Reliability and Validity of Cutaneous Sarcoidosis Outcome Instruments Among Dermatologists, Pulmonologists, and Rheumatologists

Howa Yeung, MD; Sara Farber, BA; Belinda K. Birnbaum, MD; Jonathan Dunham, MD; Alexis Ogdie, MD, MSCE; Karen C. Patterson, MD; Aimee S. Payne, MD, PhD; Mary K. Porteous, MD; Milton D. Rossman, MD; Rebecca Sharim, MD; Junko Takeshita, MD, PhD; Victoria P. Werth, MD; Daniel B. Shin, MS; Sarah Price, BA; Misha Rosenbach, MD

IMPORTANCE Dermatologists, pulmonologists, and rheumatologists study and treat patients with sarcoidosis with cutaneous manifestations. The validity of cutaneous sarcoidosis outcome instruments for use across medical specialties remains unknown.

OBJECTIVE To assess the reliability and validity of cutaneous sarcoidosis outcome instruments for use by dermatologists and nondermatologists treating sarcoidosis.

DESIGN, SETTING, AND PARTICIPANTS We performed a cross-sectional study evaluating the use of the Cutaneous Sarcoidosis Activity and Morphology Instrument (CSAMI) and Sarcoidosis Activity and Severity Index (SASI) to assess cutaneous sarcoidosis disease severity and the Physician's Global Assessment (PGA) as a reference instrument. Four dermatologists, 3 pulmonologists, and 4 rheumatologists evaluated facial cutaneous sarcoidosis in 13 patients treated at a cutaneous sarcoidosis clinic in a 1-day study on October 24, 2014; data analysis was performed from November through December 2014.

MAIN OUTCOMES AND MEASURES Interrater and intrarater reliability and convergent validity, with correlation with quality-of-life measures as the secondary outcome.

RESULTS All instruments demonstrated excellent intrarater reliability. Interrater reliability (reported as intraclass correlation coefficient [95% CI]) was good for the CSAMI Activity scale (0.69 [0.51-0.87]) and PGA (0.66 [0.47-0.85]), weak for the CSAMI Damage scale (0.26 [0.11-0.52]), and excellent for the modified Facial SASI (0.78 [0.63-0.91]). The CSAMI Activity scale and modified Facial SASI showed moderate correlations (95% CI) with the PGA (0.67 [0.57-0.75] and 0.57 [0.45-0.66], respectively). The CSAMI Activity scale but not the modified Facial SASI showed significant correlations (95% CI) with quality-of-life instruments, such as the Dermatology Life Quality Index (Spearman rank correlation, 0.70 [0.25-0.90]) and the Skin Stigma raw score of the Sarcoidosis Assessment Tool (Pearson product moment correlation, 0.56 [0.01-0.85]).

CONCLUSIONS AND RELEVANCE The CSAMI and SASI were reliable and valid in assessing cutaneous sarcoidosis among our diverse group of specialists. The CSAMI Activity score also correlated with quality-of-life measures and suggested construct validity. These results lend credibility to expand the use of the CSAMI and SASI by dermatologists and nondermatologists in assessing cutaneous sarcoidosis disease activity.
Sarcoidosis is an uncommon multisystem inflammatory disease characterized by noncaseating granulomatous infiltrates, with cutaneous manifestations occurring in 25% to 30% of cases. Although studies of treatments for cutaneous sarcoidosis have been based on limited observational evidence, multiple randomized clinical trials on the efficacy of systemic treatments have emerged. In the absence of a standardized, validated, and widely accepted instrument to measure cutaneous sarcoidosis severity, many of these clinical trials used different ad hoc outcome instruments to assess objective clinical severity and therapeutic end points, thus limiting the validity of the results and comparability across trials.

Two outcome instruments for cutaneous sarcoidosis have undergone psychometric validation. The Sarcoidosis Activity and Severity Index (SASI) was the first instrument developed to measure cutaneous disease severity by a group of pulmonologists with expertise in treating sarcoidosis and has since been incorporated into multiple clinical studies. The Cutaneous Sarcoidosis Activity and Morphology Instrument (CSAMI) was proposed by a group of dermatologists led by the principal investigator of our study (M.R.) with experience in treating cutaneous sarcoidosis and was designed to capture disease activity and damage. Among dermatologists, both instruments demonstrated excellent intrarater reliability, acceptable interrater reliability, and convergent validity, whereas the CSAMI also suggested correlation with quality-of-life measures. Given the multisystem nature of sarcoidosis necessitating interdisciplinary evaluation, management, and research, a criterion standard outcome measure for the cutaneous manifestations of sarcoidosis should display satisfactory psychometric properties for use among various specialists managing this disease, including pulmonologists, rheumatologists, and dermatologists. To our knowledge, no study has examined the validity of cutaneous sarcoidosis outcome instruments across specialists.

Our primary objectives were to assess the intrarater and interrater reliabilities and convergent validity of the CSAMI and SASI in a group of pulmonologists, rheumatologists, and dermatologists. Our secondary objective was to evaluate the instruments' correlations with quality-of-life measures to address construct validity.

Methods

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

Physician Participants

Four dermatologists, 3 pulmonologists, and 4 rheumatologists—all board-certified, attending-level physicians except 1 pulmonology fellow and 1 rheumatology fellow—with experience in diagnosing and managing sarcoidosis were invited to complete this 1-day study in October 2014. All participants completed a training session on the assessment of cutaneous sarcoidosis using the CSAMI, SASI, and Physician's Global Assessment (PGA), which constituted the 3 outcome instruments.

Patient Participants

Patients from the Cutaneous Sarcoidosis Clinic at the Hospital of the University of Pennsylvania were recruited via telephone. All eligible patients had a clinical diagnosis of cutaneous sarcoidosis that was supported by histopathologic findings. Patients were selected by one of us (M.R.) to include a range of sarcoidosis presentation and severity. However, because the SASI only evaluates sarcoidosis lesions on the face, we limited recruitment of patient participants to those with facial involvement. Thirteen patients participated in the study (1 patient did not show), and all completed the following 3 self-administered surveys on the effect of cutaneous sarcoidosis on health-related quality of life: the Skindex-29, Dermatology Life Quality Index (DLQI), and Sarcoidosis Assessment Tool (SAT).

Study Design

In this cross-sectional study, physicians rated sarcoidosis in each patient using the CSAMI, SASI, and PGA. Each physician then rerated sarcoidosis in 2 patients with each instrument.

Cutaneous Sarcoidosis Activity and Morphology Instrument

The CSAMI consists of 2 separate scales measuring inflammatory activity and disease damage. The Activity scale measures inflammation, induration, and/or depression; surface changes, such as scaling and ulceration; and area of involvement. The Damage scale measures dyspigmentation and scarring. When limited to evaluation of facial areas, CSAMI scores may range from 0 to 75 (maximum activity) for the Activity scale and 0 to 10 (maximum damage) for the Damage scale. Morphologic types of cutaneous sarcoidosis lesions, including the presence of lupus pernio, were also documented. Erythema nodosum was not assessed given the focus on facial lesions.

Sarcoidosis Activity and Severity Index

The SASI measures erythema, induration, desquamation, and area of involvement for each of the 4 facial quadrants and the nose, thus producing 5 separate sets of scores per patient. The instrument has since been modified to be incorporated in clinical trials as a single severity score. The modified Facial SASI score is calculated by summing the erythema, induration, and desquamation scores; multiplying by the respective area scores; and calculating the mean across all 5 facial regions. The modified Facial SASI scores may range from 0 to 72 (maximum activity/severity).

Physician’s Global Assessment

At present, no criterion standard instrument exists with which we may assess criterion validity; alternatively, we use the PGA to assess convergent validity, with the expectation that valid outcome instruments would correlate positively with the PGA in reflecting overall physician-assessed disease severity. The PGA used in this study is a linear visual analog scale ranging from 0 (no evidence of disease) to 10 (extremely severe disease). The PGA has been used to rate the physician’s overall impression of disease severity in instrument validation stud-
ies for inflammatory skin disorders, such as dermatomyositis, pemphigus, and cutaneous sarcoidosis.9,14,15

Quality-of-Life Metrics
The Skindex-2911 and DLQI12 are validated and widely used dermatology-specific quality-of-life metrics. The Skindex-29 is a 29-item survey with Emotions, Symptoms, and Functioning domains, each ranging from 0 (no effect on quality of life) to 100 (effect always experienced).11 The DLQI is a 10-item survey assessing the effect of skin diseases on the patient’s life, with scores ranging from 0 (no effect) to 30 (extremely large effect).12 The SAT is a novel health-related quality-of-life instrument developed specifically for sarcoidosis, focusing on common issues affecting patients owing to pulmonary, ophthalmic, and cutaneous involvement.13,16,17 The SAT has demonstrated good internal consistency and convergent validity with the DLQI and patient-assessed global disease severity; a minimal clinically important difference has also been established.13 Raw sum scores were used as an alternative to converted scores based on item-response theory models.16 Our study focused on the Skin Concerns and Skin Stigma scales of the SAT to measure cutaneous sarcoidosis-specific quality-of-life burden; their maximal ranges are 0 to 40 and 0 to 20, respectively, with higher scores representing greater quality-of-life burden. These quality-of-life metrics were all used to assess construct validity, with the expectation that valid disease severity outcome instruments will correlate positively with patient-assessed effect of skin disease on quality of life.

Statistical Analysis
We performed data analyses from November through December 2014. Scores from each instrument were summarized descriptively. We assessed normality assumptions with skewness and kurtosis tests. The reliability of each instrument was analyzed using the intraclass correlation coefficient (ICC).18 Intrarater and interrater reliability ICCs were calculated using 1-way and 2-way random-effects models, respectively, and interpreted as poor (<0.40), fair to good (0.40-0.75), and excellent (>0.75).18,19 Intrarater and interrater reliabilities of CSAMI morphologic types were analyzed using κ statistics.20,21 Convergent and construct validities were assessed by comparing the CSAMI and SASI with other physician- and patient-reported outcome measures, such as the PGA and quality-of-life surveys, using Pearson product moment correlation (r value) or Spearman rank correlation (p value) as appropriate, with bootstrap confidence intervals.22 Correlation statistics were interpreted as slight (0-0.2), fair (>0.2 to 0.4), moderate (>0.4 to 0.6), substantial (>0.6 to 0.8), and almost perfect (>0.8) agreement.23 We used mixed-effects linear regression to confirm the linearity of the associations with the PGA, adjusting for interrater and intrarater variations as random effects and PGA scores as a fixed effect. Post hoc reliability analyses stratified by dermatologists and nondermatologists were performed. Statistical approach was independently reviewed and confirmed by a biostatistician who was not involved in the initial analysis (D.B.S.). All statistics were analyzed using STATA (version 12.1; StataCorp LP). We planned to include at least 12 patients with 11 physician ratings per patient to detect an interrater ICC of 0.7 with 80% power, when the ICC is 0.4 under the null hypothesis, using a 2-tailed F test with a significance level of .05.

Results
Of 22 eligible patients who were approached, 13 were available and participated in this study. Their mean (SD) age was 51.2 (10.6) years; 3 patients were male. Twelve patients were African American and 1 patient was white. Sarcoidosis involvement was documented in a mean (SD) of 1.7 (1.3) extracutaneous organ systems. The patients presented with a wide spectrum of skin disease severity as evidenced by the range of PGA scores from 0.2 to 9.8 (Table 1). Patients’ cutaneous morphologic features, as determined by one of us (M.R.), and current treatments are displayed in eTable 1 in the Supplement.

Intrarater and Interrater Reliabilities
All 3 study instruments demonstrated excellent intrarater reliability. Intrarater reliability (presented as ICC [95% CI]) was good for the CSAMI Activity scale (0.69 [0.51-0.87]) and PGA (0.66 [0.47-0.85]), weak for the CSAMI Damage scale (0.26 [0.11-0.52]), and excellent for the modified Facial SASI (0.78 [0.63-0.91]) (Table 2). The morphologic types overall demonstrated substantial intrarater reliability (κ = 0.66 [95% CI, 0.47-0.84]) and moderate interrater reliability (κ = 0.46 [95% CI, 0.33-0.59]). The predominant morphologic type selected also showed substantial intrarater reliability (κ = 0.66 [95% CI, 0.35-0.90]) and fair interrater reliability (κ = 0.35 [95% CI, 0.23-0.50]). The presence of lupus pernio displayed substantial intrarater reliability (κ = 0.74 [95% CI, 0.46-1.00]) and fair interrater reliability (κ = 0.34 [95% CI, 0.15-0.55]).

Table 1. Disease Severity Ratings of Patient Participants

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Score, Study Range</th>
<th>Score, Maximum Range</th>
<th>Median Score (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSAMI Activity scale</td>
<td>0-47</td>
<td>0-75</td>
<td>14 (10-18)</td>
</tr>
<tr>
<td>Inflammation</td>
<td>0-12</td>
<td>0-15</td>
<td>3 (2-5)</td>
</tr>
<tr>
<td>Induration and/or depression</td>
<td>0-12</td>
<td>0-15</td>
<td>3 (2-5)</td>
</tr>
<tr>
<td>Surface changes</td>
<td>0-7</td>
<td>0-15</td>
<td>1 (0-3)</td>
</tr>
<tr>
<td>Area</td>
<td>0-19</td>
<td>0-30</td>
<td>6 (4-8)</td>
</tr>
<tr>
<td>CSAMI Damage scale</td>
<td>0-10</td>
<td>0-10</td>
<td>2 (1-4)</td>
</tr>
<tr>
<td>Modified Facial SASI</td>
<td>0-43.2</td>
<td>0-72</td>
<td>3.2 (1.4-5.6)</td>
</tr>
<tr>
<td>PGA</td>
<td>0.2-9.8</td>
<td>0-10</td>
<td>3.6 (2.0-5.4)</td>
</tr>
</tbody>
</table>

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

* Rescored disease severity scores for intrarater reliability analysis were excluded for baseline descriptions.
Convergent Validity
The CSAMI Activity scale and modified Facial SASI demonstrated moderate correlations with the PGA (Table 2). The CSAMI Damage scale did not correlate with the PGA. Mixed-effects regression modeling also demonstrated that a 1-unit increase in PGA score significantly predicted linear increases in the CSAMI Activity scale (regression coefficient $\beta$, 2.42 [95% CI, 2.03-2.80]; $P < .001$) and modified Facial SASI (regression coefficient $\beta$, 1.55 [95% CI, 1.20-1.90]; $P < .001$) scores.

Construct Validity With Quality-of-Life Measures
The mean (SD) SkinIndex-29 Emotions, Symptoms, and Functioning domain scores were 65.8 (19.8), 44.6 (15.7), and 42.9 (16.7), respectively. The mean (SD) SAT Skin Concerns and Skin Stigma raw sum scores were 13.3 (11.8) and 11.1 (6.4), respectively. The median DLQI score was 2 (interquartile range, 1-7).

The CSAMI Activity scale demonstrated a strong correlation with the DLQI ($\rho = 0.70$ [95% CI, 0.25-0.90]) and moderate correlation with the SAT Skin Stigma raw score ($r = 0.56$ [95% CI, 0.01-0.85]). Several CSAMI Activity scale components also demonstrated significant correlations with the SkinIndex-29 Functioning, DLQI, and SAT Skin Concerns components (eTable 2 in the Supplement). The PGA demonstrated a strong correlation with the SkinIndex-29 Functioning component ($r = 0.75$ [95% CI, 0.34-0.92]). The modified Facial SASI failed to correlate with any health-related quality-of-life measures.

Exploratory Analyses
Post hoc analyses of interrater and intrarater reliabilities were performed with stratification between dermatologists and nondermatologists (Table 3). Intrarater reliability for the CSAMI Activity and Damage scales, modified Facial SASI, and PGA were excellent among dermatologists and nondermatologists. Intrarater reliability for the CSAMI Activity and Damage scales and modified Facial SASI trended higher among dermatologists than among nondermatologists; interrater reliability for the PGA was comparable between the 2 groups.

Discussion
This study was unique in its inclusion of pulmonologists, rheumatologists, and dermatologists who treat patients with sarcoidosis; by demonstrating reliability and convergent validity in this diverse group, our data lend credibility to expand the use of the CSAMI and SASI by dermatologists and nondermatologists in assessing cutaneous sarcoidosis disease activ-
Reliability and Validity of Cutaneous Sarcoidosis Outcome Instruments

Andrew A. Dunham, MD; Lay G. Ogdie, MD, MSc; noodles Sharim, MD, MPH; Bruce E. Farber, MD; Charles C. Payne, MD; Peter S. Porteous, BSc, MD; David L. Rossman, MD; Martine R. Patterson, MD; Yu-Juen Yeung, ScD; Mark Rosenbach, MD, MBA; Lewis B. Takeshita, MD; Michael A. Werth, MD, PhD; Robert Shin, MD; Andrea Price, MD; and Jennifer Rosenblatt, MPH

Objective: To assess the reliability and validity of the Cutaneous Sarcoidosis Activity Measure (CSAMI) and Sarcoidosis Assessment Tool (SASI) to evaluate disease severity in a large, diverse group of hematologic/oncologic and dermatologic specialists.

Methods: A secondary analysis of the Cutaneous Sarcoidosis Activity Measure International Multicenter Study. A total of 502 specialists from 24 institutions participated. Agreement among dermatologists was compared with that of nondermatologists. Intrarater reliability of the instruments was assessed. Convergent validity was assessed using the Physician’s Global Assessment (PGA), CSAMI, and SASI. Validity was assessed by examining the relationship between disease characteristics and the instruments. Additional studies were conducted to evaluate disease severity among physicians and patients.

Results: The CSAMI and SASI were reliable and valid in assessing cutaneous sarcoidosis severity. Intrarater reliability was higher among dermatologists than among nondermatologists (kappa: 0.75 for dermatologists vs. 0.57 for nondermatologists). Convergent validity was demonstrated using the PGA, CSAMI, and SASI. Additional studies showed significant correlations between the SASI and quality-of-life measures. Although our study was not powered to detect differences among specialties, overall reliability of the instruments was still noted among dermatologists only. This result may not be surprising, given that dermatologists are more likely to be involved in treating cutaneous sarcoidosis.

Conclusions: The CSAMI and SASI were reliable and valid in assessing cutaneous sarcoidosis severity among our diverse group of specialists. Given the reliability, convergent validity, and evidence supporting correlation with dermatology-specific and cutaneous sarcoidosis-specific quality-of-life metrics, we propose that the CSAMI be used as an outcome instrument for the evaluation of cutaneous disease severity in future studies of sarcoidosis.

ARTICLE INFORMATION

Accepted for Publication: May 22, 2015.

Published Online: August 12, 2015.

Author Affiliations: Department of Dermatology, Emory University School of Medicine, Atlanta, Georgia (Yeung); Department of Dermatology, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Farber, Payne, Takeshita, Werth, Shin, Price, Rosenbach); Division of Rheumatology, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Birnbaum, Dunham, Ogdie, Sharim); Center for Clinical Epidemiology and Biostatistics, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Ogdie, Takeshita, Shin); Division of Pulmonary, Allergy, and Critical Care, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Patterson, Porteous, Rossman).

Author Contributions: Drs Yeung and Rosenbach had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Yeung, Farber, Patterson, Rosenbach. Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: Yeung, Farber. Critical revision of the manuscript for important intellectual content: Yeung, Birnbaum, Dunham, Ogdie, Patterson, Payne, Porteous, Rossman, Sharim, Takeshita, Werth, Shin, Price, Rosenbach. Statistical analysis: Yeung, Ogdie, Shin. Obtained funding: Rosenbach. Administrative, technical, or material support: Farber, Sharim, Price, Rosenbach. Study supervision: Dunham, Rosenbach.

Conflict of Interest Disclosures: Dr Rosenbach led the development of the Cutaneous Sarcoidosis Activity and Morphology Instrument. No other disclosures were reported.
Funding/Support: This study was supported in part by the Medical Dermatology Career Development Award from the Dermatology Foundation (Dr Rosenbach) and by philanthropic support from the Siegel family.

Role of the Funder/Sponsor: The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Previous Presentation: The abstract of this study was presented as a poster at the Annual Meeting of the Society of Investigative Dermatology, May 6-9, 2015, Atlanta, Georgia.

Additional Contributions: Neelam Khan, BS, Suzette Báez VanderBeek, BS, Maryte Papadopoulos, BS, Katherine Liang, BS, and Cynthia Clark, RN, PhD, Department of Dermatology, University of Pennsylvania, provided administrative support. None received compensation for their contributions.

REFERENCES