Arthritis & Rheumatology
Vol. 76, No. 11, November 2024, pp 1635–1644
DOI 10.1002/art.42939

American College of Rheumatology
Empowering Rheumatology Professionals

Development and Initial Validation of the Novel Scleroderma Clinical Trials Consortium Activity Index

Laura Ross,[1] Dylan Hansen,[2] Susanna Proudman,[3] Dinesh Khanna,[4] Ariane L. Herrick,[5] Wendy Stevens,[2] Murray Baron,[6] and Mandana Nikpour;[7] on behalf of Australian Scleroderma Interest Group (ASIG), Canadian Scleroderma Research Group (CSRG), the Scleroderma Clinical Trials Consortium Activity Index Working Group

Objective. Accurate measurement of disease activity in systemic sclerosis (SSc) remains a significant clinical challenge. The Scleroderma Clinical Trials Consortium (SCTC) convened an Activity Index (AI) Working Group (WG) to develop a novel measure of disease activity (SCTC-AI).

Methods. Using consensus methodology, we developed a conceptual definition of disease activity. Literature review and expert consensus generated provisional SCTC-AI items, which were reduced by Delphi survey. Provisional items were weighted against a combined endpoint of morbidity and mortality, using time-dependent Cox proportional hazards regression analysis of the Australian Scleroderma Cohort Study (ASCS) (n = 1,254). External validation of the SCTC-AI was performed using data collected from 1,103 Canadian Scleroderma Research Group Study participants.

Results. Disease activity in SSc was defined using consensus methodology as "aspects of disease that are reversible, or can be arrested, with time and, or effective therapy." One-hundred and forty-one provisional SCTC-AI items were generated and reduced using three rounds of Delphi survey and statistical reduction and weighting, against mortality and quality of life measures, yielding a final 24-item index with a maximum possible score of 140. Survival analysis in an external cohort showed a graded relationship between disease activity scores and survival ($P < 0.01$).

Conclusion. We present a novel instrument to quantify the burden of disease activity in SSc. We have employed a rigorous consensus-based process in combination with data-driven methods to develop an instrument that has face, content, and criterion validity. Further work is required to fully validate and confirm the construct and discriminative validity of the SCTC-AI.

INTRODUCTION

Systemic sclerosis (SSc) is a multisystem disorder characterized by a complex aetiopathogenesis of microvasculopathy, abnormal fibrosis, and inflammation.[1] There is marked clinical heterogeneity in the presentation of SSc and an absence of clear episodes of remission and relapse. Consequently, there are significant challenges in the accurate measurement of disease status in SSc.[2–5]

Measures of individual organ systems have been commonly used as primary outcomes in SSc randomized clinical trials (RCTs). However, the limitations of using single-organ outcome measures in a systemic disease are well recognized,[6] and accordingly, there has been significant effort to develop SSc-specific multisystem instruments.[7,8] The Scleroderma Clinical Trials Consortium (SCTC) convened an Activity Index (AI) Working Group (WG) to develop a novel multisystem measure of disease activity (SCTC-AI). Two previous efforts to develop an AI have used data from the

Supported by Scleroderma Clinical Trials Consortium Working Group grants, a Betty Z Benedict Award from the Scleroderma Clinical Trials Consortium and Scleroderma Research Foundation, and private philanthropic donations. Dr Nikpour's work was supported by the National Health and Medical Research Council of Australia Investigator Grant (GNT-1176538).

[1]Laura Ross, MBBS, PhD, Mandana Nikpour, MBBS, PhD: University of Melbourne and St Vincent's Hospital Melbourne, Fitzroy, Victoria, Australia; [2]Dylan Hansen, BBiomedSc, MBiostat, Wendy Stevens, MBBS: St Vincent's Hospital Melbourne, Fitzroy, Victoria, Australia; [3]Susanna Proudman, MBBS: Royal Adelaide Hospital and University of Adelaide, South Australia, Australia; [4]Dinesh Khanna, MD: University of Michigan, Ann Arbor; [5]Ariane L. Herrick, MD: The University of Manchester, Northern Care Alliance NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, United Kingdom; [6]Murray Baron, BBS, MDCM: Jewish General Hospital, McGill University, Montreal, Quebec, Canada; [7]Mandana Nikpour, MBBS, PhD: University of Melbourne and St Vincent's Hospital Melbourne, Fitzroy, Victoria, Australia; Sydney Musculoskeletal Research Flagship Centre (University of Sydney), University of Sydney School of Public Health, Sydney, New South Wales, and Royal Prince Alfred Hospital, Camperdown, New South Wales, Australia.

Drs Barron and Nikpour are co-last authors and contributed equally to this work.

Additional supplementary information cited in this article can be found online in the Supporting Information section (http://onlinelibrary.wiley.com/doi/10.1002/art.42939).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.42939.

Address correspondence via email to Mandana Nikpour, MBBS, PhD, at mandana.nikpour@sydney.edu.au.

Submitted for publication March 3, 2024; accepted in revised form June 20, 2024.

European Scleroderma Trials and Research Group (EUSTAR) cohort.[9,10] Existing indices have been shown to predict the development of increased disease severity.[10–12] Yet, concern remains over the lack of proven face, content, and construct validity of these instruments.[10,13,14] Recent analysis of the performance of the existing SSc activity indices failed to identify an existing gold-standard assessment of disease activity in SSc.[15] Consequently, these indices have not been widely applied in SSc RCTs. To build on these previous efforts, the aim of the SCTC-AI WG was to develop an instrument to measure disease activity in SSc that has face, content, and construct validity.

PATIENTS AND METHODS

Investigators and patient and public involvement statement. The SCTC-AI WG was co-chaired by Professor Murray Baron and Professor Mandana Nikpour. Membership of the WG was open to any SCTC member. WG members included 39 physicians representing Australasia, the Americas, and Europe and four experienced patient partners from Australia and Canada (see Supplementary Index 1 for full list of WG members). Patient-research partners were involved in the development of the conceptual definition of disease activity in SSc, selection of domains of disease to measure in the SCTC-AI, and nomination of provisional SCTC-AI items.

Development of the SCTC-AI. The SCTC-AI was developed using consensus and data-driven methodology, an accepted approach to development of novel multisystem outcome measures[7,8] (Figure 1). Detailed methods are provided in Supplementary Index 2. Briefly, (i) Consensus was reached as to the conceptual definition of SSc disease activity. All instrument items were required to be a measure of this agreed construct. (ii) An online Delphi exercise was performed to define the SCTC-AI disease domains and subdomains. (iii) Candidate items were generated from systematic literature review and expert input. (iv) Candidate items were reduced using the results of an online two-round Delphi exercise. In all Delphi exercises, an a priori threshold of 70% agreement was applied for items to be included or excluded. Items that did not meet the threshold for inclusion after two rounds of survey were omitted. (v) Item reduction and weighting of individual items using data derived from the Australian Scleroderma Cohort Study (ASCS). (vi) Seven provisional SCTC-AI items were not collected by ASCS. To derive weightings for these items, we performed a discrete choice experiment (DCE) to generate relative item weighting. 1000minds decision-making and conjoint analysis software was used to perform the DCE. In this step, two provisional SCTC-AI items were presented to respondents at a time, and respondents were asked to rate which item represented high disease activity. This survey included all seven provisional items not collected as part of the ASCS, in addition to relative forced vital capacity (FVC) decrease $\geq 5\%$. Conjoint analysis provided a relative ranking of these seven items, in reference to FVC decrease $\geq 5\%$. This relative ranking was used to calculate a weighted score of the item in the SCTC-AI. (vii) Validation studies were performed using data from ASCS (derivation cohort) and the Canadian Scleroderma Research Group (CSRG) cohort study (validation cohort).

Validation of the AI. The performance of the SCTC-AI was evaluated in the derivation cohort (ASCS). The ASCS is a prospective multicenter cohort study established in 2007 with up to

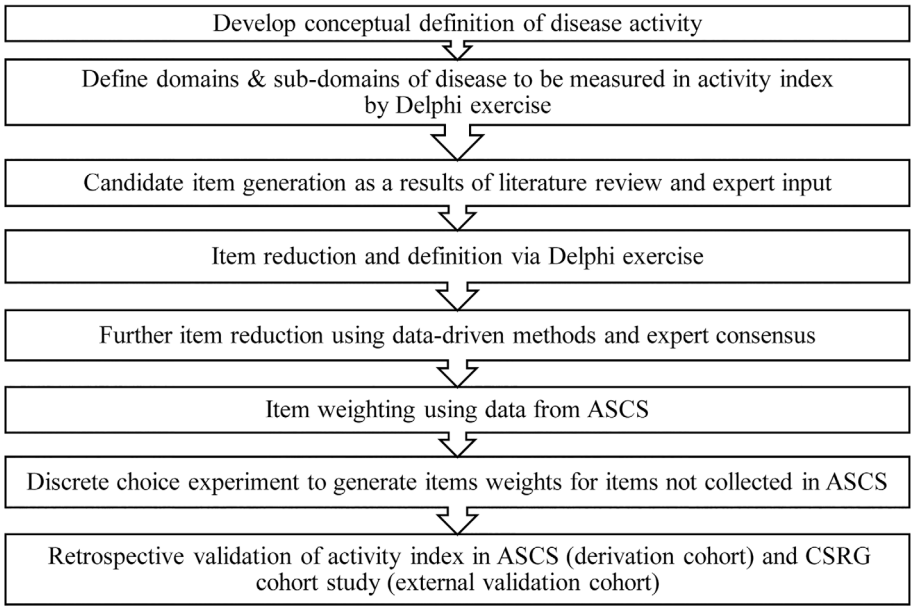

Figure 1. Expert consensus and data-driven methods to develop the Scleroderma Clinical Trials Consortium Activity Index. ASCS, Australian Scleroderma Cohort Study; CSRG, Canadian Scleroderma Research Group.

15 years of participant follow-up. In absence of a gold-standard measure of disease activity, using receiver operator curve (ROC) analyses, we calculated the area under the curve (AUC) of SCTC-AI scores against endpoints of (i) death; (ii) physical function measured by the Health Assessment Questionnaire–Disability Index (HAQ-DI); (iii) morbidity, assessed by the Short-Form-36 (SF-36) physical component scale score (PCS) (this was selected as the measure of morbidity because it quantifies the overall physical burden of illness, as reported by patients; therefore, it was felt to be the best surrogate measure of overall burden of disease to patients, in the absence of validated measure of disease morbidity); (iv) damage accrual, defined as ≥1 point increase in SCTC-DI score[7]; and (v) disease severity measured by Medsger Severity Scale.[4] Analyses were repeated in a subgroup of patients recruited within four years of disease onset to further evaluate the sensitivity and specificity of the SCTC-AI in an incident cohort of patients. A threshold of four years disease duration was applied to ensure both diffuse and limited cutaneous subtype was adequately represented in this incident cohort. Study participants were classified into high, medium, and low activity groups according to baseline SCTC-AI scores. High, medium, and low activity groups were defined by a baseline SCTC-AI score within the highest, middle, and lowest tertile, respectively. Kaplan-Meier survival estimates comparing survival in each of these groups were calculated. The patient characteristics of each group were compared using chi-square test for categorical variables or Kruskal-Wallis test for continuous variables. Multivariable multinomial logistic regression modeling was performed in the early disease cohort to identify a clinical phenotype of patients most likely to have high disease activity, selecting variables based on clinical expertise and disease manifestations found to be significantly more likely to be present in the high activity group in univariable analysis.

External validation. External validation was performed using data from the CSRG cohort study (validation cohort). The CSRG is a multicenter prospective cohort study of patients with SSc established in 2004 with annual data collection and study protocol comparable to the ASCS.[7] ROC curve analysis was performed to evaluate the performance of the SCTC-AI to predict the outcomes of death, physical function, morbidity as measured by SF-36 PCS, damage accrual, and disease severity. The AUC of the 2017 EUSTAR activity index[10] against these same endpoints was calculated to provisionally compare the performance of the SCTC-AI and the 2017 EUSTAR index. Kaplan-Meier survival analysis was used to compare the survival of those with high, moderate, and low disease activity, using the thresholds for these categories derived from the ASCS cohort.

Study data were collected and managed using REDCap electronic data capture tools hosted at the University of Melbourne. All statistical analyses were performed using STATA 15.1 (StataCorp). Ethical approval for the study was provided by St Vincent's Hospital Human Research Ethics Committee (LRR 012/21).

RESULTS

Defining the concept of disease activity. Nineteen (79%) WG members responded to the conceptual definition of disease activity survey (Supplementary Index 3). A provisional definition was developed and distributed to 31 WG members with iterative changes made based on the feedback of 21 (87%) recipients. The definition of disease activity in SSc, as determined by the SCTC-AI WG is:

> "Disease activity in SSc refers to aspects of disease, attributable to SSc, that are potentially reversible, or can be arrested, with time and/or effective therapy. Disease activity may be associated with morbidity, and uncontrolled activity may lead to organ dysfunction and mortality."

Defining disease domains and provisional AI items. Thirty-four WG members were invited to participate in a two-round Delphi exercise. Agreement was reached that the SCTC-AI should include the following nine disease domains: skin, vascular, musculoskeletal, respiratory (including pulmonary arterial hypertension), cardiac, gastrointestinal, renal, laboratory investigation results, and constitutional symptoms. In a second Delphi exercise, 141 provisional AI items were presented to the WG for consideration for inclusion, with a total of 86 items excluded as a result of the Delphi exercise and WG consensus. (Detailed results in Supplementary Indices 4–6.)

Removal of items by consensus. Provisional AI items were reviewed by the WG to consider the face validity of each item as a measure of disease activity. By consensus the WG elected to remove new-onset or worsening joint contractures, gangrene, proximal weakness, and swollen joint count from the SCTC-AI. An overrepresentation and clinical collinearity of the provisional cardiac items was observed. Therefore, the number of cardiac items were reduced to include acute or subacute systolic dysfunction, active myocarditis, acute pericarditis, and elevated troponin and B-type natriuretic peptide (BNP). Consideration was given to the inclusion of two inflammatory biomarkers, erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP). Given the superior feasibility of CRP testing, ESR was omitted, and CRP was included as a sole biomarker of inflammation.

Modification of definition of change in respiratory items. A definition of a relative FVC (percent predicted) change of –10% was nominated by the WG, recognizing that this threshold of change is of prognostic importance.[16–19] However, subsequent to the Delphi exercise, the successful trial of nintedanib was published,[20] showing a 5% worsening of FVC was associated

with poorer outcomes. Furthermore, the minimal clinical important difference of FVC in SSc is –4.8%.[21] Therefore, a –5% FVC change was considered a more appropriate measure of activity. The definition of change in FVC was adjusted to be a decrease of ≥5% over a period of six months. Items that suggested regular use of high-resolution computed tomography (HRCT) to measure radiographic progression of interstitial lung disease (ILD) were thought to potentially invite unnecessary exposure to ionizing radiation, particularly for patients enrolled in observational studies or RCTs that include participants without ILD. Therefore, it was considered more clinically appropriate to reword the respiratory domain items to be change in FVC and/or diffuse capacity of the lung for carbon monoxide (DLCO) or a progression of ILD on HRCT rather than have progressive radiographic changes as a stand-alone item.

The WG was asked to consider what change in extent of ILD on HRCT would be considered disease activity, including consideration of quantitative computed tomography (CT) measures of ILD. Quantitative CT measures were excluded by the WG because of the current lack of feasibility for the use of these measures in the assessment of SSc-ILD. A threshold of a 10% increase in extent of lung volume on HRCT was based on the nomination of this degree of change being considered significant and indicative of activity by members of the WG in the Delphi exercise.

Item weighting: Time-dependent Cox hazards regression. Data from 1,765 ASCS participants across 8,958 study visits were available for analysis (baseline derivation cohort characteristics in Supplementary Index 7). Results of the univariable relationship between SCTC-AI items and morbidity and mortality are presented in Table 1. Integer weights were calculated using the combined coefficient of the Cox regression analyses. Synovitis and tendon friction rub did not meet the rules for inclusion in the AI with either negative or nonsignificant relationships

Table 1. SCTC-AI item weighting and coefficient calculations*

Variable	Mortality coefficient (95% CI)	*P* value	Morbidity[a] coefficient (95%CI)	*P* value	Combined coefficient
Domain: Skin					
mRSS increase by at least 5 points	1.36 (0.82–1.90)	<0.001	1.18 (1.00–1.35)	<0.001	1.27
New areas of skin involvement	1.15 (0.86–1.43)	<0.001	1.15 (1.05–1.24)	<0.001	1.15
Skin worse in the last month (PRO)	0.01 (−0.31 to 0.34)	0.93	0.29 (0.12–0.45)	<0.001	0.15
Domain: Vascular					
Raynaud phenomenon worse in the last month (PRO)	−0.32 (−0.62 to −0.02)	0.03	0.25 (0.12–0.38)	<0.001	0.25
Digital ulcers present	0.59 (0.33–0.85)	<0.001	0.02 (0.14–0.18)	0.77	0.59
Higher number of digital ulcers	0.07 (0.00–0.15)	0.07	−0.01 (−0.05 to 0.04)	0.77	0.07
Critical digital ischemia	N/A	–	N/A	–	–
Domain: Musculoskeletal					
Synovitis present	−0.42 (−0.82 to −0.03)	0.03	0.01 (−0.16 to 0.19)	0.88	–
Tendon friction rubs	−0.05 (−0.76 to 0.66)	0.90	0.21 (−0.11 to 0.53)	0.18	–
Myositis	1.28 (0.72–1.84)	<0.001	0.96 (0.60–1.31)	<0.001	1.12
Domain: Respiratory					
FVC decrease by 5%	1.65 (1.36–1.93)	<0.001	1.06 (0.96–1.16)	<0.001	1.36
DLCO decrease by 10%	1.63 (1.31–1.95)	<0.001	1.06 (0.95–1.18)	<0.001	1.35
Newly diagnosed ILD of any extent on HRCT	0.97 (0.75–1.20)	<0.001	0.37 (0.22–0.53)	<0.001	0.67
At least 10% increased extent of ILD on HRCT	1.63 (1.09–2.17)	<0.001	0.84 (0.57–1.11)	<0.001	1.24
More breathless in the last month (PRO)	1.03 (0.79–1.28)	<0.001	0.27 (0.11–0.42)	<0.001	0.65
Domain: Pulmonary vascular disease					
New diagnosis PAH	0.48 (0.09–0.86)	0.02	−0.89 (−1.20 to −0.56)	<0.001	0.48
Domain: Renal					
Active scleroderma renal crisis	N/A	–	N/A	–	–
Domain: Cardiac					
Acute or subacute systolic dysfunction	0.91 (0.46–1.35)	<0.001	0.92 (0.78–1.06)	<0.001	0.92
Myocarditis	N/A	–	N/A	–	–
Pericarditis	N/A	–	N/A	–	–
Elevated troponin	2.38 (1.61–3.16)	<0.001	0.50 (0.09–1.09)	0.10	1.44
Elevated BNP/NT-proBNP	2.24 (1.39–3.10)	<0.001	0.40 (0.11–0.68)	0.01	1.32
Domain: Gastrointestinal					
Patient-reported worsening of GI symptoms	N/A	–	N/A	–	–
Domain: Laboratory values					
Abnormal CRP value	1.02 (0.76–1.29)	<0.001	0.08 (0.05–0.22)	0.23	1.02
Domain: Constitutional symptoms					
Weight decrease by 10%	1.13 (0.88–1.38)	<0.001	0.14 (0.02–0.30)	0.09	1.13

* N/A indicates item not collected in Australian Scleroderma Cohort Study. BNP, B-type natriuretic peptide; CI, confidence interval; CRP, C-reactive protein; DLCO, diffuse capacity of the lung for carbon monoxide; FVC, forced vital capacity; GI, gastrointestinal; HRCT, high-resolution computed tomography; ILD, interstitial lung disease; mRSS, modified Rodnan Skin Score; NT-proBNP, N-terminal pro B-type natriuretic peptide; PAH, pulmonary arterial hypertension; PRO, patient-reported outcome; SCTC-AI, Scleroderma Clinical Trials Consortium Activity Index
[a] Morbidity defined as a Short-Form 36 Physical Component Score below population median value.

with death and morbidity. However, investigators felt it important to include these two disease manifestations in the final SCTC-AI given the strong consensus from the WG that they be included as measures of disease activity, so they were included in the DCE for consensus weighting along with the items not collected as part of the ASCS protocol.

Item weighting: DCE. Eighteen (58%) complete responses to the online DCE survey were received. The seven unweighted items were ranked, relative to FVC decrease ≥5%, in the following order from highest to lowest activity: active scleroderma renal crisis, active myocarditis, critical digital ischemia, tendon friction rub, acute pericarditis, synovitis, and patient-reported worsening of gastrointestinal symptoms. Individual item weights were derived by multiplying the relative importance of each item by the statistically derived weight of FVC decrease ≥5% (8 points). The SCTC-AI is presented in Table 2.

Table 2. Scleroderma Clinical Trials Consortium Activity Index*

Item	Weighted score
Domain: Skin	
≥5-point increase in mRSS within the past 3 months	7
New areas of skin involvement (thickening) in any of mRSS assessment area (irrespective of total mRSS score) within the past 3 months	7
If previous mRSS unavailable: Patient-reported: Has your skin worsened in past month? (yes/no)	1
Domain: Vascular	
Current, active digital tip ulcer defined as loss of epithelialization of any degree of the epidermis, the dermis and/or subcutaneous tissue, distal to or at the PIP joint of the hands of feet and thought not due to trauma	3
Extra score for multiple digital ulcers concurrently present	+1
Critical digital ischemia defined as acute onset (within 7 days) of persistent, painful discoloration of a digit, associated with reduced capillary refill and cool temperature due to microvascular complications of SSc and not due to atherosclerosis or thromboembolic disease	9
Patient-reported: Has your Raynaud phenomenon worsened in past month? (yes/no)	1
Domain: Musculoskeletal	
Current synovitis	6
Current tendon friction rub	7
Myositis defined as presence of proximal muscle weakness on examination in combination with elevated CK and/or aldolase with the presence of myositis demonstrated on either MRI, EMG, or muscle biopsy, present within the past 3 months	6
Domain: Respiratory	
Relative decrease of ≥ 5% of FVC within past 6 months AND, OR	8
Relative decrease of ≥ 10% of DLCO within past 6 months AND, OR	7
≥ 10% increased extent of ILD on HRCT within past 6 months	7
New diagnosis of ILD on HRCT	4
If previous PFT or HRCT results unavailable: Patient-reported worsening of breathlessness in past month (yes/no)	4
Domain: Pulmonary arterial hypertension	
Presence of pulmonary arterial hypertension detected for first time by right heart catheterization, diagnosed within past 3 months	3
Domain: Renal	
Active scleroderma renal crisis with evidence of ongoing active renal crisis indicated by presence of either; poorly controlled hypertension (not due to other etiology), rising serum creatinine and/or hemolysis on blood film at the time of clinical assessment	10
Domain: Cardiac	
Acute or subacute systolic dysfunction defined as either a new presentation of cardiac failure due to systolic dysfunction, not due to other causes, or cardiogenic shock	5
Active myocarditis indicated by the presence of acute inflammation on endomyocardial biopsy, acute inflammation on cardiac MRI or the presence of new-onset symptoms such as dyspnea, chest pain or palpitations and new-onset global or regional ventricular dysfunction on TTE, not attributable to another cause, in combination with elevated troponin	9
Acute pericarditis, diagnosed clinically on basis of clinical history and examination findings, supportive EKG changes and/or imaging evidence such as a pericardial effusion with or without concurrent elevated serum troponin	6
Elevated troponin (above upper limit of normal)	8
Elevated BNP/NT-proBNP (above upper limit of normal)	8
Domain: Gastrointestinal	
Patient-reported: Have your SSc gastrointestinal symptoms worsened in past month? (yes/no)	1
Domain: Laboratory findings	
Elevated CRP (above upper limit of normal)	6
Domain: Constitutional	
Decrease in weight by 10% or more within past 6 months	6

* BNP, B-type natriuretic peptide; CK, creatine kinase; CRP, C-reactive protein; DLCO, diffuse capacity of the lung for carbon monoxide; EKG, electrocardiogram; EMG, electromyography; FVC, forced vital capacity; HRCT, high-resolution computed tomography; ILD, interstitial lung disease; MRI, magnetic resonance imaging; mRSS, modified Rodnan Skin Score; NT-proBNP, N-terminal pro B-type natriuretic peptide; PFT, pulmonary function test; PIP, proximal interphalangeal joint; SSc, systemic sclerosis; TTE, transthoracic echocardiogram.

Table 3. Discriminatory capacity of disease activity indices*

Disease activity indices	SCTC-AI		2017 EUSTAR AI	
	ASCS cohort, AUC (95% CI)	CSRG cohort, AUC (95% CI)	ASCS cohort, AUC (95% CI)	CSRG cohort, AUC (95% CI)
Mortality	0.70 (0.65–0.75)	0.68 (0.64–0.73)	0.70 (0.66–0.73)	0.73 (0.69–0.77)
HAQ-DI	0.64 (0.62–0.66)	0.64 (0.62–0.66)	0.67 (0.66–0.69)	0.66 (0.64–0.68)
SF-36 PCS	0.65 (0.63–0.67)	0.64 (0.62–0.66)	0.67 (0.65–0.68)	0.65 (0.64–0.67)
SCTC-DI	0.68 (0.66–0.70)	0.67 (0.64–0.69)	0.62 (0.61–0.64)	0.55 (0.53–0.57)
MSS	0.70 (0.60–0.64)	0.65 (0.63–0.67)	0.81 (0.80–0.82)	0.74 (0.73–0.76)

* AI, activity index; ASCS, Australian Scleroderma Cohort Study; AUC, area under the curve; CI, confidence interval; CSRG, Canadian Scleroderma Research Group; DI, damage index; EUSTAR, European Scleroderma Trials and Research Group; HAQ-DI, Health Assessment Questionnaire–Disability Index; MSS, Medsger Severity Scale; SCTC, Scleroderma Clinical Trials Consortium; SF-36 PCS, Short-Form-36 Physical Component Scale.

Validation of SCTC-AI in the ASCS derivation cohort. The median baseline SCTC-AI score for the overall ASCS cohort was 6 (range 1–10) and in the incident cohort was 6 (range 1–10). The median follow-up period for ASCS participants was 5.02 years (range 1.99–9.13 years). Five items (active scleroderma renal crisis, critical digital ischemia, myocarditis, pericarditis, patient-reported worsening of gastrointestinal symptoms) were not collected as part of the ASCS or CSRG protocols, and these items were not able to be scored (recorded as zero) at any study visit in either cohort. Two further items, troponin and BNP were not available in the CSRG cohort. Therefore, to permit comparison of results between the derivation and validation cohorts, troponin and BNP items remained unscored (recorded as zero) in both cohorts. ROC analysis showed a good discriminatory capacity for the SCTC-AI for mortality, physical function, morbidity, damage, and disease severity (Table 3).

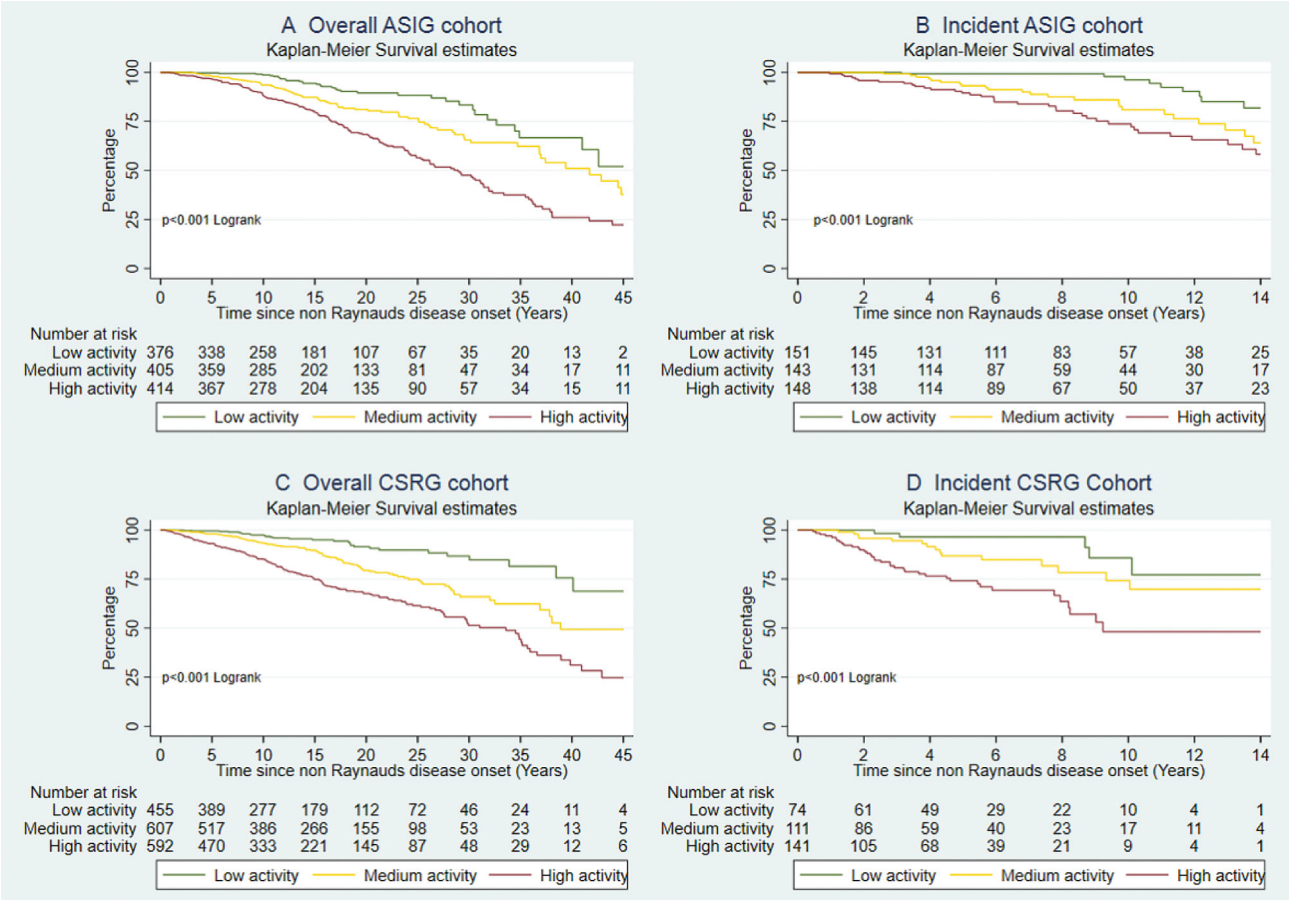

Figure 2. Kaplan-Meier survival analysis for low, medium, and high activity scores. High, medium, and low activity groupings based on a tertile split of baseline SCTC Activity Index Scores. ASCS, Australian Scleroderma Cohort Study; ASIG, Australian Scleroderma Interest Group; CSRG, Canadian Scleroderma Research Group; SCTC, Scleroderma Clinical Trials Consortium.

ASCS participants were grouped into low, medium, and high disease activity groups, based upon the tertile groupings of baseline SCTC-AI scores. Low disease activity was a SCTC-AI score ≤3, medium activity was indicated by a score between 4 and 8, and high disease activity was considered present with a SCTC-AI score ≥9. Kaplan-Meier survival analysis showed that individuals with high disease activity had the poorest prognosis, with a significant, graded risk of death associated with higher SCTC-AI scores (Figure 2). This was similarly observed in the incident subgroup of patients.

Validation of SCTC-AI in the CSRG Study validation cohort. Baseline population characteristics of the CSRG cohort are detailed in Supplementary Index 8. The baseline median SCTC-AI score was 6 (range 1–11) in the overall cohort and 7 (range 4–13) in the incident cohort. CSRG participants had a median of 5 study visits (range 3–8 study visits) over a median follow-up period of 4.10 years (range 1.98–7.24 years). Evaluation of the SCTC-AI in the CSRG cohort (validation cohort) again showed a graded relationship between higher SCTC-AI scores and risk of death (Figure 2). ROC analysis showed a good relationship between increasing SCTC-AI scores and death, function, quality of life, damage, and severity. The performance of the SCTC-AI was comparable to the 2017 EUSTAR AI[10] in both the validation and derivation cohorts (Table 3).

Identifying patients at highest risk of high disease activity. To identify those patients at highest risk of having high disease activity scores, we analyzed the clinical features of the low, medium, and high activity groups in the incident (disease duration <4 years at recruitment) ASCS cohort (Supplementary Index 9). Multivariable regression analysis identified that, in those with incident SSc, higher baseline modified Rodnan Skin Score (mRSS) and patient-reported skin activity, Scl70 antibody positivity, lower FVC, digital ulcers, and synovitis at baseline were more likely to be in the high activity group (Table 4).

DISCUSSION

Using a combination of consensus and data-driven methods we were able to develop a 24-item SSc-specific AI. There have been long-standing efforts to develop multisystem instruments to assess disease activity, damage, and severity in SSc.[4,7,9,10] This is the first such effort to use a combination of consensus and data-driven methods to first define the construct of disease activity in SSc and then develop a multisystem instrument. An important strength of this project is the careful definition of the construct of disease activity. Drawing together an international group of experts in the clinical management and research of SSc as well as patient representatives, we have generated the first consensus definition of disease activity in SSc. Using this definition as the central core construct, we have been able to generate a multisystem index, measuring nine domains of disease while paying careful attention to the face, content, and construct validity of each item. Such careful attention to the core construct of disease activity was essential to develop a measure that could differentiate disease activity from damage. Particular attention was paid to the potential reversibility of individual items included in the SCTC-AI, because reversibility is a key aspect of disease activity and can be considered a distinguishing feature of activity compared to disease-associated damage. It was for this reason certain items with unproven reversibility such as proximal weakness, joint contractures, and digital gangrene were removed from the SCTC-AI.

Involvement of the patient perspective in the assessment of disease response is integral to demonstrating efficacy of novel therapies. We have previously demonstrated that patient-reported symptoms correlate strongly with objective measures of disease progression,[22,23] and patient-reported outcomes are justifiably included in the SCTC-AI. Furthermore, there was consensus among the project steering committee that assessment of change was an important component of the measurement of disease activity. There may be instances in which a previous investigation result is unavailable or has occurred at a time point too distant to make the assessment of current disease activity possible. Therefore, inclusion of patient-reported change was an

Table 4. Risk factors for high disease activity scores*

Variable	Baseline activity 4–8		Baseline activity >9	
	RRR (95% CI)	*P* value	RRR (95% CI)	*P* value
Age at disease onset	1.00 (0.99–1.00)	0.52	1.01 (1.00–1.02)	0.08
Male sex	1.05 (0.65–1.70)	0.85	1.20 (0.73–1.98)	0.47
Scl70 positive	1.49 (0.91–2.45)	0.11	2.17 (1.31–3.57)	<0.01
mRSS	1.02 (1.00–1.04)	0.07	1.03 (1.01–1.05)	0.01
FVC (% predicted)	0.99 (0.98–1.00)	<0.01	0.96 (0.96–0.97)	<0.01
Digital ulcers	1.93 (1.26–2.88)	<0.01	3.28 (2.16–4.99)	<0.01
Synovitis	1.85 (1.27–2.71)	<0.01	4.47 (3.04–6.59)	<0.01
Patient-reported worsening of skin thickening	1.17 (0.76–1.80)	0.48	2.05 (1.32–3.18)	<0.01

* All clinical variables are baseline values recorded at first study visit. CI, confidence interval; FVC, forced vital capacity mRSS, modified Rodnan skin score; RRR, relative risk ratio; Scl70, antitopoisomerase I antibody.

important addition to the SCTC-AI to account for changes in specific aspects of disease status when previous physician assessments are unavailable.

The SCTC-AI WG included patient-research partners to inform the conceptual understanding of disease activity in SSc as well as assist in the generation of relevant SCTC-AI items. Furthermore, we have weighted SCTC-AI items against a patient-reported measure of physical health and quality of life to ensure that the SCTC-AI is capturing aspects of SSc that are meaningful to patients. Improvement in how patients feel, function, or survive is an inherent requirement of regulatory authorities for the approval of novel therapeutics. The extensive consultation with SSc experts, involvement of patient-research partners, and weighting of SCTC-AI items against patient-reported outcomes were intentional steps in the development of this outcome measure to ensure the final instrument's content validity. Content validity of previously proposed activity indices is limited by the absence of patient involvement in the development phase and with the inclusion of cross-sectional assessments of items such as mRSS and DLCO.[10] The majority of SSc experts in the SCTC-AI WG reported that single measures of skin thickness and pulmonary function represented damage rather than disease activity. Thus, statistically significant relationships with endpoints such as morbidity and mortality may be able to be demonstrated with both existing and the new SCTC-AI; however, such results do not assess the content validity of the overall instrument nor necessarily prove the construct validity of a measure of disease activity.

Distinction between disease activity and damage is a well-recognized challenge in SSc, particularly in the absence of validated biomarkers used in other rheumatic diseases that establish a threshold of disease burden, such as a swollen joint count, elevated serum markers of inflammation, or proteinuria. Both activity and damage can be simultaneously present in the same organ in patients with SSc. It was recognized that absolute thresholds of either clinical examination or investigation findings that might denote active disease were frequently lacking; therefore, the WG decided that change in disease parameters would need to be used as an indicator of disease activity. For this reason, many of the SCTC-AI items are "change" items because it was felt that the documentation of change in an item would more reliably distinguish activity from damage.

In the development of the SCTC-DI, the presence of certain disease manifestations for >6 months was required before they could be defined as disease damage.[7] Damage is unidirectional and only increases with time, whereas disease activity can improve or worsen with time. The SCTC-AI has been deliberately designed with the expectation that scores will both increase and decrease over time. Certain items such as a recent diagnosis of pulmonary arterial hypertension (PAH) can only be scored once in the SCTC-AI at the time of first diagnosis. At a subsequent visit, this item score would be removed, and the overall SCTC-AI may improve even though the PAH is still present. However, at that point, PAH should be considered damage and would therefore be scored indefinitely as damage by the SCTC-DI.[7] If PAH substantially progresses, leading to right heart failure, then that progression which likely represents ongoing disease activity, can still be captured by other SCTC-AI items such as ventricular systolic dysfunction and elevated BNP.

The ultimate goal of therapy in SSc is to target treatment to periods of high activity and maintain a low activity or even disease remission state. The SCTC-AI aims to provide a method of quantifying the burden of disease activity that can assist in the enrichment of clinical trial populations to select those patients most likely to respond to the intervention under investigation. The concept of treat-to-target, which has improved the outcomes in other rheumatic diseases, is not yet possible in SSc partly because there are no validated instruments to define appropriate treatment targets. If appropriate sensitivity to change of the SCTC-AI can be demonstrated in future studies, the SCTC-AI could measure response to therapy and potentially provide a treatment target once a validated threshold of low disease activity is defined.

The SCTC-AI has only been provisionally validated in the ASCS and CSRG studies because not all SCTC-AI items are presently collected in these cohorts. The absence of some items in these cohorts is a limitation of our study. In development of the SCTC-AI, experts were asked to consider measurement of activity, irrespective of what variables were available for analysis in the derivation cohort in order to ensure the face and content validity of the overall SCTC-AI. Further evaluation of the construct and criterion validity of the SCTC-AI is required in prospective studies with all SCTC-AI items collected. Furthermore, there is a need to evaluate the performance of the SCTC-AI in an RCT cohort of patients rather than an observational cohort study, in which demonstration of change in activity levels is most important to be able to demonstrate the efficacy of novel therapies.

Prospective evaluation of the feasibility of the SCTC-AI is required. It is not possible to prove or disprove the feasibility of this instrument from retrospective analyses. The SCTC-AI does not prescribe an investigation schedule nor specify the frequency of investigations that should be performed in an RCT. However, time limits have been specified for the clinical presentation of certain items to be attributable to SSc disease activity. This was intended to aid in the delineation between disease activity and damage in SSc and in keeping with the need for the long-term presence of specific items to be rated as "damage" according to the SCTC-DI.[7] The overall length of the SCTC-AI could conceivably limit the feasibility of widespread use of the index; however, its structure is in keeping with other multisystem instruments used to measure disease status in other complex autoimmune diseases such as systemic lupus erythematosus and systemic vasculitis.[24,25] Future studies may identify redundancy of certain SCTC-AI items, enabling the development of a shortened, modified SCTC-AI with similar performance to the overall index.

The SCTC-AI in its current form is limited by including a provisional rather than definitive measure of gastrointestinal disease activity. Reaching consensus regarding appropriate and valid measures of gastrointestinal disease activity was not possible in the Delphi exercises performed in this study. Gastrointestinal items considered by the WG included gastrointestinal symptoms such as sicca symptoms, bloating, vomiting, gastric antral vascular ectasia, small intestinal bacterial overgrowth, malnutrition, pseudo-obstruction, fecal calprotectin, and imaging measures of gastrointestinal dysmotility. Literature review demonstrated no existing validated measures of gastrointestinal disease activity. There was strong consensus from the WG that gastrointestinal disease be included in the SCTC-AI; however, it was repeatedly raised that symptoms and investigation abnormalities generally reflect disease-associated damage rather than disease activity. Consideration was given to inclusion of other multidimensional scores such as the SCTC University of California, Los Angeles Gastrointestinal 2.0 Questionnaire.[26] However, this instrument is an assessment of severity rather than activity, and when considering the ease of use of the final AI, the principal investigators elected to not include other multidimensional instruments within the SCTC-AI. For this reason, the WG made the decision to use self-reported worsening or improvement in GI symptoms after performing a substudy to determine whether this patient-reported outcome is clinically meaningful.[23]

Over time, as understanding of the underlying pathogenic mechanisms of SSc improves, it is likely that revision of certain SCTC-AI items will be required. As serological and radiographic biomarkers of disease are identified, the specificity of items for disease activity should improve. For example, the future widespread adoption of quantitative chest CT assessment could allow for the adoption of a minimal threshold of change such as 2% change in ILD extent.[27] The validation of novel imaging techniques such as positron emission tomography (PET) to assess fibroblast activation[28] may even permit the removal of "change" items to assess disease activity because these novel biomarkers may permit cross-sectional evaluation of disease activity.

In conclusion, we present a novel instrument, the SCTC-AI, to quantify the burden of disease activity in SSc. With careful attention paid to face validity at each step, we have employed a rigorous, consensus-based methodology in combination with data-driven methods to develop an instrument that has face and content validity. Future work is required to confirm the construct and discriminative validity, responsiveness to change, and reliability of the SCTC-AI.

ACKNOWLEDGMENTS

The Australian Scleroderma Cohort Study is supported by Actelion Australia, Scleroderma Australia, Scleroderma Victoria, Arthritis Australia, Musculoskeletal Australia, The Scleroderma Clinical Trials Consortium, St Vincent's Hospital Research Endowment Fund, The Australian Rheumatology Association, philanthropic donations, GlaxoSmithKline, Roche, Pfizer, Bayer, CSL Biotherapies, and Bristol-Myers Squibb. The Canadian Scleroderma Research Group (CSRG) is supported by the Canadian Institutes of Health Research, Scleroderma Canada, the Scleroderma Society of Ontario, the Scleroderma Association of Saskatchewan, Scleroderma Manitoba, the Scleroderma Society of Nova Scotia, the Scleroderma Association of British Columbia, the Cure Scleroderma Foundation, the Canadian Blood and Marrow Transplant Group, and the Lady Davis Institute of Medical Research of the Jewish General Hospital in Montreal, Quebec, Canada. The CSRG has also received educational grants from Actelion pharmaceuticals and Mallinckrodt. No commercial or pharmacological company listed provided direct support to the Activity Index Working Group to support the development of the SCTC Activity Index.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr Nikpour had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Ross, Proudman, Baron, Nikpour.
Acquisition of data. Ross, Proudman, Khanna, Herrick, Stevens, Baron, Nikpour.
Analysis and interpretation of data. Ross, Hansen, Baron, Nikpour.

REFERENCES

1. Cutolo M, Soldano S, Smith V. Pathophysiology of systemic sclerosis: current understanding and new insights. Expert Rev Clin Immunol 2019;15(7):753–764.
2. Steen VD, Medsger TA Jr. The value of the Health Assessment Questionnaire and special patient-generated scales to demonstrate change in systemic sclerosis patients over time. Arthritis Rheum 1997;40(11):1984–1991.
3. Clements P. Measuring disease activity and severity in scleroderma. Curr Opin Rheumatol 1995;7:517–521.
4. Medsger TA Jr, Bombardieri S, Czirjak L, et al. Assessment of disease severity and prognosis. Clin Exp Rheumatol 2003;21(suppl 29): S42–S46.
5. Ross L, Baron M, Nikpour M. The challenges and controversies of measuring disease activity in systemic sclerosis. J Scleroderma Relat Disord 2018;3(2):115–121.
6. Pope JE. The future of treatment in systemic sclerosis: can we design better trials? Lancet Rheumatol 2020;2(3):e185–e194.
7. Ferdowsi N, Huq M, Stevens W, et al; Schleroderma Clinical Trials Consortium Damage Index Working Group, The Australian Scleroderma Interest Group, Canadian Scleroderma Research Group; Scleroderma Clinial Trials Consortium Damage Index Working Group, The Australian Sleroderma Interest Group, Canadian Scleroderma Research Group. Development and validation of the Scleroderma Clinical Trials Consortium Damage Index (SCTC-DI): a novel instrument to quantify organ damage in systemic sclerosis. Ann Rheum Dis 2019;78(6):807–816.
8. 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 2016;68(2):299–311.
9. Valentini G, Della Rossa A, Bombardieri S, et al. European multicentre study to define disease activity criteria for systemic sclerosis. II. Identification of disease activity variables and development of preliminary activity indexes. Ann Rheum Dis 2001;60:592–598.
10. Valentini G, Iudici M, Walker UA, et al. The European Scleroderma Trials and Research group (EUSTAR) task force for the development of

revised activity criteria for systemic sclerosis: derivation and validation of a preliminarily revised EUSTAR activity index. Ann Rheum Dis 2017; 76(1):270–276.

11. Fasano S, Riccardi A, Messiniti V, et al. Revised European Scleroderma Trials and Research Group Activity Index is the best predictor of short-term severity accrual. Ann Rheum Dis 2019;78(12):1681–1685.
12. Doyen M, Houssiau FA, Lauwerys BR, et al. Comparison of the disease activity score and the revised EUSTAR activity index in diffuse cutaneous systemic sclerosis patients. Clin Exp Rheumatol 2020;38-(suppl 125):S53–S58.
13. Ross L, Stevens W, Wilson M, et al. Performance of the 2017 EUSTAR activity index in an scleroderma cohort. Clin Rheumatol 2020;39(12):3701–3705.
14. Valentini G, Bencivelli W, Bombardieri S, et al. European Scleroderma Study Group to define disease activity criteria for systemic sclerosis. III. Assessment of the construct validity of the preliminary activity criteria. Ann Rheum Dis 2003;62:901–903.
15. Groseanu L, Petrescu S, Balanescu A, et al. Do we have good activity indices in systemic sclerosis? Curr Rheumatol Rev 2022;18(1):39–47.
16. Moore OA, Proudman SM, Goh N, et al. Quantifying change in pulmonary function as a prognostic marker in systemic sclerosis-related interstitial lung disease. Clin Exp Rheumatol 2015;33(4)(suppl 91):S111–S116.
17. Hoffmann-Vold AM, Allanore Y, Alves M, et al; EUSTAR collaborators. Progressive interstitial lung disease in patients with systemic sclerosis-associated interstitial lung disease in the EUSTAR database. Ann Rheum Dis 2021;80(2):219–227.
18. Goh NS, Hoyles RK, Denton CP, et al. Short-term pulmonary function trends are predictive of mortality in interstitial lung disease associated with systemic sclerosis. Arthritis Rheumatol 2017;69(8):1670–1678.
19. Volkmann ER, Tashkin DP, Sim M, et al; SLS I and SLS II study groups. Short-term progression of interstitial lung disease in systemic sclerosis predicts long-term survival in two independent clinical trial cohorts. Ann Rheum Dis 2019;78(1):122–130.
20. Distler O, Highland KB, Gahlemann M, et al; SENSCIS Trial Investigators. Nintedanib for systemic sclerosis-associated interstitial lung disease. N Engl J Med 2019;380(26):2518–2528.
21. Kafaja S, Clements PJ, Wilhalme H, et al. Reliability and minimal clinically important differences of forced vital capacity: Results from the Scleroderma Lung Studies (SLS-I and SLS-II). Am J Respir Crit Care Med 2018;197(5):644–652.
22. Ross L, Stevens W, Wilson M, et al. Can patient-reported symptoms be used to measure disease activity in systemic sclerosis? Arthritis Care Res (Hoboken) 2020;72(10):1459–1465.
23. Ross L, Proudman S, Walker J, et al. Evaluation of patient and physician assessments of gastrointestinal disease activity in systemic sclerosis. J Rheumatol 2023;50(4):519–525.
24. Mukhtyar C, Lee R, Brown D, et al. Modification and validation of the Birmingham Vasculitis Activity Score (version 3). Ann Rheum Dis 2009;68(12):1827–1832.
25. Isenberg DA, Rahman A, Allen E, et al. BILAG 2004. Development and initial validation of an updated version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. Rheumatology (Oxford) 2005;44(7):902–906.
26. Khanna D, Hays RD, Maranian P, et al. Reliability and validity of the University of California, Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract Instrument. Arthritis Rheum 2009; 61(9):1257–1263.
27. Volkmann ER, Tashkin DP, Roth MD, Goldin J, Kim GHJ. Early radiographic progression of scleroderma: lung disease predicts long-term mortality. Chest 2022;161(5):1310–1319.
28. Bergmann C, Distler JHW, Treutlein C, et al. ^{68}Ga-FAPI-04 PET-CT for molecular assessment of fibroblast activation and risk evaluation in systemic sclerosis-associated interstitial lung disease: a single-centre, pilot study. Lancet Rheumatol 2021;3(3):e185–e194.

APPENDIX A: ACTIVITY INDEX WORKING GROUP MEMBERS

Members of the Activity Index Working Group are as follows: A/Prof Shervin Assassi, Dr Cosimo Bruni, Dr Soumya Chatterjee, Prof Lorinda Chung, A/Prof Chris Derk, Dr Ivan Foeldvari, Dr Nava Ferdowsi, Dr Andreu Fernandez-Codina, A/Prof Tracy Frech, Dr Jessica Gordon, Dr Faye Hant, A/Prof Marie Hudson, A/Prof Sindhu Johnson, Prof Dinesh Khanna, Dr Andrea Low Hsiu Ling, Dr Ashima Makol, Dr Chris Mecoli, Prof Thomas Medsger, Prof Peter A Merkel, Dr John Pauling, Prof Janet Pope, Dr Tatiana Rodriquez-Reyna, Dr Lesley Saketkoo, A/Prof Nora Sandorfi, Dr Amanda Saracino, Dr James Seibold, Dr Vicki Shanmugam, Prof Virginia Steen, Dr Wendy Stevens, Dr Gemma Strickland, and Prof Gabriele Valentini.