASI) and the Physician Global Assessment (PGA). Interrater reliability was assessed for each instrument and was compared to results obtained from a prior study involving sarcoidosis experts evaluating the same patient population in an in-person setting. *Results:* Interrater reliability (presented as ICC [95%CI]) was poor for the CSAMI Activity scale (0.36 [0.16 - 0.65]) and the CSAMI Damage scale (0.17 [0.04 - 0.43]) and was fair for the Modified Facial SASI (0.59 [0.36 - 0.82]) and the PGA (0.47 [0.23 - 0.74]). All results were inferior to those obtained from the prior studies validating these instruments for in-person use. *Conclusions:* Given the superiority of these instruments when utilized in person, it is recommended to have an on-site sarcoidosis expert evaluate cutaneous sarcoidosis lesions whenever possible. *(Sarcoidosis Vasc Diffuse Lung Dis 2017; 34: 165-169)*

Key words: cutaneous sarcoidosis, teledermatology, outcome instruments, reliability

Introduction

Sarcoidosis is a multisystem inflammatory disorder of unknown etiology characterized by the formation of noncaseating granulomas in affected organs (1). Cutaneous involvement is seen in 25-30% of sarcoidosis patients and is the presenting feature in one third of all cases (1,2). Historically, management decisions for cutaneous sarcoidosis have been based on small retrospective series due to a lack of randomized clinical trials (1). While there has been a recent emergence of such trials (3-6), validated outcome instruments for cutaneous sarcoidosis were rarely used, limiting the strength and applicability of these results.

Two outcome measures for cutaneous sarcoidosis have been developed and validated. The Sarcoidosis

Received: 10 May 2016
Accepted after revision: 27 October 2016
Correspondence: Misha Rosenbach, MD
Department of Dermatology and Medicine,
Perelman School of Medicine, University of Pennsylvania,
3600 Spruce St, Second Floor, Maloney Building,
Philadelphia, PA 19104
E-mail: misha.rosenbach@uphs.upenn.edu

Activity and Severity Index (SASI) was the first tool developed and has been used in clinical trials (4,7-9). Contemporaneously, the Cutaneous Sarcoidosis Activity and Morphology Index (CSAMI), was also introduced and validated. The CSAMI specifically aims to measure both disease activity and resultant/persistent damage (a possible marker of fibrotic disease) as well as information about the cutaneous lesion morphology and presence of cutaneous findings with important prognostic information, such as lupus pernio or erythema nodosum (10). When evaluated among dermatologists, both instruments exhibited excellent intra-rater reliability, acceptable inter-rater reliability and convergent validity with the Physician Global Assessment. The CSAMI also correlated with health-related quality of life measures (10). A recent study at our institution involving an interdisciplinary group of dermatologists and non-dermatologist physicians yielded similar results (11). Specifically, all instruments exhibited good to excellent intra and inter-rater reliability, as well as moderate correlations with the PGA.

Given the high rates of pulmonary involvement, clinical trials for sarcoidosis have focused primarily on pulmonary endpoints. However, the accessibility of skin lesions, and low-cost, validated, reliable tools to assess cutaneous disease have led to increasing consideration of using cutaneous sarcoidosis improvement as a primary endpoint in clinical trials. Most sarcoidosis clinical trial principal investigators are pulmonologists, as these subspecialists care for the vast majority of sarcoidosis patients. It may be problematic to coordinate the in-person assessment of cutaneous sarcoidosis lesions by dermatologists in these clinical trials. The difficulty in obtaining prompt dermatology assessments by dermatologists because of physician shortages has rendered teledermatology increasingly useful for such purposes (12-14). To our knowledge, no study has assessed the validity of cutaneous outcome instruments via remote store-and-forward teledermatology. Accurate evaluation of cutaneous disease activity via remote physician evaluation of photographs could improve future clinical trial design as all sites could store and forward patient photographs for expert evaluation.

Our primary objective was to assess the inter-rater reliability of CSAMI when used by a group of sarcoidosis experts to evaluate store-and-forward photographs of patients with cutaneous sarcoidosis. These results were compared to results obtained from live sarcoidosis experts evaluating the same patient set to assess for the validity of the CSAMI in teledermatology.

Methods

The study was approved by the University of Pennsylvania institutional review board. Written informed consent was obtained from all patients.

Physician Participants

Five dermatologists with sarcoidosis expertise and two pulmonologists with extensive experience in cutaneous sarcoidosis from several large University hospitals were invited to participate in this study in February 2015. To limit the bias related to familiarity with specific instruments, investigators who previously participated in pilot validation studies of both the SASI (R.P.B. and M.A.J.) and the CSAMI (E.J.K., K.A.W., and M.R.) were included in this study. All participants underwent a standardized, self-directed computer training session developed by the primary investigator (M.R.) on the assessment of cutaneous sarcoidosis using the three outcome instruments.

Patient Participants

Fourteen subjects with clinical and pathological evidence of cutaneous sarcoidosis were recruited via telephone from the Cutaneous Sarcoidosis Clinic at the Hospital of the University of Pennsylvania to participate in a single day study to assess the psychometric properties of the outcome instruments when used by an interdisciplinary group of physicians (11). In order to be enrolled in this study, subjects were required to consent to have their skin lesions photographed for use in this photovalidation study. Subjects were selected by the principal investigator to include a range of cutaneous sarcoidosis presentations and severity. Since the SASI was validated based on evaluation of sarcoidosis lesions on the face only, we included photographs of facial lesions only for this study. Thirteen subjects participated in the single-day study, and all consented to be part of this photovalidation sub-study (11). All photographs were taken

on the study day and were securely stored on the principal investigator's computer. The Information Services Department at the University of Pennsylvania created a double-secure, HIPAA-compliant, temporary webpage to host the patient photographs. Each physician was given a unique username and password that could be used to access the subject photographs.

Study Design

The physicians rated each patient photograph using the CSAMI, SASI and the Physician's Global Assessment (PGA). Physicians could use a zoom feature to aid in their assessments. A sample patient photograph with the magnifying feature displayed is shown in Figure 1. The webpage and subject photographs were removed from the server at the conclusion of the study to protect patient confidentiality.

Cutaneous Sarcoidosis Activity and Morphology Index

The CSAMI captures cutaneous sarcoidosis disease activity and damage via two separate scales (10). The Activity scale evaluates inflammation, induration and/or depression, surface change and area of involvement; the Damage scale measures post-inflammatory changes including dyspigmentation and scarring. When limited to facial lesions, the Activity score may range from 0-75 and the Damage score from 0-10. The predominant morphologic types of the lesions, as well as the presence or absence of lupus pernio, were also documented. The presence of erythema nodosum was not evaluated given the focus on facial lesions in this study.

Sarcoidosis Activity and Severity Index

The original SASI measures erythema, induration, desquamation and area of activity in the four quadrants of the face as well as the nose, yielding 5 different scores for a single patient (7). SASI has since been modified for ease of use in clinical trials to result in a single score per patient by summing erythema, induration and desquamation for each quadrant, multiplying by the respective area of involvement within that quadrant and averaging the results across all 5 facial regions (4,8-11). Scores for the modified SASI range from 0-72.

Physician Global Assessment

The PGA is a linear visual analog scale measuring disease activity from 0 (no skin disease) to 10 (severe skin disease). The PGA has been utilized in prior studies validating outcome instruments for inflammatory skin conditions including pemphigus vulgaris, dermatomyositis and cutaneous sarcoidosis as a measure for physicians' overall perception of skin disease (10,11,15,16).

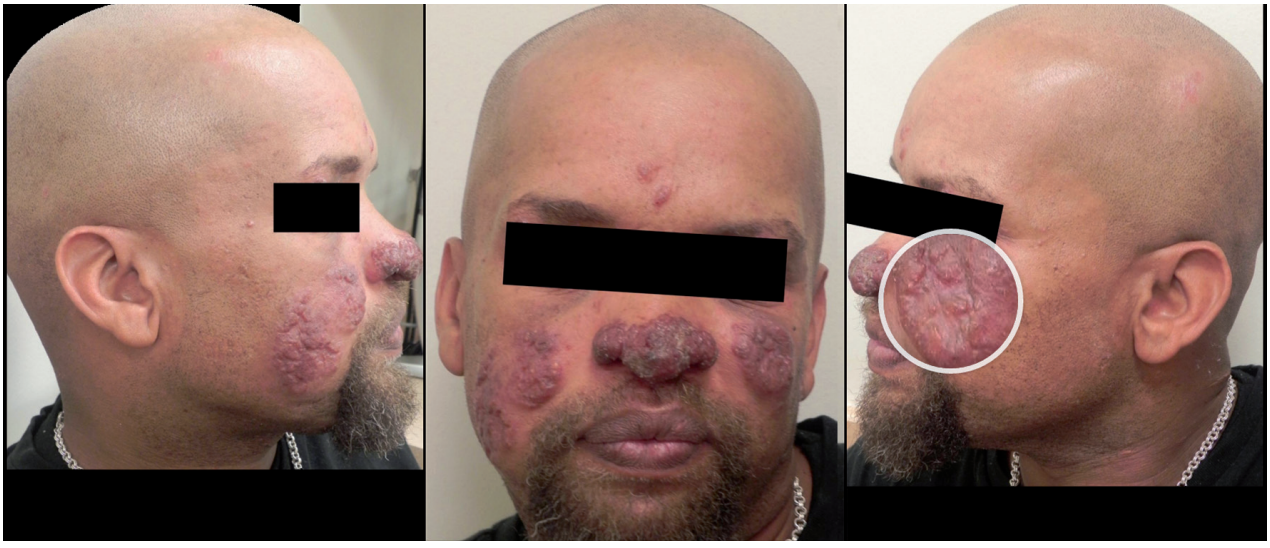

Fig. 1. Sample patient photograph demonstrating the magnifying feature.

Statistical Analysis

Data analysis was performed in February 2016. Scores from each instrument were summarized descriptively. The interrater reliability of each instrument was analyzed using the intraclass correlation coefficient (ICC) using two-way random-effects models and interpreted as poor (<0.40), fair to good (0.40-0.75), and excellent (>0.75) (17).

Results

Thirteen subjects participated in this study. The mean subject age was 51.2 years (SD 10.6). Twelve subjects were African American, and one patient was Caucasian. Three patients were male. The subjects had documented extra-cutaneous sarcoidosis in a mean of 1.7 (SD 1.3) organ systems. The subjects represented a spectrum of cutaneous disease, as illustrated by the range of disease severity scores seen in Table 1.

Interrater Reliability

Interrater reliability (presented as ICC [95%CI]) was poor for the CSAMI Activity scale (0.36 [0.16-0.65]) and was fair for the Modified Facial SASI (0.59 [0.36-0.82]) and the PGA (0.47 [0.23-0.74]) (Table 2).

Table 1. Disease severity ratings of patient participants

	Range (Max. Range)	Median (IQR)
CSAMI Activity	0-38 (0-75)	15 (10-20)
Inflammation	0-10 (0-15)	4 (2-6)
Induration/Depression	0-9 (0-15)	3 (2-5)
Surface changes	0-5 (0-15)	0 (0-1)
Area	0-16 (0-30)	6 (4-10)
CSAMI Damage	0-9 (0-10)	2 (0-3)
Modified Facial SASI	0-27.4 (0-72)	3.6 (1.6-7.2)
Erythema	0-4 (0-4)	0.6 (0.2-1.2)
Induration	0-4 (0-4)	0.6 (0.4-1.2)
Desquamation	0-2.8 (0-4)	0 (0-0.2)
Area	0-5 (0-6)	1 (0.6-1.8)
PGA	0.1-9.7 (0-10)	4.8 (2.9-6.5)

Abbreviations: IQR, interquartile range; CSAMI, Cutaneous Sarcoidosis Activity and Morphology Instrument; SASI, Sarcoidosis Activity and Severity Index; PGA, Physician's Global Assessment

Table 2. Inter-rater reliability of cutaneous sarcoidosis severity measures in a teledermatologic setting

	Inter-rater reliability (N = 7 MD's) ICC (95% CI)
CSAMI Activity	0.36 (0.16-0.65)
Inflammation	0.26 (0.09-0.54)
Induration/Depression	0.40 (0.19-0.68)
Surface changes	0.60 (0.38-0.82)
Area	0.22 (0.07-0.49)
CSAMI Damage	0.17 (0.04-0.43)
Modified Facial SASI	0.59 (0.36-0.82)
Erythema	0.59 (0.33-0.82)
Induration	0.79 (0.62-0.92)
Desquamation	0.52 (0.30-0.77)
Area	0.19 (0.05-0.46)
PGA	0.47 (0.23-0.74)

Abbreviations: ICC, intraclass correlation coefficient; CSAMI, Cutaneous Sarcoidosis Activity and Morphology Instrument; SASI, Sarcoidosis Activity and Severity Index; PGA, Physician's Global Assessment.
ICC interpretation: <0.4, poor; 0.4-0.75, fair to good; and >0.75, excellent

Discussion

To our knowledge, this is the first study that has been performed to assess the validity of cutaneous sarcoidosis outcome instruments in the setting of store-and-forward teledermatology. Prior studies have psychometrically validated the CSAMI and SASI when assessed by dermatologists, pulmonologists, and rheumatologists with expertise in sarcoidosis in face-to-face evaluations (7,10,11). Specifically, the CSAMI Activity scale has consistently demonstrated good to excellent interrater reliability, whereas the Modified Facial SASI scales has shown acceptable reliabilities (11). All instruments in our study demonstrated poor to fair reliability when used by all sarcoidosis experts in a teledermatologic setting. This is an important finding to consider when planning clinical trials; based on these results, it would be ideal to have an on-site dermatologist available to evaluate sarcoidosis patients in clinical trials whenever possible.

Several limitations must be considered when reviewing our pilot study. The quality of the images used in the study was not standardized. Comparison of the reliabilities of CSAMI erythema (inflammation) vs. dyspigmentation (damage) are shown in Table 2, and both are inferior as compared to SASI scores. Subtle erythema and dyspigmentation may be conflated in images of darker skin types, which may explain this lower reliability in images. However, er-

ythema alone was assessed by both the CSAMI and SASI scores using the same set of images and thus could not account for a bias in reliability.

The majority of physician participants in the study were from institutions other than our own; therefore, they were not familiar with all the outcome measures and this may have affected the reliability of this study. In an effort to surmount this problem, we used remote training tools to educate these physicians concerning the outcome instruments. We acknowledge this training may have failed to capture some of the nuances of these instruments that can be addressed in an interactive, live training session, and future studies could consider incorporating a portfolio of sample images with in-person sarcoidosis expert scores as an additional training modality. Our study was limited to seven experienced physician participants, and larger studies involving additional sarcoidosis experts may be warranted to generalize these results. However, the fact that even these experts failed to demonstrate good to excellent inter-rater reliability suggests that remote photographic assessment of cutaneous sarcoidosis lesion is an inadequate outcome measure in clinical trials.

Our pilot study evaluated the reliability of cutaneous sarcoidosis outcome instruments for teledermatologic use. Both the CSAMI and SASI demonstrated fair to poor interrater reliability when used by a group of sarcoidosis experts to evaluate patients' skin disease remotely through store-and-forward teledermatology. Because validated tools exhibiting good to excellent intra- and inter-rater reliability have been developed involving face-to-face evaluations, we believe that presently such tools should be used by dermatologists, pulmonologists, and/or rheumatologists with expertise in the evaluation of cutaneous sarcoidosis in clinical trials.

Acknowledgements

The authors would like to thank Mr. Nick Davis for building and managing the secure website.

Funding:
This research was made possibility by a generous philanthropic gift from the Siegel family.

Conflict of interest:
Dr. Rosenbach led the development of the Cutaneous Sarcoidosis Activity and Morphology Instrument; Drs. Baughman and Judson led the development of the Sarcoidosis Activity and Severity Index.

References

1. Haimovic A, Sanchez M, Judson MA, Prystowsky S. Sarcoidosis: A comprehensive review and update for the dermatologist: Part I. cutaneous disease. J Am Acad Dermatol 2012; 66(5): 699.e1-699.e18+717-718.
2. Iannuzzi MC, Rybicki BA, Teirstein AS. Sarcoidosis. New Engl J Med 2007; 357(21): 2153-2165+2108.
3. Judson MA, Baughman RP, Costabel U, et al. Efficacy of infliximab in extrapulmonary sarcoidosis: Results from a randomised trial. Eur Respir J 2008; 31(6): 1189-1196.
4. Drake WP, Oswald-Richter K, Richmond BW, et al. Oral antimycobacterial therapy in chronic cutaneous sarcoidosis: A randomized single-masked placebo-controlled study. JAMA Dermatol 2013; 149(9): 1040-1049.
5. Droitcourt C, Rybojad M, Porcher R, et al. A randomized, investigator-masked, double-blind, placebo-controlled trial on thalidomide in severe cutaneous sarcoidosis. Chest 2014; 146(4): 1046-1054.
6. Pariser RJ, Paul J, Hirano S, Torosky C, Smith M. A double-blind, randomized, placebo-controlled trial of adalimumab in the treatment of cutaneous sarcoidosis. J Am Acad Dermatol 2013; 68(5): 765-773.
7. Baughman RP, Judson MA, Teirstein A, et al. Chronic facial sarcoidosis including lupus pernio: Clinical description and proposed scoring systems. Am J Clin Dermatol 2008; 9(3): 155-161.
8. Baughman RP, Judson MA, Ingledue R, Craft NL, Lower EE. Efficacy and safety of apremilast in chronic cutaneous sarcoidosis. Arch Dermatol 2012; 148(2): 262-264.
9. Judson MA, Baughman RP, Costabel U, et al. Safety and efficacy of ustekinumab or golimumab in patients with chronic sarcoidosis. Eur Resp J 2014; 44(5): 1296-1307.
10. Rosenbach M, Yeung H, Chu EY, et al. Reliability and convergent validity of the cutaneous sarcoidosis activity and morphology instrument for assessing cutaneous sarcoidosis. JAMA Dermatol 2013; 149(5): 550-556.
11. Yeung H, Farber S, Birnbaum BK, et al. Reliability and validity of cutaneous sarcoidosis outcome instruments among dermatologists, pulmonologists, and rheumatologists. JAMA Dermatol 2015; 151(12): 1317-1322.
12. Coates SJ, Kvedar J, Granstein RD. Teledermatology: From historical perspective to emerging techniques of the modern era part I: History, rationale, and current practice. J Am Acad Dermatol 2015; 72(4): 563-576.
13. Kimball AB, Resneck JS, Jr. The US dermatology workforce: A specialty remains in shortage. J Am Acad Dermatol 2008; 59(5): 741-745.
14. Nelson CA, Takeshita J, Wanat KA, et al. Impact of store-and-forward (SAF) teledermatology on outpatient dermatologic care: A prospective study in an underserved urban primary care setting. J Am Acad Dermatol 2016; 74(3): 484-490.
15. Rosenbach M, Murrell DF, Bystryn J, et al. Reliability and convergent validity of two outcome instruments for pemphigus. J Invest Dermatol 2009; 129(10): 2404-2410.
16. Goreshi R, Okawa J, Rose M, et al. Evaluation of reliability, validity, and responsiveness of the CDASI and the CAT-BM. J Invest Dermatol 2012; 132(4): 1117-1124.
17. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull 1979; 86(2): 420-428.