Yang Monica Mu (Orcid ID: 0000-0002-2693-3648)
Yuan Yiwei (Orcid ID: 0000-0001-9898-7920)
Khanna Dinesh (Orcid ID: 0000-0003-1412-4453)
McMahan Zsuzsanna Hortobagyi (Orcid ID: 0000-0001-6461-8940)
Hinchcliff Monique Evangeline (Orcid ID: 0000-0002-8652-9890)

Clinical Phenotypes of Patients with Systemic Sclerosis with Distinct Molecular Signatures in Skin

Monica Yang MD[1], Vivien Goh MD[2], Jungwha Lee MS PhD[2], Monica Espinoza BS[3], Yiwei Yuan BS[3], Mary Carns MS[2], Kathleen Aren MA MPH[2], Lorinda Chung MD MS[4], Dinesh Khanna MD MS[5], Zsuzsanna H. McMahan MD MHS[6], Rishi Agrawal, MD[7], Lauren Beussink Nelson RDCS MPH[8], Sanjiv J Shah MD[8], Michael L. Whitfield PhD[3], Monique Hinchcliff MD MS[9]

[1]Department of Medicine, Division of Rheumatology, University of California San Francisco, San Francisco, California

[2]Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois

[3]Department of Biomedical Data Science, Geisel School of Medicine Dartmouth, Lebanon, New Hampshire

[4]Department of Internal Medicine, Stanford University, Palo Alto, California

[5]University of Michigan Scleroderma Program, Ann Arbor, Michigan

[6]Department of Internal Medicine, Division of Rheumatology, Johns Hopkins University, Baltimore, Maryland

[7]Department of Radiology, Northwestern University Feinberg School of Medicine, Chicago, Illinois

[8]Department of Internal Medicine, Division of Cardiology, Northwestern University Feinberg School of Medicine, Chicago, Illinois

[9]Department of Internal Medicine, Section of Rheumatology, Allergy & Immunology, Yale School of Medicine, New Haven, Connecticut

Corresponding Authors:

Monica Yang, MD

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/acr.24998

Rheumatology Fellow, Department of Medicine
University of California San Francisco
400 Parnassus Ave
San Francisco, CA, 94143
Email: Monica.yang@ucsf.edu
Phone: 908-444-3047

Word count: 4009

Funding: This research was supported by the National Institute of Health (NIH)/National Institutes of Arthritis and Musculoskeletal Skin Diseases T32 AR079068-01, K23 AR059763, K24 AR063120, K23 AR071473, and R01 AR073270, Scleroderma Research Foundation, Burroughs-Welcome PUP Big Data in the Life Sciences Training Program and the Dr. Ralph and Marian Falk Medical Research Trust.

Competing Interests: L.C. participated in consulting for Boehringer Ingelheim, Mitsubishi Tanabe, Genentech, Kyverna and Eicos, and received honoraria from Boehringer Ingelheim. S.J.S. participated in consulting for Abbott, Amgen, Aria, Axon, Bayer, Bristol-Myers Squibb, Boehringer Ingelheim, Boston Scientific, Boxer Capital, Cardiora, CVRx, Cyclerion, Cytokinetics, Edwards Lifesciences, Eidos, Eisai, eKO ai, GSK, Imara, Intellia, Ionis, Janssen, Keyto, Lilly Medical, Merck, MyoKardia, NGM, Novo Nordisk, Protena, Regeneron, Rivus Pharma, Sanofi, Shifamed, Tenax, United Therapeutics, and Third Rock Ventures, and receives royalties from UpToDate. D.K. participated in consulting for Bayer, Bristol-Myers Squibb, CSL Behring, Horizon, GSK, Boehringer Ingelheim, Corbus, Genentech, Chemomab, Theraly, Prometheus, and Astra Zeneca, and in sponsored research agreement with BMS, Pfizer and Horizon and owns stock in Eicos Sciences. M.L.W. participated in consulting for Celdara Medical LLC, Bristol-Myers Squibb, Corbus Pharmaceuticals, UCB, Third Rock Ventures and

Acceleron, and received honoraria from Abbvie. M.H. received honoraria from AbbVie and Boehringer Ingelheim. All other authors declare no competing interests.

Objective: Systemic sclerosis (SSc) patients are classified according to degree of skin fibrosis (limited and diffuse cutaneous) and serum autoantibodies. We undertook the present multicenter study to determine if intrinsic subset (IS) classification based upon skin gene expression yields additional valuable clinical information.

Methods: SSc patients and healthy participants (HP) were classified as Normal-like, Limited, Fibroproliferative and Inflammatory IS using a previously trained classifier. Clinical data were obtained (serum autoantibodies, pulmonary function testing, modified Rodnan Skin scores [mRSS], and high-resolution chest computed tomography [HRCT]). Statistical analyses were performed to compare patients classified by IS, traditional cutaneous classification, and serum autoantibodies.

Results: 223 participants (165 SSc [115 dcSSc and 50 lcSSc] and 58 HP) were classified. Inflammatory IS patients had higher mRSS (22.1±9.9, p <0.001) than other IS and dcSSc (19.4±9.4, p= 0.05) despite similar disease duration (median [IQR] months 14.9[19.9] vs 18.4[31.6], p=0.48). In multivariable modeling, no significant association between mRSS and RNA Pol III (p=0.07) or Scl-70 (p=0.09) was found. Radiographic ILD was more prevalent in Fibroproliferative IS compared to other IS (91%, p=0.04) with similar prevalence between lcSSc and dcSSc (67% vs. 76%, p=0.73). Positive Scl-70 antibody was the strongest ILD predictor (p<0.001). Interestingly, all lcSSc/Fibroproliferative patients were demonstrated radiographic ILD.

Conclusions: Classification by IS identifies patients with distinct clinical phenotypes versus traditional cutaneous or autoantibody classification. IS classification identifies subgroups of SSc patients with more radiographic ILD (Fibroproliferative), higher mRSS

(Inflammatory) and milder phenotype (Normal-like), and may provide additional clinically useful information to current SSc classification systems.

Significance and Innovation

- The intrinsic subset skin gene expression classification system is a newer SSc classification that may help identify patients with radiographic interstitial lung disease and more skin fibrosis compared to the traditional cutaneous classification (limited cutaneous and diffuse cutaneous) and serum autoantibody status.
- Significance lies in the ability to enrich SSc clinical trials for patients with similar molecular programs underlying skin disease who may respond to targeted treatment.
- Innovation lies in the potential to identify SSc subgroups that may not otherwise be identified by the traditional cutaneous classification and serum autoantibodies including patients with lcSSc who are at increased risk for interstitial lung disease.

Systemic Sclerosis (SSc/scleroderma) is a phenotypically diverse collagen vascular disease whose clinical hallmark is fibrosis of the skin and internal organs. Autoimmunity with autoantibody production and vascular dysfunction are thought to occur earlier in the disease course followed by progressive fibrosis and organ dysfunction in a subset of patients. Although two SSc subtypes (limited/lcSSc and diffuse cutaneous/dcSSc), have been clinically identified based on the pattern of skin fibrosis as defined by Leroy et al.,[1] this traditional cutaneous classification system does not reliably predict disease severity, progression and therapeutic response.[2] Furthermore in light of uniformly negative recent trials results in SSc cutaneous disease,[3,4] there is an unmet need for identifying and enrolling subgroups of SSc patients that may enhance treatment efficacy. In 2008, a new SSc classification system termed intrinsic subset (IS) classification based upon skin gene expression was identified.[5] The current study makes use of high quality and comprehensive clinical data to determine the clinical phenotype of SSc patients classified according to IS and how this novel classification system compares and complements traditional cutaneous and autoantibody classification.

Four IS: Normal-like, Limited, Fibroproliferative, and Inflammatory have been defined.[5] Since their discovery, these IS have been validated in several additional SSc cohorts and found to be expressed in skin and esophageal tissues.[6–10] Furthermore, the IS model has been increasingly utilized in post-hoc analysis of clinical studies and trial data to understand the presence/absence of skin disease improvement as assessed by the modified Rodnan skin score (mRSS).[11–13] However, while the gene expression IS are distinct from traditional cutaneous classification, the clinical phenotype and

significance of each individual IS are not well characterized, and the clinical utility of IS classification is not established.

The study objective was to define the clinical characteristics and markers of disease severity in SSc patients classified by IS and examine the additive value of IS classification to existing systems, including traditional cutaneous and autoantibody classification. Demographic, clinical characteristics (*e.g.,* laboratory, mRSS, pulmonary function tests [PFT], chest high-resolution computed tomography [HRCT], echocardiogram [echo]) and patient-reported outcomes (PROs) instrument data were used. We hypothesized IS classification would identify distinct groups of SSc patients with similar clinical phenotypes and provide more granular information to the existing classification systems with regards to severity of skin and lung disease.

Methods

Study Population

This study reports the baseline findings of a large, multi-center, prospective cohort study that was approved by the Institutional Review Board at Northwestern University (STU00004428), University of Michigan (UM00084385), Stanford University (IRB00026772) and Johns Hopkins University (NA_00087035). Study participants provided written informed consent. Participant-level data were prospectively obtained for SSc patients and healthy participants (HP). Paired skin biopsies of the non-dominant forearm were performed: one for histology and the other for DNA microarray analyses to measure gene expression. All SSc patients fulfilled the American College of Rheumatology/European League Against Rheumatism 2013 classification criteria.[14]

DNA microarray preparation

RNA was prepared from skin biopsies as previously reported.[5,10] Tissue homogenization was performed using Qiagen TissueLyserII. RNA purification was carried out in QIAcube with Qiagen's RNeasy Mini Kit and the Agilent2100 Bioanalyzer assessed RNA integrity. Samples had RNA integrity numbers >7. RNA concentration was measured with Thermo Scientific NanoDrop2000 Spectrophotometer and 200ng total RNA was amplified and labeled with Agilent Quick-Amp Labeling Kits. Cy3-labeled sample and Cy5-labeled Universal Human Reference RNA (Stratagene) were co-hybridized to Agilent Human Genome (4×44K) Microarrays (G4112F). Data were Log_2 Lowess normalized and filtered for probes with intensity ≥2-fold over local background in Cy3 or Cy5 channels. Data were multiplied by −1 to convert to Log_2(Cy3/Cy5) ratios. Probes with >20% missing data were excluded. Gene expression data are available on Gene Expression Omnibus (GSE59787).

Determination of Intrinsic Subsets

Intrinsic subsets were assigned using a previously trained Glmnet machine learning classifier.[15] Briefly, the training set consisted of three independent skin gene expression datasets, each processed separately using GenePattern. Missing values were imputed using K-Nearest Neighbors, the CollapseDataset module was run using median collapse mode, and genes were median-centered. Datasets were merged using only genes present in all three sets. Glmnet along with random forest, KernSmooth, and caret packages implemented in R were used to train supervised classifiers. 10x, 3-fold cross-validation was used to train the model and assess robustness. The test set consisted of three additional independent SSc skin gene expression datasets. As

previously published, the average classification accuracy of the single sample Glmnet machine learning classifier utilized was 87.1% compared to IS classification based upon unsupervised clustering algorithms of paired samples.[15]

Patient-level Data

Clinical information including demographic, laboratory, PFT, chest HRCT, echo and PRO instruments were collected within 3 months of the baseline skin biopsy used for gene expression analyses. Clinical characteristics included SSc subtype (limited vs. diffuse cutaneous), SSc disease duration (defined as months since the first non-Raynaud event), current immunosuppression use, and skin biopsy date. Laboratory data include anti-nuclear antibody (ANA) status and measurement of SSc-specific antibodies including anticentromere (ACA), anti-topoisomerase I (Scl-70), and anti-RNA polymerase III (RNA Pol III) (both Specialty Laboratories, Valencia, CA), and markers of inflammation including platelet count, erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP) (performed at the laboratory at each institution).

PFT values of interest included force vital capacity percent predicted (FVC%) and diffusing capacity of carbon monoxide precent predicted (DLCO%). Comprehensive clinical echo exams with Doppler were obtained for patients at Northwestern University and scored by one research echocardiographer (LBN) according to standardized research protocols. All echo measurements were overread by a board-certified cardiologist (SJS). A designated chest radiologist with ILD expertise (RA) scored HRCT exams for the degree of ground glass opacification (GGO) and fibrosis for each lung lobe to generate a total lung disease score (TLS), as previously described.[16,17] Presence or absence of ILD was also determined by (RA) based upon the presence of

morphologic features matching a recognized pattern of disease (i.e., NSIP, UIP). A second chest radiologist scored and classified a subset of studies for validation.[18] Both assessors were blinded to clinical and laboratory data. For subjects who underwent testing at an outside institution, HRCT and echo digital images were obtained and analyzed by the radiologist (RA) and/or echocardiographer (LBN) when available.

Three PRO questionnaires that have been validated in SSc were administered at the baseline visit. These included the Patient-Reported Outcomes Measurement Information System 29-Item General Health Profile (PROMIS-29), Functional Assessment of Chronic Illness Therapy Dyspnea (FACIT-Dyspnea) and Saint George's Respiratory Questionnaire (SGRQ).[19,20]

Statistical Analysis

Descriptive statistics were calculated, and cross-sectional analyses were performed. Categorical measures were summarized using counts and percentages, while continuous measures were summarized by mean and standard deviation (SD) or median and interquartile range [IQR]. Student's *t*-test, Analysis of Variance, Wilcoxon/Mann-Whitney test or Kruskal-Wallis test (continuous variables) and chi-square test or Fisher's exact test (categorical variables) were used as appropriate. The significance level was defined as $p < 0.05$. The data were analyzed using SAS v9.4 (SAS Institute, Inc., Cary, NC).

Multiple linear regression models were used to evaluate continuous outcomes (mRSS, TLS, FVC%, DLCO%) individually and exact logistic regression models were used to evaluate binary outcome (radiographic ILD presence) among IS. Within the models, covariates included disease duration (<36 vs. ≥36 months), SSc subtype

(diffuse vs. limited), SSc antibodies (ACA, Scl-70, and RNA Pol III), smoking status (current/past vs. never), and current immunosuppression (yes vs. no).

Results

SSc Cohort and Intrinsic Subset Classification

The study cohort consisted of 58 HP and 165 SSc patients including 50 (30%) with lcSSc and 115 (70%) with dcSSc. The mean (SD) age of SSc patients and HP was 50±11 years, and 42 ± 13, p=0.26, respectively. Most participants were female (83%) and white (77%). The median [IQR] SSc disease duration was 23.0[46.3] months for the total group, and 49.1[91.5] mo for lcSSc and 18.4[31.7] mo for dcSSc. The mean (SD) mRSS was 15.2 ± 10.2 for the total group, and 5.6 ± 2.9 for lcSSc and 19.4 ± 9.4 for dcSSc patients. Distribution of serum antibodies included 10% ACA, 27% Scl-70, and 30% RNA Pol III positive. Demographic data by SSc clinical subtype and IS classification are shown in Table 1.

Intrinsic subsets were assigned using a previously trained IS classifier.[15] A dendrogram depicting the distribution of patients classified according to IS and pattern of skin gene expression is shown in Supplemental Figure 1A. The gene signature denoting the Inflammatory IS totaled 217 genes and consisted of signaling molecules such as IL-17, IL-3, CCR7, CCL2, and STAT1. Pathway analysis database accessed through G:Profiler revealed the top upregulated molecular pathways included leukocyte activation, regulation of T cell activation, granulocyte activation, and chemokine signaling pathways. Fibroproliferative IS gene signature consisted of 741 genes including CD69, RTF, CCDC53, PIK3C3, and pathway analysis identified signatures

related to extracellular matrix formation, collagen deposition, and cell cycle processes. The normal-like IS signature consisted of 128 genes such as RPL37, KLF14, TMEM256, and pathway analysis showed an upregulation in organic compound processes, lipid biosynthesis, and lipid metabolic processes (Supplemental Figure 1B). Only two patients in the cohort were classified as Limited IS, likely due to intentional recruitment of patients with active skin disease, and thus there were insufficient data to characterize this subset further.

Of the 58 HP, 54 (93%) were classified as Normal-like and four (7%) were classified as Fibroproliferative. Of the 165 SSc patients, 58 belonged to Normal-like, two to Limited, 72 to Inflammatory, and 33 to Fibroproliferative IS. Patients with dcSSc were classified in the Inflammatory (56%), Normal-like (23%) and Fibroproliferative (21%) IS, while lcSSc patients belonged to the Normal-like (64%), Fibroproliferative (18%), Inflammatory (14%) and Limited (4%) IS.

Demographics and Serologic Findings

Demographic and serologic differences were observed when classifying SSc patients by IS compared to traditional cutaneous classification (Table 2). Mean age and sex among IS were similar. There was a trend toward higher percentage of white patients in the Normal-like IS and higher percentage of black patients in the Fibroproliferative IS, but the results were not significant (p=0.22). Disease duration differed among IS with Inflammatory IS having the shortest disease duration at 15[20] months followed by Fibroproliferative IS (21[35] mo). Limited IS had the longest at 121[8] mo followed by Normal-like (45[84] mo) ($p < 0.001$). Similar findings were seen within traditional cutaneous classification with dcSSc having shorter disease duration

than lcSSc patients (18[32] vs. 49[92] mo, p <0.001). At the time of skin biopsy, 24 of 165 (14.6%) SSc patients were on immunosuppression, primarily mycophenolate mofetil (88%), followed by methotrexate (8%) and azathioprine (4%).

SSc-associated antibodies differed between IS and were distinct from traditional cutaneous classification. Anticentromere antibodies were more common in Normal-like IS while Scl-70 and RNA Pol III antibodies were more common in Fibroproliferative and Inflammatory IS patients, respectively (p<0.001). LcSSc patients had more ACA positivity and dcSSc patients had more RNA PolIII positivity (p<0.001); however, there was no difference in Scl-70 positivity between lcSSc and dcSSc patients (p=0.29). Platelet count was elevated in the Inflammatory IS compared to other IS (p<0.001). CRP levels demonstrated a similar trend (p=0.20) while ESR was similar between IS (p=0.82). Traditionally classified patients demonstrated a similar inflammatory marker pattern, with significantly elevated platelet count (p=0.004) and CRP (p=0.002) in dcSSc compared to lcSSc, but no difference in ESR between subtypes (p=0.63).

Skin, Pulmonary and Cardiac Characteristics

Skin score differed significantly among IS with Inflammatory IS having significantly higher mRSS (22.1 ± 9.9, p<0.001) followed by Fibroproliferative (12.3 ± 7.1), and then Normal-like (8.6 ± 5.9) IS (Table 3). The Inflammatory IS had higher mRSS than dcSSc subgroup (22.1 ± 9.9 vs. 19.4 ± 9.4, p=0.05) (Figure 1A). Within the dcSSc subgroup, when patients were further stratified by IS, the Inflammatory IS also demonstrated higher mRSS compared to Fibroproliferative and Normal-like IS (23.8 ± 8.8 vs. 15 ± 6.2 vs. 12.2 ± 6.3, p<0.001) (Table 4). mRSS among lcSSc patients stratified by IS were similar.

Utilizing multivariable modeling of mRSS, Inflammatory IS continued to have significantly higher mRSS (adjusted mean = 18.1, 95%CI 15.2-21.1) compared to Fibroproliferative (11.3, 8.0-14.6) and Normal-like (10.0, 6.9-13.1) IS with IS and cutaneous classification having the most significant effects in the model (both p<0.001) (Table 5). RNA Pol III positivity, Scl-70 positivity, and disease duration were not significantly associated with increased mRSS (p=0.07, p=0.09, p=0.70, respectively). Black race also was not associated with increased mRSS (p=0.41).

With regards to pulmonary findings, Fibroproliferative IS had significantly more lung disease compared to other IS with 91% of patients having radiographic ILD present on HRCT (vs. Inflammatory 65% and Normal-like 62%, p=0.038) despite having shorter disease duration than Normal-like IS (p=0.03) and comparable disease duration with Inflammatory IS (p=0.22). In contrast, there was no significant difference of radiographic ILD among SSc patients classified traditionally, (dcSSc 76% vs. lcSSc 67%, p =0.73) (Figure 1B). Within traditional cutaneous classification, the pulmonary findings of Fibroproliferative IS persisted. Among dcSSc patients, dcSSc/Fibroproliferative IS had the highest TLS (12.5 ± 9.7) and higher positive radiographic ILD (89%) compared to other IS (p=0.009, p=0.04 respectively), despite only 54% of the group being Scl-70 positive. Among lcSSc patients, all lcSSc/Fibroproliferative patients had positive ILD diagnosis on HRCT compared to lcSSc/Normal-like (62%) and lcSSc/Inflammatory (67%) (p<0.001) despite only 40% of patients being Scl-70 positive.

Utilizing multivariable modeling, IS was significantly associated with ILD presence (p=0.016) with the Fibroproliferative IS having a higher likelihood of positive radiographic ILD compared to Normal-like (odds ratio/OR =5.58, 95% CI 1.29 - 46.90,

p=0.01) than Inflammatory IS vs. Normal-like (OR = 1.33, 95% CI 0.45-3.94, p=0.38). While traditional cutaneous classification was not significantly associated with increased TLS (p=0.31) or positive radiographic ILD (p=0.64), Scl-70 positivity had a strong association with both (p<0.001, p<0.001) and was associated with lower DLCO% (p=0.01). Neither IS nor Scl-70 positivity was associated with lower FVC% (p=0.43, p=0.47, respectively) (Table 5).

Echocardiogram findings including left ventricular ejection fraction (LVEF), LV mass, tricuspid annular plane systolic excursion (TAPSE), right ventricular fractional area change (RV FAC), pulmonary artery systolic pressure (PASP) and diastolic dysfunction (DD) were similar among different IS groups (p>0.1) as well as among lcSSc and dcSSc patients (p>0.1). Normal-like IS demonstrated a trend toward increased prevalence of diastolic dysfunction however, the result was not significant (p=0.19).

Patient Reported Outcomes Instruments

The PROMIS-29 questionnaire assesses seven health domains: physical function, anxiety, depression, fatigue, sleep disturbance, satisfaction with social role, and pain and includes four questions (items) for each domain plus a 1-10 pain scale. There were significant differences in scores for four of seven domains for patients classified by IS (physical function, p=0.02; sleep disturbance, p=0.04; satisfaction with social participation, p=0.013; and pain, p=0.002) but differences in only one domain for patients traditionally classified (social participation, p=0.001) (Supplemental Table 1). Patients in the Inflammatory and Fibroproliferative IS reported more limitations in physical function, sleep disturbances, pain, and less satisfaction with social participation

compared to the Normal-like IS. The FACIT-dyspnea and SGRQ questionnaires, which were limited by fewer responses, did not demonstrate significant differences between IS or traditional cutaneous classification (Supplemental Table 1).

Discussion

Systemic sclerosis is a clinically heterogeneous disease with some patients demonstrating relatively normal organ function and low symptom burden while others experience life-threatening and disabling disease. Recent clinical trial results for SSc skin disease have been disappointing with patients who were randomized to several promising agents (*e.g.,* abatacept, lenabasum, tocilizumab) failing to meet the primary clinical endpoint.[3,4,13] One reason for negative trial results may be our inability to identify SSc patients with similar phenotypes in whom targeted treatments are more likely to be effective. The study aim was to elucidate the clinical phenotypes of SSc patients classified according to the newer IS classification system based upon skin gene expression and examine how they compare and add to traditional cutaneous and antibody status classifications. The goal was to determine whether IS classification may be advantageous for identifying SSc patients with similar disease phenotype. We demonstrate IS classification identifies distinct clinical SSc phenotypes that have both shared and unique features with clinically available subtypes.

Utilizing genome-wide gene expression analysis of the skin, Milano et al. was the first to classify patients among four IS that appeared biologically relevant and distinct from the traditional cutaneous classification system.[5] The Inflammatory IS most highly expressed genes associated with the presence of inflammatory infiltrates and increased

immune response.[5] In the Fibroproliferative IS, genes associated with cell proliferation were highly expressed and interestingly, genes associated with fatty acid and lipid synthesis were downregulated, which is a known SSc hallmark and a putative fibrotic pathogenic mechanism.[21,22] Recent studies have sought to understand the clinical implications of IS and how they may impact treatment decisions.[8,9,12] However, many of these studies have been limited by small sample size and lack of comprehensive patient-level data. This study utilized high-quality prospectively collected patient-level information including echo and chest HRCT data interpreted according to standardized research protocols.

When compared to Normal-like and Limited IS, the Fibroproliferative and Inflammatory IS appeared to have more severe SSc as assessed by mRSS, PFTs, chest HRCT, acute phase reactants, although each with a distinct phenotype. The Fibroproliferative IS consisted of 73% dcSSc patients and although not significant, a quarter of the Fibroproliferative patients were black, which in epidemiologic studies has been associated with more severe SSc and worse prognosis.[23,24] Moreover, Fibroproliferative IS had greater radiographic ILD prevalence and had higher TLS and lower FVC% and DLCO% than the Inflammatory IS, despite having more dcSSc patients in the latter IS and similar disease durations. The Inflammatory IS, which consisted of 90% dcSSc patients, had the highest platelet count and CRP compared to other IS, and a significantly higher mRSS than the other IS and the dcSSc group. Taken together, the Fibroproliferative IS appears to represent patients with more lung fibrosis in whom antifibrotic treatments may be most effective while the Inflammatory IS

identifies patients with more active skin disease in whom drugs targeting innate or adaptive immune responses may be more appropriate.

Patients with dcSSc and Scl-70 positivity are thought to be at higher risk for ILD.[2,25] We found the prevalence of radiographic ILD was similar between lcSSc and dcSSc patients. In multivariable models, patients with Scl-70 antibodies compared to those with ACA or RNA Pol III were more likely to have radiographic ILD. Moreover, positive Scl-70 was a better predictor of radiographic ILD compared to IS classification. However, we demonstrate that IS classification identifies additional lcSSc patients (lcSSc/Fibroproliferative) with radiographic ILD independent of antibody status. In fact, five lcSSc/Fibroproliferative patients had radiographic ILD within 29.9 ± 20.1 months of disease duration, and only two had Scl-70 antibodies. Furthermore, of these lcSSc patients, three had normal FVC% (≥ 80% predicted), and one had normal DLCO% (≥ 60% predicted), and chest HRCT may not have been pursued in these patients. The limitations of PFT has been demonstrated in our previous work involving 265 [188(71%) with radiographic ILD] SSc patients where we showed that 59 (31%) had "normal" FVC%, and 65 out of 151 (43%) had "normal" DLCO%.[18] Moreover, lcSSc/Fibroproliferative compared to lcSSc/Inflammatory and lcSSc/Normal-like patients were more likely to have radiographic ILD ($p<0.001$). Thus, while Scl-70 antibody appears to be the strongest risk factor for ILD in our cohort and has the benefit of being readily available, IS may allow identification of an additional at-risk group (*i.e.,* lcSSc/Fibroproliferative) for ILD. Future studies will determine which SSc subtype/IS combination have the highest risk of developing progressive ILD.

Among dcSSc patients, the Fibroproliferative IS also had more severe ILD compared to other IS. Using granular chest HRCT data, specifically, GGO and Fibrosis scores,[16] we found that dcSSc/Fibroproliferative patients had the highest GGO (8.1 ± 5.5) and Fibrosis scores (4.5 ± 4.9) while scores were lowest in lcSSc/Inflammatory patients (GGO 3.2±4.5, Fibrosis 1.3 ± 1.5). When faced with choosing the most appropriate FDA-approved treatment for SSc-ILD (tocilizumab, an IL-6 receptor antagonist,[3] versus nintedanib, a tyrosine kinase receptor antagonist[26]), tocilizumab might be best for the dcSSc/Fibroproliferative subset patients with high GGO scores while nintedanib may be more effective for patients with high Fibrosis scores. The verity of this hypothesis warrants testing.

With regards to skin disease, IS and cutaneous subtype had the strongest association with high mRSS ($p<0.001$); however, disease duration, Scl-70 positivity, and RNA Pol III positivity were not significantly associated. Among IS and cutaneous classification, dcSSc/Inflammatory patients had the highest mRSS (23.8 ± 8.8) compared to dcSSc/Fibroproliferative (15.0 ± 6.2) and dcSSc/Normal-like (12.2 ± 6.3) ($p <0.001$). In addition, dcSSc/Inflammatory patients had a higher mean mRSS than either dcSSc (19.4 ± 9.4) or Inflammatory IS (22.1 ± 9.9) alone, suggesting that together, IS and cutaneous classification may identify patients with the most severe skin disease, but determining the predictive ability of the dual classification will be important. Similar to dcSSc/Inflammatory patients, lcSSc/Inflammatory patients had the highest mean mRSS (6.1± 2.8) compared to lcSSc/Normal-like (5.8 ± 3.2) and lcSSc/Fibroproliferative (5.1 ± 2.2) groups, although the differences were not statistically nor likely clinically significant. However, we propose that consideration be given to treating skin disease in

SSc patients who demonstrate an inflammatory skin gene expression signature with anti-inflammatory agents, such as mycophenolate mofetil and abatacept, that have been previously demonstrated to impact mRSS regardless of whether they have lcSSc or dcSSc.[10,27] Of course, additional longitudinal analyses are necessary to evaluate the effectiveness of this approach.

Our study findings may help explain the recent negative results of the Phase III Lenabasum trial where dcSSc patients were randomized to receive oral lenabasum, a cannabinoid type 2 receptor agonist, vs. placebo.[4] The primary outcome (the ACR Combined Response Index in diffuse cutaneous systemic sclerosis [CRISS] score [28]) and secondary outcomes, including mRSS, were not met. However, results of post-hoc analyses from the previous Phase II trial found that patients lacked elevated IL-6 levels, and the majority of patients belonged to the Fibroproliferative IS.[29] Since CB2 receptors are primarily located on immune cells, the potential inclusion of Fibroproliferative IS rather than Inflammatory IS patients in the Phase III trial may explain the negative results. Forthcoming analyses of the skin gene expression data will shed light on this.

Study strengths include our multicenter study design and the collection of skin gene expression and comprehensive clinical data for a robust SSc cohort that included patients from four large academic centers. We were able to examine several clinical data domains including serologic, pulmonary, cardiac and PRO findings that had not previously been examined. While we aimed to obtain comprehensive data from all patients, there were a subset of patients lacking either echo (N=47) or HRCT (N=57), which is a study limitation. In the present study, we purposefully recruited patients with active skin disease in the opinion of the treating physician. It is possible that a subset of

patients classified as lcSSc, especially those with positive Scl-70 antibodies, had progressive skin disease over time and will be subsequently reclassified as having dcSSc. Analyses of longitudinal data are underway that will address this possibility. We utilized a validated IS classifier with an accuracy rate of 87.1%, which introduces the potential for misclassification. However, in comparison to previously utilized IS classifiers, this classifier can be applied to single compared to paired samples to increase utility and feasibility in studies and clinical trials. Overall, we did not correct for performing multiple statistical tests given the exploratory nature of the analyses and the intention to generate additional testable hypotheses. Moreover, the cross-sectional study design precluded determination of the prognostic value of IS classification. We recognize that our study is not the definitive study on the importance of IS classification, and work is underway analyzing prospectively collected longitudinal data.

In conclusion, we utilized molecular gene expression signatures of the skin to classify SSc patients into distinct IS that demonstrated unique clinical characteristics and varying disease severity that may provide added value to traditional cutaneous and antibody classifications. In addition to these existing classification models, the IS classification identifies patients with more skin fibrosis (Inflammatory IS) and radiographic ILD (Fibroproliferative IS) with Normal-like and possibly Limited IS patients exhibiting milder disease phenotype. Taken together, these findings demonstrate that the IS classification provides additive value that may be utilized in combination with the traditional cutaneous and serum autoantibody classification to aid clinicians and researchers in clinically impactful SSc patient stratification.

References:

1. LeRoy EC, Black C, Fleischmajer R, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. *J Rheumatol*. 1988;15(2):202-205.

2. Walker UA, Tyndall A, Czirják L, et al. Clinical risk assessment of organ manifestations in systemic sclerosis: a report from the EULAR Scleroderma Trials And Research group database. *Ann Rheum Dis*. 2007;66(6):754-763. doi:10.1136/ard.2006.062901

3. Khanna D, Lin CJF, Furst DE, et al. Tocilizumab in systemic sclerosis: a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Respir Med*. 2020;8(10):963-974. doi:10.1016/S2213-2600(20)30318-0

4. Spiera R, Kuwana M, Khanna D, et al. OP0171 PHASE 3 TRIAL OF LENABASUM, A CB2 AGONIST, FOR THE TREATMENT OF DIFFUSE CUTANEOUS SYSTEMIC SCLEROSIS (DCSSC). *Ann Rheum Dis*. 2021;80(Suppl 1):102. doi:10.1136/annrheumdis-2021-eular.1795

5. Milano A, Pendergrass SA, Sargent JL, et al. Molecular subsets in the gene expression signatures of scleroderma skin. *PloS One*. 2008;3(7):e2696. doi:10.1371/journal.pone.0002696

6. Pendergrass SA, Lemaire R, Francis IP, Mahoney JM, Lafyatis R, Whitfield ML. Intrinsic gene expression subsets of diffuse cutaneous systemic sclerosis are stable in serial skin biopsies. *J Invest Dermatol*. 2012;132(5):1363-1373. doi:10.1038/jid.2011.472

7. Johnson ME, Mahoney JM, Taroni J, et al. Experimentally-Derived Fibroblast Gene Signatures Identify Molecular Pathways Associated with Distinct Subsets of Systemic Sclerosis Patients in Three Independent Cohorts. Liossis SN, ed. *PLOS ONE*. 2015;10(1):e0114017. doi:10.1371/journal.pone.0114017

8. Taroni JN, Martyanov V, Huang CC, et al. Molecular characterization of systemic sclerosis esophageal pathology identifies inflammatory and proliferative signatures. *Arthritis Res Ther*. 2015;17:194. doi:10.1186/s13075-015-0695-1

9. Taroni JN, Greene CS, Martyanov V, et al. A novel multi-network approach reveals tissue-specific cellular modulators of fibrosis in systemic sclerosis. *Genome Med*. 2017;9(1):27. doi:10.1186/s13073-017-0417-1

10. Hinchcliff M, Huang CC, Wood TA, et al. Molecular Signatures in Skin Associated with Clinical Improvement during Mycophenolate Treatment in Systemic Sclerosis. *J Invest Dermatol*. 2013;133(8):1979-1989. doi:10.1038/jid.2013.130

11. Gordon JK, Martyanov V, Magro C, et al. Nilotinib (Tasigna™) in the treatment of early diffuse systemic sclerosis: an open-label, pilot clinical trial. *Arthritis Res Ther*. 2015;17:213. doi:10.1186/s13075-015-0721-3

12. Hinchcliff M, Toledo DM, Taroni JN, et al. Mycophenolate Mofetil Treatment of Systemic Sclerosis Reduces Myeloid Cell Numbers and Attenuates the Inflammatory Gene Signature in Skin. *J Invest Dermatol*. 2018;138(6):1301-1310. doi:10.1016/j.jid.2018.01.006

13. Khanna D, Spino C, Johnson S, et al. Abatacept in Early Diffuse Cutaneous Systemic Sclerosis: Results of a Phase II Investigator-Initiated, Multicenter, Double-Blind, Randomized, Placebo-Controlled Trial. *Arthritis Rheumatol Hoboken NJ*. 2020;72(1):125-136. doi:10.1002/art.41055

14. van den Hoogen F, Khanna D, Fransen J, et al. 2013 Classification Criteria for Systemic Sclerosis: An American College of Rheumatology/European League Against Rheumatism Collaborative Initiative: ACR/EULAR Classification Criteria for SSc. *Arthritis Rheum*. 2013;65(11):2737-2747. doi:10.1002/art.38098

15. Franks JM, Martyanov V, Cai G, et al. A Machine Learning Classifier for Assigning Individual Patients With Systemic Sclerosis to Intrinsic Molecular Subsets. *Arthritis Rheumatol Hoboken NJ*. 2019;71(10):1701-1710. doi:10.1002/art.40898

16. Kazerooni EA, Martinez FJ, Flint A, et al. Thin-section CT obtained at 10-mm increments versus limited three-level thin-section CT for idiopathic pulmonary fibrosis: correlation with pathologic scoring. *AJR Am J Roentgenol*. 1997;169(4):977-983. doi:10.2214/ajr.169.4.9308447

17. Strollo D, Goldin J. Imaging lung disease in systemic sclerosis. *Curr Rheumatol Rep*. 2010;12(2):156-161. doi:10.1007/s11926-010-0095-0

18. Showalter K, Hoffmann A, Rouleau G, et al. Performance of Forced Vital Capacity and Lung Diffusion Cutpoints for Associated Radiographic Interstitial Lung Disease in Systemic Sclerosis. *J Rheumatol*. 2018;45(11):1572-1576. doi:10.3899/jrheum.171362

19. Hinchcliff M, Beaumont JL, Thavarajah K, et al. Validity of two new patient-reported outcome measures in systemic sclerosis: Patient-Reported Outcomes Measurement Information System 29-item Health Profile and Functional Assessment of Chronic Illness Therapy-Dyspnea short form. *Arthritis Care Res*. 2011;63(11):1620-1628. doi:10.1002/acr.20591

20. Hinchcliff ME, Beaumont JL, Carns MA, et al. Longitudinal evaluation of PROMIS-29 and FACIT-dyspnea short forms in systemic sclerosis. *J Rheumatol*. 2015;42(1):64-72. doi:10.3899/jrheum.140143

21. Marangoni RG, Korman BD, Wei J, et al. Myofibroblasts in murine cutaneous fibrosis originate from adiponectin-positive intradermal progenitors. *Arthritis Rheumatol Hoboken NJ*. 2015;67(4):1062-1073. doi:10.1002/art.38990

22. Varga J, Marangoni RG. Systemic sclerosis in 2016: Dermal white adipose tissue implicated in SSc pathogenesis. *Nat Rev Rheumatol*. 2017;13(2):71-72. doi:10.1038/nrrheum.2016.223

23. Steen V, Domsic RT, Lucas M, Fertig N, Medsger TA. A clinical and serologic comparison of African American and Caucasian patients with systemic sclerosis. *Arthritis Rheum*. 2012;64(9):2986-2994. doi:10.1002/art.34482

24. Morgan ND, Shah AA, Mayes MD, et al. Clinical and serological features of systemic sclerosis in a multicenter African American cohort: Analysis of the genome research in African American scleroderma patients clinical database. *Medicine (Baltimore)*. 2017;96(51):e8980. doi:10.1097/MD.0000000000008980

25. Nihtyanova SI, Sari A, Harvey JC, et al. Using Autoantibodies and Cutaneous Subset to Develop Outcome-Based Disease Classification in Systemic Sclerosis. *Arthritis Rheumatol*. 2020;72(3):465-476. doi:10.1002/art.41153

26. Flaherty KR, Wells AU, Cottin V, et al. Nintedanib in Progressive Fibrosing Interstitial Lung Diseases. *N Engl J Med*. 2019;381(18):1718-1727. doi:10.1056/NEJMoa1908681

27. Chakravarty EF, Martyanov V, Fiorentino D, et al. Gene expression changes reflect clinical response in a placebo-controlled randomized trial of abatacept in patients with diffuse cutaneous systemic sclerosis. *Arthritis Res Ther*. 2015;17(1):159. doi:10.1186/s13075-015-0669-3

28. Khanna D, Berrocal VJ, Giannini EH, et al. The American College of Rheumatology Provisional Composite Response Index for Clinical Trials in Early Diffuse Cutaneous Systemic Sclerosis. *Arthritis Rheumatol Hoboken NJ*. 2016;68(2):299-311. doi:10.1002/art.39501

29. Spiera R, Hummers L, Chung L, et al. Safety and Efficacy of Lenabasum in a Phase II, Randomized, Placebo-Controlled Trial in Adults With Systemic Sclerosis. *Arthritis Rheumatol Hoboken NJ*. 2020;72(8):1350-1360. doi:10.1002/art.41294

Table/Figures:

Table 1: Clinical characteristics for study participants by disease state and systemic sclerosis subtype

Table 2: Demographic and clinical characteristics of SSc patients by intrinsic subset.

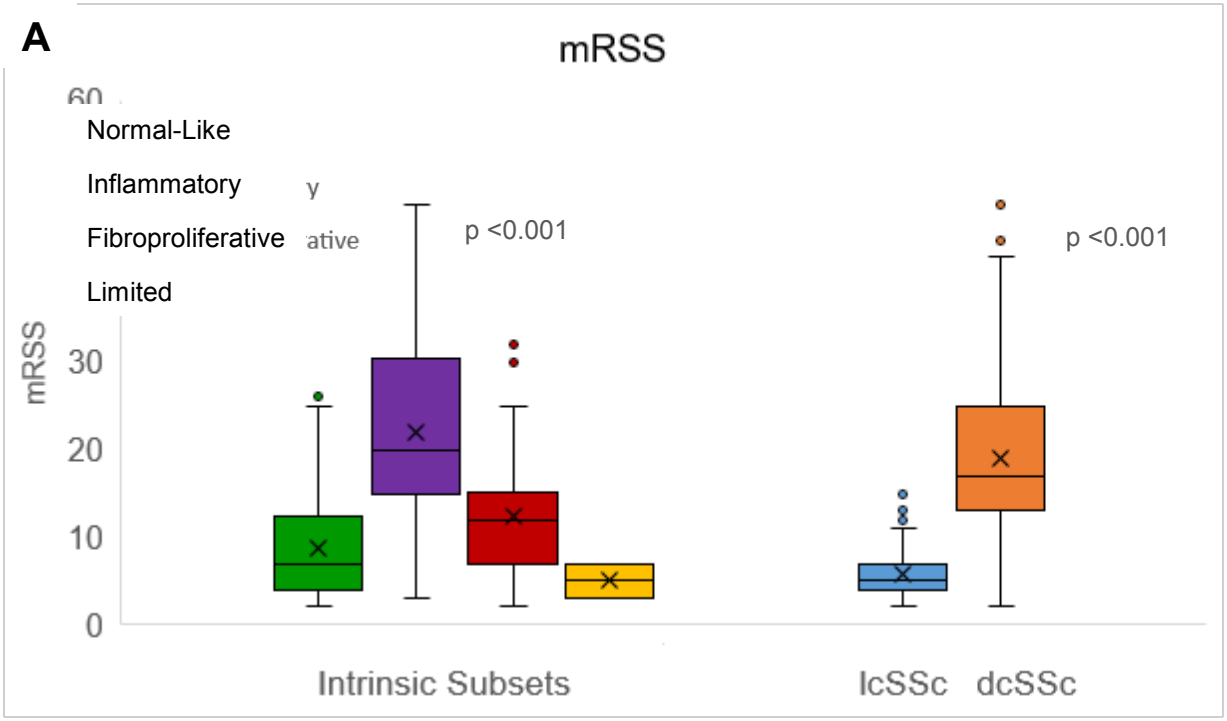
Table 3: Skin, pulmonary and cardiac findings of SSc patients by intrinsic subset and traditional cutaneous classification.

Figure 1: Skin and pulmonary manifestations in SSc patients by intrinsic subset compared to traditional cutaneous classification. A) Comparison of modified Rodnan skin score (mRSS) between IS and traditional classification. B) Comparison of prevalence of interstitial lung disease (ILD) on HRCT between IS and traditional classification.

Table 4: Clinical characteristics of SSc patients classified first by traditional cutaneous classification and then by intrinsic subset.

Table 5: Clinical Predictors of SSc Skin and Pulmonary Outcomes. Shown are the clinical predictors, and their associated difference in modified Rodnan skin score, total lung score, force vital capacity percent predicted, diffusion capacity of carbon monoxide,

and presence of radiographic interstitial lung disease as resulted from multiple regression modeling.



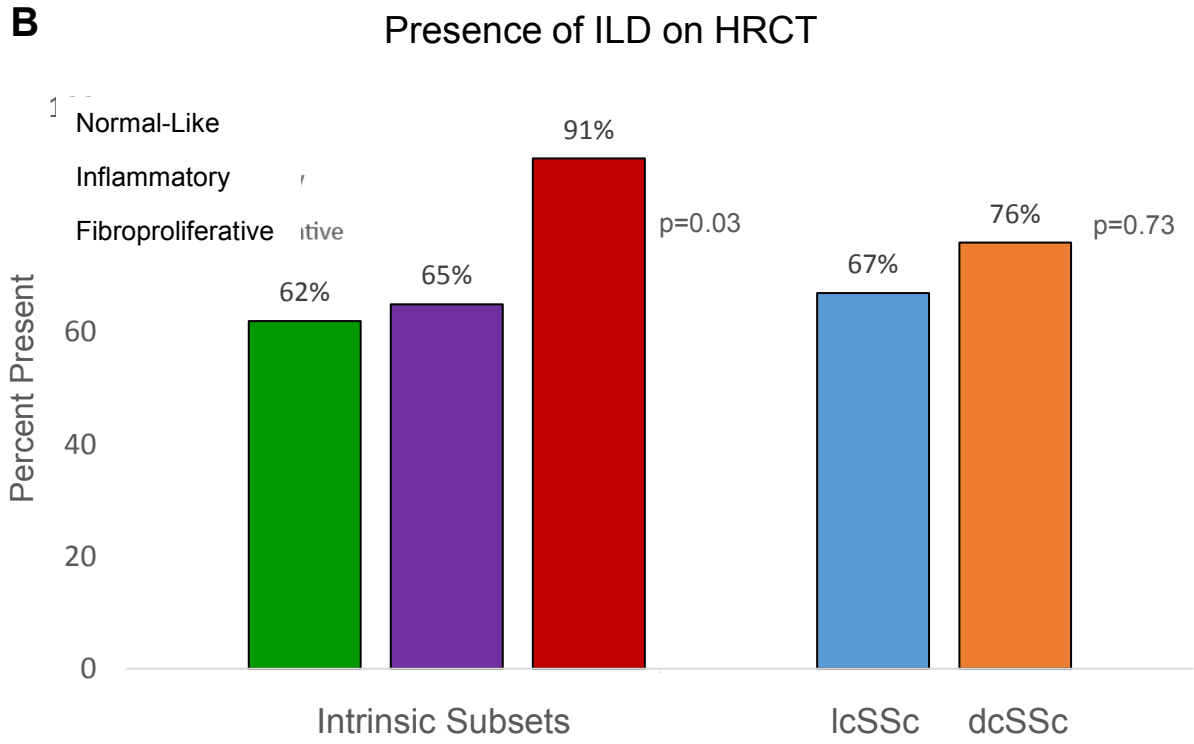


Table 1.

Clinical Characteristic	Healthy Participants (n=58)	All SSc (n=165)	Limited cutaneous SSc (n=50)	Diffuse cutaneous SSc (n=115)	P-value (HP vs. SSc)	P-value (dcSSc vs. lcSSc)
Age, years (mean ± SD)	42 ± 13	50 ± 11	49 ± 12	51 ± 11	<0.001 [a]	0.26 [a]
Sex, female, n (%)	42 (72)	137 (83)	44 (88)	93 (81)	0.08 [b]	0.26 [b]
Race, White, n (%)	47 (81)	125 (77)	43 (86)	82 (71)	0.41 [b]	0.04 [b]
MRSS, mean ± SD		15.2 ± 10.2	5.6 ± 2.9	19.4 ± 9.4		<0.001 [a]
SSc disease duration, mo, median [IQR]		23 [46]	49 [92]	18 [32]		<0.001 [c]
ANA positive, n (%)		153 (93)	47 (94)	106 (92)		>0.999 [b]
ANA pattern, n (%)						<0.001 [b]
Centromere		14 (10)	11 (27)	3 (3)		
Speckled		57 (39)	13 (32)	44 (46)		
Homogenous		40 (27)	13 (32)	27 (28)		
Nucleolar		25 (17)	4 (10)	21 (22)		
Autoantibodies, n (%)						
ACA		16 (10)	13 (26)	3 (3)		<0.001 [b]
Scl-70		44 (27)	10 (20)	34 (30)		0.25 [b]
RNA Pol III		50 (30)	2 (4)	48 (43)		<0.001 [b]
Platelet count (x10^9/L), median [IQR]		297 [96]	278 [93]	307 [92]		0.004 [c]
ESR (mm/h), median [IQR]		18 [32]	13 [24]	20 [32]		0.63 [c]
CRP (mg/L), median [IQR]		0.50 [0.55]	0.25 [0.35]	0.60 [0.65]		0.002 [c]

[a] Student's t-test, [b] Pearson's Chi-squared test, [c] Wilcoxon rank sum test. SSc = systemic sclerosis, SD = Standard deviation, IQR = interquartile range, ANA = anti-nuclear antibody, ACA = anticentromere, Scl-70 = anti-topoisomerase I, RNA pol III = anti-RNA polymerase III, ESR = erythrocyte sedimentation rate, CRP = C-reactive protein.

Table 2.

Clinical Characteristic	Intrinsic Subset Classification				
	Normal-like (n=58)	Limited (n=2)	Inflammatory (n=72)	Fibro-proliferative (n=33)	P-value
Age, years (mean ± SD)	49.5 ± 11.8	56.5 ± 2.1	52.7 ± 9.8	47.3 ± 12.9	0.09 [a]

Sex, female, n (%)	51 (88)	2 (100)	56 (78)	28 (85)	0.45 [b]
Race, n (%)					
White	47 (81)	2 (100)	52 (72)	24 (73)	0.22 [b]
Black	6 (10)	0 (0)	7 (10)	8 (24)	
Asian	2 (4)	0 (0)	3 (4)	0 (0)	
Hispanic	1 (2)	0 (0)	9 (13)	1 (3)	
SSc subtype, n (%)					
Limited cutaneous	32 (64)	2 (4)	7 (14)	9 (18)	<0.001 [b]
Diffuse cutaneous	26 (23)	0 (0)	65 (57)	24 (21)	
SSc disease duration, mo, median [IQR]	45 [84]	121 [8]	15 [20]	21[35]	<0.001 [a]
ANA, positive, n (%)	54 (93)	2 (100)	68 (94)	29 (88)	0.54 [b]
ANA pattern, n (%)					0.10 [b]
Centromere	9 (19)	1 (0)	3 (5)	2 (7)	
Speckled	20 (42)	2 (50)	28 (48)	8 (29)	
Homogenous	15 (31)	2 (50)	14 (24)	10 (37)	
Nucleolar	4 (8)	0 (0)	14 (24)	7 (26)	
Autoantibodies, n (%)					
ACA	11 (19)	0 (0)	3 (4)	2 (6)	0.04 [b]
Scl-70	15 (27)	0 (0)	14 (19)	15 (46)	0.04 [b]
RNA Pol III	8 (14)	0 (0)	38 (53)	4 (12)	<0.001 [b]
Platelet count (x10^9/L), median [IQR]	287 [89]	217 [24]	326 [132]	276 [76]	<0.001 [c]
ESR (mm/h), median [IQR]	18 [31]	30 [4]	19.5 [39]	14 [23]	0.82 [c]
CRP (mg/L), median [IQR]	0.5 [0.6]	0.25 [0]	0.6 [0.8]	0.5 [0.5]	0.20 [c]

[a] Analysis of variance [b] Fisher's exact test, [c] Kruskal-Wallis test. SD = Standard deviation, SSc = systemic sclerosis, IQR = interquartile range, ANA = anti-nuclear antibody, ACA = anticentromere, Scl-70 = anti-topoisomerase I, RNA pol III = anti-RNA polymerase III, ESR = erythrocyte sedimentation rate, CRP = C-reactive protein.

Table 3.

	Intrinsic Subset Classification					**Traditional Cutaneous Classification**		
Mean ± SD or n (%)	Normal-like	Limited	Inflammatory	Fibro-Proliferative	P-value	Limited cutaneous SSc	Diffuse cutaneous SSc	P-value
Distribution	N=58	N=2	N=72	N=33		N=50	N=115	

mRSS	8.6 ± 5.9	5.0 ± 2.8	22.1 ± 9.9	12.3 ± 7.1	<0.001[a]	5.7 ± 3.0	19.4 ± 9.4	<0.001 [b]
PFT	N=53	N=2	N=69	N=33		N=45	N=112	
FVC%	82 ± 18	77 ± 46	79 ± 18	75 ± 16	0.41 [a]	84 ± 18	78 ± 18	0.04 [b]
FEV1%	82 ± 17	86 ± 16	82 ± 14	76 ± 16	0.42 [a]	84 ± 17	80 ± 15	0.13 [b]
TLC%	89 ± 19	94 ± 51	88 ± 17	82 ± 18	0.35 [a]	92 ± 19	85± 19	0.06 [b]
DLCO%	69 ± 20	50 ± 35	72 ± 20	63 ± 22	0.15 [a]	71 ± 21	68 ± 21	0.38 [b]
Chest HRCT	N=35	N=2	N=48	N=23		N=31	N=77	
Fibrosis Score	3.6 ± 4.1	6.5 ± 4.9	2.2 ± 2.9	4.3 ± 4.6	0.08 [a]	3.6 ± 3.6	3.0 ± 3.9	0.47 [b]
GGO Score	6.2 ± 5.4	5.5 ± 2.1	4.3 ± 4.4	7.4 ± 5.4	0.09 [a]	5.7 ± 5.0	5.5 ± 5.1	0.87 [b]
Total Lung Score	9.2 ± 8.6	12.0 ± 7.1	6.5 ± 6.4	11.7 ± 9.2	0.06 [a]	8.8 ± 7.8	8.5 ± 8.1	0.86 [b]
Radiographic ILD present	21 (62)	2 (100)	31 (65)	21 (91)	0.04 [c]	21 (67)	31 (76)	0.73 [d]
Echo	N=42	N=2	N=48	N=26		N=38	N=80	
LVEF (%)	63 ± 7	60 ± 3	62 ± 5	62 ± 5	0.85 [a]	63 ± 7	62 ± 5	0.47 [b]
LV mass (g/m^2)	77 ± 18	62 ± 2	81 ± 25	79 ± 18	0.60 [a]	76 ± 16	79 ± 23	0.43 [b]
TAPSE (cm)	2.2 ± 0.4	2.0 ± 0.1	2.2 ± 0.4	2.1 ± 0.4	0.68 [a]	2.1 ± 0.4	2.2 ± 0.5	0.71 [b]
RV FAC (%)	43 ± 5	43 ± 7	44 ± 5	43 ± 5	0.99 [a]	42 ± 6	44 ± 5	0.20 [b]
PASP (mm Hg)	32 ± 9	41 ± 16	31 ± 10	28 ± 8	0.15 [a]	32 ± 11	30 ± 9	0.39 [b]
DD, n (%)	23 (55)	0 (0.0)	19 (39)	9 (35)	0.20 [c]	17 (45)	34 (43)	0.82 [d]

[a] Analysis of variance, [b] Student's t-test, [c] Fisher's exact test, [d] Pearson's Chi-squared test. SD = standard deviation. mRSS = modified Rodnan skin score, PFT = pulmonary function tests, FVC= forced vital capacity, FEV1= forced expiratory volume in one second, TLC=total lung capacity, DLCO=diffusion for carbon monoxide, HRCT=high-resolution computed tomography, GGO = ground glass opacities, ILD = interstitial lung disease, Echo = echocardiogram, LVEF = left ventricular ejection fraction, LV = left ventricle, TAPSE=tricuspid annular plane systolic excursion, RV FAC=right ventricular fractional area change, PASP = pulmonary artery systolic pressure, DD = diastolic dysfunction.

Table 4

Mean ± SD or n (%)	**lcSSc (N=50)**					**dcSSc* (N=115)**			
	Normal-like	Limited	Inflammatory	Fibro-Proliferative	P-value	Normal-like	Inflammatory	Fibro-Proliferative	P-value
Distribution	32 (64.0)	2 (4.0)	7 (14.0)	9 (18.0)	<0.001 [b]	26 (22.6)	65 (56.5)	24 (20.9)	<0.001 [b]

Skin Score	N=32	N=2	N=7	N=9		N=26	N=65	N=24	
mRSS	5.8 ± 3.2	5.0 ± 2.0	6.1 ± 2.8	5.1 ± 2.2	0.77 [a]	12.2 ± 6.3	23.8 ± 8.8	15 ± 6.2	<0.001 [a]
Antibodies	N=32	N=2	N=7	N=9		N=26	N=65	N=24	
ACA	10 (20.0)	0 (0.0)	1 (2.0)	2 (4.0)	0.02 [b]	1 (0.8)	2 (1.7)	0 (0.0)	0.79 [b]
Scl-70	7 (14.0)	0 (0.0)	1 (2.0)	2 (4.0)	0.25 [b]	8 (6.9)	13 (11.3)	13 (11.3)	0.007 [b]
RNA pol III	1 (2.0)	0 (0.0)	1 (2.0)	0 (0.0)	0.79 [b]	7 (6.1)	37 (32.1)	4 (3.5)	<0.001 [b]
PFT	N=28	N=2	N=6	N=9		N=24	N=64	N=24	
FVC%	83 ± 18	50 ± 25	85 ± 14	83 ± 10	0.96 [a]	79 ± 16	79 ± 18	70 ± 15	0.11 [a]
DLCO%	71 ± 19	41 ± 11	82 ± 20	66 ± 19	0.13 [a]	66 ± 19	70 ± 19	61 ± 22	0.13 [a]
Chest HRCT	N=17	N=2	N=6	N=5		N=18	N=42	N=18	
Fibrosis Score	3.5 ± 3.9	6.5 ± 3.5	1.3 ± 1.5	3.6 ± 2.6	0.30 [a]	3 ± 3.9	2.35 ± 3.0	4.5 ± 4.9	0.05 [a]
GGO Score	5.8 ± 5.9	5.5 ± 1.5	3.2 ± 4.5	5.0 ± 2.4	0.49 [a]	5.4 ± 4.5	4.45 ± 4.6	8.1 ± 5.5	0.007 [a]
Total lung score	8.5 ± 8.9	12.0 ± 5.0	4.5 ± 3.9	8.6 ± 4.45	0.45 [a]	8.4 ± 7.9	6.81 ± 6.6	12.5 ± 9.7	0.009 [a]
Radiographic ILD present	10 (62)	2 (100)	4 (67)	5 (100)	<0.001 [b]	11(68)	27 (64)	16 (89)	0.04 [b]
Echo	N=24	N=2	N=6	N=6		N=18	N=42	N=20	
TAPSE (cm)	2.1 ± 0.4	2.01 ± 0.1	2.2 ± 0.4	2.1 ± 0.1	0.89 [a]	2.2 ± 0.5	2.19 ± 0.4	2.0 ± 0.4	0.14 [a]
PASP (mmHg)	34 ± 11	41 ± 11	30 ± 10	27 ± 3	0.19 [a]	31 ± 4	31 ± 9	29 ± 9	0.48 [a]
Laboratories	N=30	N=2	N=6	N=9		N=25	N=62	N=23	
Platelet count (10^9/L)	275 ± 60	217 ± 12	278 ± 85	278 ± 45	0.97 [c]	283 ± 49	346 ± 104	269 ± 55	<0.001 [c]
ESR (mm/h)	24 ± 20	30 ± 2	16 ± 18	30 ± 34	0.47 [c]	22 ± 14	27 ± 22	22 ± 23	0.33 [c]
CRP (mg/L)	0.5 ± 0	0.5 ± 0	0.5 ± 0	0.6 ± 0.26	0.06 [c]	1.7 ± 3.1	4.6 ± 13.9	0.9 ± 0.7	0.11 [c]

*Within dcSSc subgroup, no patients were classified as Limited IS. [a] Analysis of variance, [b] Fisher's exact test, [c] Kruskal-Wallis test. lcSSc = limited cutaneous SSc; dcSSc = diffuse cutaneous SSc; mRSS= modified Rodnan skin score; ACA = anticentromere serum antibody; Scl-70 = anti-topoisomerase I serum antibody; RNA pol III = RNA polymerase III serum antibody; PFT = pulmonary function tests; FVC % = forced vital capacity % predicted; DLCO % = diffusion capacity for carbon monoxide % predicted; HRCT = high-resolution computed tomography; GGO = ground glass opacities; ILD = interstitial lung disease; echo = echocardiography; TAPSE = tricuspid annular plane systolic excursion; PASP = pulmonary artery systolic pressure; ESR = erythrocyte sedimentation rate; CRP = c-reactive protein.

Table 5.

	Skin and Pulmonary Outcome Variables				
Predictor Variables	**mRSS** β (95% CI) [a] N=163	**Total Lung Score** β (95% CI) [a] N=106	**FVC%** β (95% CI) [a] N=155	**DLCO%** β (95% CI) [a] N=155	**Radiographic ILD** OR (95% CI) [b] N = 105
Intrinsic subset:					

Inflammatory	8.15 (5.24, 11.06)	-1.47 (-5.03, 2.09)	-0.67 (-8.05, 6.72)	2.89 (-5.81, 11.6)	1.33 (0.45, 3.94)
Proliferative	1.30 (-1.89, 4.49)	0.16 (-4.01, 4.33)	-4.93 (-12.89, 3.03)	-3.96 (-13.38, 5.47)	5.58 (1.29, 46.90)
Normal-like (ref)	—	—	—	—	—
SSc Subtype: Diffuse vs. Limited Cutaneous	9.42 (6.54, 12.30)	1.85 (-1.69, 5.38)	-5.61 (-13.18, 1.95)	-6.36 (-15.3, 2.57)	1.32 (0.46, 3.68)
SSc Disease duration: >36mo vs. ≤36mo	0.48 (-1.95, 2.91)	4.34 (1.14, 7.55)	-1.93 (-8.01, 4.16)	-6.09 (-13.32,1.14)	1.37 (0.50, 3.66)
Scl-70: Positive vs. Negative	2.19 (-0.33 ,4.71)	7.47 (4.12, 10.82)	-2.12 (-9.3, 5.06)	-8.78 (-17.25, -0.32)	13.19 (2.4, 403.4)
RNA Pol III: Positive vs. Negative	2.62 (-0.09, 5.34)	1.39 (-1.90, 4.68)	-1.29 (-8.75, 6.16)	1.19 (-7.61, 9.99)	1.57 (0.49, 5.08)
ACA: Positive vs. Negative	—	—	5.9 (-4.49, 16.29)	2.75 (-9.39, 14.89)	—
Immunosuppressant (Current): Yes vs. No	0.08 (-3.13, 3.3)	2.04 (-1.66, 5.74)	-2.8 (-10.94, 5.34)	-4.77 (-14.76, 5.22)	—
Smoker (Ever): Yes vs. No	1.16 (-1.9, 4.22)	0.15 (-3.39, 3.69)	9.05 (1.47, 16.62)	0.85 (-8.29, 9.99)	—

[a] Multiple Linear Regression, [b] Exact Logistic Regression. OR = odds ratio, mRSS= modified Rodnan skin score, FVC % = forced vital capacity % predicted, DLCO % = diffusion capacity for carbon monoxide % predicted, ILD = interstitial lung disease, SSc = systemic sclerosis, Scl-70 = anti-topoisomerase I serum antibody, RNA pol III = RNA polymerase III serum antibody, ACA = anticentromere serum antibody, '—' denotes variables that were not included to permit model convergence, ref = reference variable.