Quantifying Organ Damage
in Systemic Sclerosis:
rationale, methodology and item generation

Dr Tien Siong Tay
ORCID number: 0000-0002-3415-4724

A thesis submitted in fulfilment of the requirements for the degree of
Master of Philosophy at the University of Melbourne

Department of Medicine
University of Melbourne
1 April 2018
# Table of Contents

Abstract 3  
Preface 5  
Acknowledgements 6  
Candidate’s contributions 8  
Publications and Presentations relating to this thesis 9  
Abbreviations 11  

Chapter 1 Literature Review, Hypothesis and the Aims of this Thesis 16  
Chapter 2 Early Accrual of Organ Damage in Systemic Sclerosis using *ad hoc* criteria: rationale for development of a disease Damage Index 21  
Chapter 3 Methodology for Systematic Review and Appraisal of the Psychometric Properties of Measures of Disease Status in Systemic Sclerosis 34  
Chapter 4: Systematic Review of Measures of Health Status in Systemic Sclerosis to Inform Item Generation for Development of a Disease Damage Index 37  
Chapter 5: Implications from the Findings of this Thesis and Future Directions 66
Abstract

Purpose:
The overall purpose of this thesis is to form a rationale for the development of a disease damage index in systemic sclerosis and to inform item generation for inclusion in this index through: (1) appraisal of existing measures of disease status in systemic sclerosis (SSc); and (2) quantifying the accrual of organ damage in patients with early SSc using *ad hoc* criteria.

Methods:
(1) A systematic review of Medline (1966–2015), EMBASE (1974–2015), and Cochrane Library (inception–2015) was undertaken to identify indices of disease status in SSc. The study focused on objective measures and excluded non-English articles. Measures were reviewed for content, whether they measured activity, damage and/or severity and if they were validated according to the OMERACT filter; (2) Patients in the Australian Scleroderma Cohort Study enrolled within two years of onset of the first non-Raynaud’s SSc symptom were included. Organ damage was defined by a group of six experts as substantial and permanent loss of organ function due to SSc.

Results:
(1) Of the 4558 articles retrieved through the search, 58 articles were identified for review. A further 44 articles were found through a search of the bibliography of relevant articles. The systematic review identified the following 10 “composite” (multi-organ) indices: two disease activity indices, six disease severity scales, and two combined response indices. There was no disease ‘damage index’ for SSc. The Valentini disease activity index has face validity, partial validity for content and criterion validity, and sensitivity to change, but not for reliability. The Medsger disease severity scale has face, content and criterion validity but not construct validity, sensitivity to change and reliability. (2) The study identified 278 patients with early SSc. Among these, 38% had diffuse SSc. Damage was more common in the diffuse than in the limited disease subtype in the skin/musculoskeletal (75% vs. 25.2%, *p*<0.001) and lung (31.4% vs. 19.9%, *p* = 0.035) domains at year seven of follow-up. The rates of damage accrual were highest in the skin/musculoskeletal, gastrointestinal and respiratory
systems at year two (29.1%, 18.7%, 14.4%), increasing at year five (41.4%, 30.6%, 21.2%) and declining thereafter to year seven (43.9%, 32.7%, 23.0%). In particular, there was early accrual of damage due to joint contracture (22.3%), gastrointestinal dysmotility (11.5%) and pulmonary fibrosis with forced vital capacity <70% predicted (9.7%) at year two. The highest accrual rate of organ-specific damage from years two to seven was seen in faecal incontinence followed by proximal muscle weakness and pulmonary fibrosis.

Conclusions:

(1) The study identified a number of composite and organ-specific indices in SSc, incorporating mostly objective measures, developed to quantify disease activity, severity, and response in clinical trials. However, none of the indices was developed to exclusively quantify organ damage. Most of the existing indices require further validation according to the OMERACT filter; (2) Substantial accrual of organ damage occurs early in the course of disease, particularly in diffuse SSc. This provides a strong rationale for the development and validation of a SSc damage index.
Preface

Systemic sclerosis (SSc) is an autoimmune disease characterised by inflammation, vasculopathy and excessive collagen production, leading to progressive skin and internal organ fibrosis. It has a wide range of prevalence varying between 7 to 489 per million (1) and a South Australian study has reported a prevalence of 220 cases per million per year (2). Its incidence ranges from 1.2 to 23 cases per million per year (3). Recent meta-analysis of cohort studies of patients with SSc demonstrated a high mortality with a pooled standardised mortality ratio of 3.5 (4). In a recent review, the leading causes of SSc-related deaths were pulmonary arterial hypertension and interstitial lung disease (5). Compared to other rheumatic diseases, such as systemic lupus erythematosus (SLE), rheumatoid arthritis, ankylosing spondylitis, psoriatic arthritis and systemic vasculitis, SSc has the highest loss of life expectancy (6) and a significant reduction in quality of life (7-9).

There are existing measures of disease activity (10-12), disease severity (13-19) and outcome measures for clinical trials in SSc (20, 21). However, there is no index to describe and quantify organ damage in SSc.

The first chapter of this thesis provides a literature review and describes the aims of the thesis. Chapter 2 quantifies the accrual of organ damage in patients with early SSc based on ad hoc criteria using the Australian Scleroderma Cohort Study database. Chapter 3 describes the methodology used to for a systematic review of the psychometric properties of the existing measures of health status in SSc. Chapter 4 discusses the results of the systematic review in relation to item generation for potential inclusion in a survey of experts from the Scleroderma Clinical Trials Consortium (SCTC) for the development of a disease damage index. Chapter 5 presents the implications of the findings from this thesis and the future directions of the project.
Acknowledgements

I thank my supervisors, Dr Mandana Nikpour and Associate Professor David Prior, for the excellent guidance and support throughout my postgraduate studies. I also thank Dr Wendy Stevens for providing excellent clinical supervision and advice with guidance, support and writing of manuscripts. I thank Associate Professor Susanna Proudman for her professional advice with the development of the project and the writing of the manuscripts.

I thank the following individuals: Dr Nava Ferdowski for her assistance in the subsequent part of the project; Mr Molla Huq, Ms Candice Rabusa and Ms Michelle Wilson for their assistance with the statistical analysis and database queries.

I am grateful to members of the SCTC Working Group for their valuable input as well as the ASIG members who have contributed to the Australian Scleroderma Cohort Study. The names and affiliations of the members of the SCTC Working Group for Damage Index and ASIG members who contributed to the project are listed at the end of this section. I would like to thank the following experts: Associate Professor David Prior, Professor David Celermajer, Professor Eli Gabbay, Professor Trevor Williams, Professor Michael Kamm, Dr Peter de Cruz and Associate Professor Robyn Langham for their contribution to the definition of organ damage in the cardiovascular, respiratory and gastrointestinal and nephrology domains, respectively.

Lastly, I am grateful to my wife and children for their forbearance, encouragement and support during my post-graduate studies.

Funding

I thank the Australian Scleroderma Interest Group and Scleroderma Australia for funding my fellowship in 2013.
Scleroderma Clinical Trials Consortium Damage Index (SCTC-DI) Working Group members:
Mandana Nikpour (Co-chair), The University of Melbourne, Australia; Murray Baron (Co-Chair), Lady David Institute for Medical Research and Jewish General Hospital, Canada; Tracy Frech, University of Utah, USA; Marie Hudson, Lady David Institute for Medical Research and Jewish General Hospital, Canada; Susanna Proudman, The University of Adelaide, Australia; Wendy Stevens, St Vincent’s Hospital, Melbourne, Australia; The University of Melbourne, Australia; Tien Tay, The University of Melbourne, Australia; Nava Ferdowsi, St Vincent’s Hospital, Melbourne, Australia; Souyma Chatterjee, Cleveland Clinic, Ohio, USA; Lorinda Chung, Stanford University, USA; Jessica K Gordon, Hospital for Special Surgery, USA; Anna Haemel, University of California San Francisco, USA; Sindhu R Johnson, University of Toronto, Canada; Dinesh Khanna, University of Michigan, USA; Thomas A Medsger, Department of Medicine, University of Pittsburgh, U.S.A; Peter Merkel, University of Pennsylvania, USA; John Pauling, Royal National Hospital for Rheumatic Diseases, Bath, UK; Janet E Pope, University of Ontario, Canada; Tatiana Rodriguez-Reyna, National Institute of Medical Sciences, Mexico; Lesley Saketkoo, Louisiana State University, USA; James R Seibold, Scleroderma Research Consultants, USA; Ankoor Shah, Duke University, USA; Virginia Steen, Georgetown University, USA; Gemma Strickland, Barwon Health, Australia.

Australian Scleroderma Interest Group members:
Catherine Hill, Adelaide, South Australia; Sue Lester, Adelaide, South Australia; Peter Nash, Sunshine Coast, Queensland; Gene Ngian, Melbourne Victoria; Mandana Nikpour, Melbourne, Victoria; Susanna Proudman, Adelaide, South Australia; Maureen Rischmueller, Adelaide, South Australia; Janet Roddy, Perth, Western Australia; Joanne Sahhar, Melbourne, Victoria; Wendy Stevens, Melbourne, Victoria; Gemma Strickland, Geelong, Victoria; Vivek Thakkar, Liverpool, New South Wales, Jenny Walker, Adelaide, South Australia; Jane Zochling, Hobart, Tasmania.
**Candidate’s Contribution**

Dr Mandana Nikpour, Dr Wendy Stevens, Associate Professor David Prior and the candidate jointly contributed to the development of the thesis topic.

The candidate was involved in (1) writing the systematic review and the creation of a repository of articles for item generation, (2) writing the paper on the accrual of organ damage in patients with early SSc using an *ad hoc* criteria, with assistance in data analysis from Mr Molla Huq, and (3) liaising with SCTC Damage Index Working Group members and the non-rheumatology experts to determine the conceptual definition of organ damage, as well as the definitions for the various items for inclusion in the Delphi exercise. During the fellowship year, the candidate was involved in the clinical assessment of SSc patients in the Scleroderma Clinic at St Vincent’s Hospital and systematic data collection for the *ad hoc* damage accrual study.
Publications Relating to this Thesis


Presentations at Australian and International Conferences

Podium presentations


Poster presentations


Oral presentation at Hospital Grand Medical Rounds


2. Tay T. Scleroderma: Overview, Updates, and Systemic Sclerosis Damage Index at the Medical Grand Round at Western Hospital Melbourne (23 May 2014)
Abbreviations

6MWT 6 minute walk test
ASCS Australian Scleroderma Cohort Study
ASIG Australian Scleroderma Interest Group
CRISS Combined Response Index in Systemic Sclerosis
Cs Capillary score
CRP C-reactive protein
CTD Connective Tissue Disease
DAT Dermal atrophy
DLCO Diffusion capacity for carbon monoxide
DP Dyspigmentation
ECG Electrocardiogram
ED Erectile dysfunction
eGFR Estimated glomerular filtration rate
EMG Electromyography
EPOSS Pulmonary Arterial Hypertension related to Systemic Sclerosis
ESKD End-stage kidney disease
ESR Erythrocyte sedimentation rate
EULAR European League Against Rheumatism
EUSTAR European Scleroderma Trials and Research
FEV1 Forced Expiratory Volume in 1 second
FVC Forced vital capacity
FTP Finger-to-palm
GC Giant capillaries
GGO Ground glass opacity
GI Gastrointestinal
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAMIS</td>
<td>Hand Mobility in Scleroderma</td>
</tr>
<tr>
<td>HAQ-DI</td>
<td>Health Assessment Questionnaire Disability Index</td>
</tr>
<tr>
<td>HC</td>
<td>Honeycombing</td>
</tr>
<tr>
<td>HRCT</td>
<td>High resolution computed tomography</td>
</tr>
<tr>
<td>ILD</td>
<td>Interstitial lung disease</td>
</tr>
<tr>
<td>LoSDI</td>
<td>Localized Scleroderma Skin Damage Index</td>
</tr>
<tr>
<td>LTM</td>
<td>Latent linear trajectory model</td>
</tr>
<tr>
<td>LUS</td>
<td>Lung ultrasound</td>
</tr>
<tr>
<td>MCP</td>
<td>Metacarpal joint</td>
</tr>
<tr>
<td>MHE</td>
<td>Micro-haemorrhages</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic resonance imaging</td>
</tr>
<tr>
<td>mRSS</td>
<td>modified Rodnan skin score</td>
</tr>
<tr>
<td>MSK</td>
<td>Musculoskeletal</td>
</tr>
<tr>
<td>MT</td>
<td>Micro-thrombosis</td>
</tr>
<tr>
<td>MUST</td>
<td>Malnutrition Universal Screening Tool</td>
</tr>
<tr>
<td>NEMO</td>
<td>Number of micro-haemorrhages or micro-thrombosis</td>
</tr>
<tr>
<td>NVC</td>
<td>Nailfold videocapillaroscopy</td>
</tr>
<tr>
<td>NYHA</td>
<td>New York Heart Association</td>
</tr>
<tr>
<td>OMERACT</td>
<td>Outcome Measures for Rheumatology</td>
</tr>
<tr>
<td>OSIRIX</td>
<td>Open-source dicom viewer software</td>
</tr>
<tr>
<td>PIP</td>
<td>Proximal inter-phalangeal joint</td>
</tr>
<tr>
<td>PFT</td>
<td>Pulmonary function test</td>
</tr>
<tr>
<td>PAH</td>
<td>Pulmonary arterial hypertension</td>
</tr>
<tr>
<td>PRIMSA-P</td>
<td>Preferred reporting items for systematic review and meta-analysis protocols</td>
</tr>
<tr>
<td>PV</td>
<td>Partially validated</td>
</tr>
<tr>
<td>RHC</td>
<td>Right heart catheterisation</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomised controlled trial</td>
</tr>
<tr>
<td>ROM</td>
<td>Range of movement</td>
</tr>
<tr>
<td>RP</td>
<td>Raynaud’s phenomenon</td>
</tr>
<tr>
<td>SAT</td>
<td>Subcutaneous atrophy</td>
</tr>
<tr>
<td>SCTC</td>
<td>Scleroderma Clinical Trials Consortium</td>
</tr>
<tr>
<td>S-HAQ</td>
<td>Scleroderma Health Assessment Questionnaire</td>
</tr>
<tr>
<td>SF-36</td>
<td>36-item Short Form Health Survey</td>
</tr>
<tr>
<td>SLE</td>
<td>Systemic lupus erythematosus</td>
</tr>
<tr>
<td>SLS</td>
<td>Scleroderma Lung Study</td>
</tr>
<tr>
<td>sPAP</td>
<td>Systolic pulmonary arterial pressure</td>
</tr>
<tr>
<td>SSc</td>
<td>Systemic sclerosis</td>
</tr>
<tr>
<td>STPR</td>
<td>Skin thickness progression rate</td>
</tr>
<tr>
<td>TLC</td>
<td>Total Lung Capacity</td>
</tr>
<tr>
<td>UCLA</td>
<td>University of California Los Angeles</td>
</tr>
<tr>
<td>VAS</td>
<td>Visual Activity Scale</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
</tbody>
</table>


Chapter One

Brief Literature Review, Hypothesis

and the Aims of Thesis
1.1 Brief Literature Review

Systemic sclerosis [(SSc) scleroderma] is an autoimmune disease characterised by inflammation, vasculopathy, and fibrosis, resulting in progressive loss of function in various organ systems. The type and severity of organ involvement in SSc varies among patients and the disease phenotype is influenced by autoantibodies, racial, genetic, and environmental factors. Unlike other connective tissue diseases, such as systemic lupus erythematosus (SLE), which have a relapsing–remitting course, the lack of clearly defined episodes of flare and remission in SSc has made it difficult to describe and quantify disease “activity”. Hence, there is no gold standard instrument for measuring disease activity in SSc (1-3).

The disease activity indices developed by Valentini et al. (4, 5) and Minier et al. (6) are not fully validated according to the OMERACT filter. These indices are only partially valid for content as they do not cover gastrointestinal and renal systems. The Valentini disease activity score is only partially validated for criterion validity as it has been tested against mortality in a case-control study which demonstrated the prognostic value of anti-topoisomerase-1 antibody in patients with SSc compared with other rheumatic conditions (7). The Valentini disease activity index was assessed for construct validity in two studies (6, 8). In this study, Valentini et al. showed a moderate to strong agreement between the ranking by the four SSc experts and the actual disease activity scores of the patients for the whole cohort \( r = 0.428-0.720, p \) ranging from < 0.001 to 0.018) (8). However, there was a low to moderate level of agreement between the ranking and the actual activity scores when patients were stratified into diffuse and limited disease groups (lowest \( r = 0.357, p > 0.05 \)). Subsequently, Minier et al. found that the Valentini disease activity index is associated with several clinical and laboratory markers of disease activity (6). However, reliability has not been assessed. Sensitivity to change was partially validated as it was only assessed in a small study reporting the efficacy of rituximab in SSc (9). Internal and external validation have been undertaken in other studies (4, 10).
Disease “severity” captures the total impact of disease on organ function due to disease activity and damage (11-16). The main limitation of disease severity indices is the lack of distinction between features that are actively progressing (activity), and irreparable organ damage (damage) that may not be reversible with therapy. More recently, outcome measures have been developed for clinical trials in SSc (17, 18). However, these tools are designed to measure response to therapy rather than disease status at any particular point in the course of disease.

To the best of our knowledge, we are not aware of any index that measures organ damage per se in SSc. In addition, there is a paucity of published data on the impact of organ damage on morbidity and mortality. Measures of organ damage have been developed for other rheumatic disease, such as SLICC/ACR DI for SLE and VDI for systemic vasculitis, and they have been shown to prognosticate morbidity and mortality (19, 20).

1.2 Hypotheses

The hypotheses for this thesis are as follows:

1. That the accrual of organ damage in SSc is substantial and occurs early in the disease course providing a rationale for the development of the first disease Damage Index in order to systematically identify and quantify organ damage in this disease; and

2. That a systematic review and appraisal of the psychometric properties of existing measures of disease status in SSc will inform generation of items for potential inclusion in a disease Damage Index.

1.3 Aims of this Thesis

The aims of this thesis are as follow:
1. To determine the frequency and rate of accrual of organ damage in patients with early SSc using ad hoc criteria for organ damage; and

2. To identify and appraise the psychometric properties of existing measures of disease status in SSc through a systematic review.

References


Chapter Two

Early Accrual of Organ Damage in Systemic Sclerosis using *ad hoc* criteria

Incorporating data from:

Introduction

There is a paucity of data on the accrual and the implications of organ damage in SSc. Clinically evident examples of disease damage in SSc include joint contracture, digital amputation, oesophageal stricture, pseudo-obstruction, interstitial lung disease and pulmonary arterial hypertension. Several studies have highlighted the prevalence of joint contracture (1) and digital amputation (2-4) and the increased hospitalisation and adverse impact on health-related quality of life due to these complications (5, 6). The Pittsburgh Scleroderma Cohort (7) and the Scleroderma Lung Study (8) demonstrated that SSc-related interstitial lung disease develops within two years following the first non-Raynaud’s symptom. Pulmonary arterial hypertension which affects approximately 10% of patients with SSc, is another major cause of SSc-related mortality with significant burden on health-related quality of life (9, 10).

Objective of this study

The objective of this study is to assess the extent and rate of accrual of organ damage in patients with early SSc from the Australian Scleroderma Cohort Study, according to an ad hoc definition of significant organ damage.

Results

Among the 1696 patients enrolled in the Australian Scleroderma Cohort Study with SSc at the time of analysis, we identified 278 patients who had SSc of less than two years’ duration. Among these, 38% had diffuse SSc. Damage was more common in the diffuse than in the limited disease subtype in the skin/musculoskeletal (75% vs. 25.2%, p<0.001) and lung (31.4% vs. 19.9%, p = 0.035) domains at year seven. The rates of damage accrual were highest in the skin/musculoskeletal, gastrointestinal and respiratory systems at year two (29.1%, 18.7%, 14.4%), increasing at year five (41.4%, 30.6%, 21.2%) and declining thereafter to year seven (43.9%, 32.7%, 23.0%). In particular, there was early accrual of damage due to joint contracture (22.3%), gastrointestinal dysmotility (11.5%) and
pulmonary fibrosis with forced vital capacity < 70% predicted (9.7%) at year two. The highest accrual rate of organ-specific damage from years two to seven was seen in faecal incontinence followed by proximal muscle weakness and pulmonary fibrosis.

We undertook a sensitivity analysis comparing the accrual of organ damage since the onset of Raynaud’s phenomenon and less than 2 years since the onset of the first non-Raynaud’s manifestation, and found no difference in this cohort of patients with early disease.

**Discussion**

Our study findings have demonstrated an early accrual of organ damage in SSc, particularly in the first four years of the disease which remained stable until year seven. The accrual of organ damage in SSc reflects past and on-going disease activity. There is a lag between the accrual of organ damage and past disease activity, which is the highest in the first four years of diffuse disease and up to seven years for limited disease (11). More than half of our cohort accrued at least one form of organ damage at year two and this increased to almost 70% at year seven. We also showed that organ damage in SSc was the highest in the skin/musculoskeletal system followed by the gastrointestinal and lung domains. Specific SSc-related organ damage included joint contracture, gastrointestinal dysmotility and proximal muscle weakness. Furthermore, the rate of accrual was partly influenced by the availability of effective treatments for the organ damage, such as oesophageal stricture.

The detection of mild organ damage in SSc disease prior to the development of the first non-Raynaud’s phenomenon is limited based on the current laboratory and imaging tests used in routine care. However, it is possible to detect and monitor capillaroscopic changes using nailfold capillaroscopy (NVC), a non-invasive tool, from the onset of Raynaud’s phenomenon. In a prospective cohort of 140 SSc patients, Avouac et al (12) performed sequential nailfold capillaroscopy at baseline followed by three subsequent annual visits. Avouac et al (12) reported
nailfold capillaroscopic morphological changes and microhemorrhages in 90% of the patients over the 3-year study period. In this cohort of patients with a mean disease duration of 9 years, 126 patients (90%) demonstrated a low capillary density at study entry. Follow-up over the next 3 years detected significant NVC changes in 72 (51%) patients, 20% of whom showed progressive capillary loss. The authors found that progressive loss of capillaries predicted overall disease progression (HR = 4.35, 95% CI: 1.87–10.12), occurrence of new digital ulcers (HR = 5.33, 95% CI: 1.69–16.71), lung vascular progression (HR = 18.53, 95% CI: 1.28–78.33), progression of skin fibrosis (HR = 4.22, 95% CI: 1.24–14.36), and worsening of the Medsger severity score (HR = 5.26, 95% CI: 1.78–15.52). More studies are needed to detect and monitor capillaroscopic changes from the onset of Raynaud’s phenomenon (pre-symptomatic phase) and in the initial years following the development of the first non-Raynaud’s symptom (post-symptomatic phase).

As irreversible organ damage is the hallmark of SSc, a composite index of damage in multiple organs would be a useful measure of the consequences of this disease. Our study provides a strong rationale for the development of a SSc damage index (SSc-DI) using robust methodology. Findings from recent studies suggested that nailfold capillary drop out is an independent marker of poor outcome and progression of organ involvement in SSc and should be considered for inclusion in a disease damage index for SSc.

References


Early accrual of organ damage in systemic sclerosis: rationale for development of a disease damage index

Tien Tay1,2, Molla Huq1,2, Nava Ferdowsi1,2, Wendy Stevens3, Joanne Sahar3,4, Gene-Siew Ngian1, Janet Roddy5, Jane Zochling6, Jenny Walker7, Susanna M. Proudman8,9, Mandana Nikpour1,2

1 Department of Rheumatology, St. Vincent’s Hospital, Fitzroy, VIC - Australia
2 Department of Medicine, University of Melbourne, Fitzroy, VIC - Australia
3 Department of Rheumatology, Monash Health, Clayton, VIC - Australia
4 Department of Medicine, Monash University, Clayton, VIC - Australia
5 Royal Perth Hospital, Perth, WA - Australia
6 Menzies Institute for Medical Research, Hobart, Tasmania - Australia
7 Department of Rheumatology, Flinders Medical Centre, Bedford Park, South Australia - Australia
8 Rheumatology Unit, Royal Adelaide Hospital, North Terrace, Adelaide - Australia
9 Discipline of Medicine, University of Adelaide, Adelaide - Australia

ABSTRACT

Introduction: Systemic sclerosis (SSc) is characterized by irreversible organ damage rather than fluctuating disease activity. However, there is no validated measure of damage in SSc. We aimed to quantify the accrual of organ damage in patients with early SSc.

Methods: Patients enrolled in the Australian Scleroderma Cohort Study with less than 2 years of SSc since the onset of the first non-Raynaud’s symptom were included. Organ damage was defined by a group of six experts as substantial and permanent loss of organ function due to SSc.

Results: We identified 278 patients with early SSc. Among these, 38% had diffuse SSc. Damage was more common in the diffuse than in the limited disease subtype in the skin/musculoskeletal (75% vs 25.2%, p<0.001) and lung (31.4% vs. 19.9%, p = 0.035) domains at year seven. The rates of damage accrual were highest in the skin/musculoskeletal, gastrointestinal and respiratory systems at year two (29.1%, 18.7%, 14.4%), increasing at year five (41.4%, 30.6%, 21.2%) and declining thereafter to year seven (43.9%, 32.7%, 23.0%). In particular, there was early accrual of damage due to joint contracture (22.3%), gastrointestinal dysmotility (11.5%) and pulmonary fibrosis with forced vital capacity <70% predicted (9.7%) at year two. The highest accrual rate of organ-specific damage from years two to seven was seen in fecal incontinence followed by proximal muscle weakness and pulmonary fibrosis.

Conclusions: Substantial accrual of organ damage occurs early in the course of disease, particularly in diffuse SSc. This provides the rationale for the development of a SSc damage index.

Keywords: Damage index, Joint contracture, Pulmonary fibrosis, Systemic sclerosis

Introduction

Systemic sclerosis ([SSc]) scleroderma is an autoimmune disease characterized by inflammation, vasculopathy, and fibrosis, resulting in progressive tissue damage and loss of function in various organ systems. Measures of disease status in SSc are potentially useful as end points in observational studies and clinical trials as well as to assess the interval progression of disease. Furthermore, outcome measures are useful for comparison of cohorts, enrichment of studies to focus on specific subsets of SSc patients, and to measure the burden of disease in epidemiological studies.

Unlike other rheumatic diseases, such as systemic lupus erythematosus and rheumatoid arthritis, which have a relapsing-remitting course, the lack of clearly defined episodes of flare and remission in SSc has made it difficult to describe and quantify disease “activity”. There are two disease activity indices in SSc (1, 2) but neither is considered a gold standard measure for disease activity in SSc (3, 4).

Combined response indices by the Composite Response Index for Clinical Trials (CRISS) and the Expert Panel on Outcome Measures in Pulmonary Arterial Hypertension related to SSc (EPOSS) have been developed to assess response to
novel therapies in clinical trials in SSc (5, 6). In addition, several measures of disease severity have been developed to assess the total impact of activity and damage in SSc (7-12). However, none of these measures distinguishes potentially reversible disease activity from irreversible organ damage in SSc. Our recent systematic review identified no instrument to quantify organ damage in SSc (3). This is in part due to a lack of definition of organ damage in SSc.

Clinically evident examples of disease damage in SSc include joint contracture, digital amputation, esophageal stricture, pseudo-obstruction, interstitial lung disease and pulmonary arterial hypertension. Several studies have highlighted the prevalence of joint contracture (13) and digital amputation (14-16) and the increased hospitalization and adverse impact on health-related quality of life of these complications (17, 18). The Pittsburgh Scleroderma Cohort (19) and the Scleroderma Lung Study (20) demonstrated that SSc-related interstitial lung disease (ILD), a major cause of SSc-related mortality (21), develops within two years following the first non-Raynaud's symptom. Pulmonary arterial hypertension (PAH), which affects approximately 10% of patients with SSc, is another major cause of SSc-related mortality with significant burden on health-related quality of life (21, 22).

The objective of this study is to assess the extent and rate of accrual of organ damage in patients with early SSc from the Australian Scleroderma Cohort Study, according to an ad hoc definition of significant organ damage.

Methods

Study population

The Australian Scleroderma Cohort Study (ASCS) is a large multicenter observational study of risk and prognostic factors for cardiopulmonary outcomes in SSc. Patients fulfilling the 2013 classification criteria for SSc were included in this study (23). All patients undergo an annual clinical assessment, transthoracic echocardiography and pulmonary function test. We also collected demographic details, comprehensive data related to SSc, co-morbidities and medication history. The ASCS is approved by the human research ethics committees of the 13 participating Australian centers. All patients provide written informed consent at recruitment.

Inclusion and exclusion criteria

We included patients aged 18 years and older with early SSc, defined as having disease duration of less than two years since the onset of the first non-Raynaud's symptom, with at least one annual follow-up visit, from ASCS between December 2007 and June 2016.

Definitions

A preliminary definition of significant organ damage was agreed by a multidisciplinary panel of six experts from rheumatology, cardiology, respiratory and gastroenterology. Significant organ damage for each organ system was defined as substantial and permanent loss of organ function due to the SSc disease process, or its treatment. Organ-specific damage was defined as below.

Skin/musculoskeletal domain

Damage in the skin and musculoskeletal domain was defined as one or more of joint contracture, digital gangrene or amputation or myositis with muscle weakness and atrophy. Joint contracture, digital gangrene, digital amputation and muscle atrophy were based on clinical assessment. Myositis was defined based on characteristic clinical features with either elevated serum creatine kinase (based on individual laboratory reference range) or a diagnostic muscle biopsy. Proximal muscle weakness was assessed clinically using the 5-point system where 0 = no contraction, 1 = muscle flicker but no movement, 2 = movement possible but not against gravity, 3 = movement possible against gravity, 4 = movement possible against resistance, and 5 = normal strength.

Gastrointestinal domain

Damage in the gastrointestinal domain was defined as one or more of esophageal stricture, gastrointestinal dysmotility, pseudo-obstruction or fecal incontinence. Esophageal stricture was defined based on endoscopy or barium swallow. Gastrointestinal dysmotility was defined as either suspected (based on clinical assessment with a response to antibiotics) or definite (confirmed on barium studies or scintigraphy). Pseudo-obstruction was defined by symptoms such as vomiting or constipation, with dilatation of the small and/or large bowel on imaging. Fecal incontinence was defined as incontinence at least three times per week.

Cardiovascular domain

Damage in the cardiovascular domain was defined as one or more of myocardial disease, conduction defect and left ventricular dysfunction not due to other causes.

Myocardial disease was defined as either suspected (based on conduction defect or left ventricular dysfunction or diastolic dysfunction) or definite cases (based on endomyocardial biopsy). Conduction defect was confirmed on electrocardiogram (ECG) or Holter monitoring. Left ventricular dysfunction was defined as having left ventricular ejection fraction (LVEF) of 50% or less on transthoracic echocardiography. We excluded other causes of left ventricular dysfunction, such as ischemia or infarction, based on clinical assessment and echocardiography. Screening for coronary heart disease was based on history, baseline ECG and annual echocardiography. More invasive procedures, such as coronary angiography, were undertaken on a case-by-case basis on clinical grounds.

Pulmonary domain

Damage in the pulmonary domain was defined as the presence of clinically significant ILD. This was defined as typical changes of lung fibrosis, on high resolution computed tomography (HRCT) of the chest with extent of ILD greater than 20%, or with forced vital capacity less than 70% predicted.

© 2017 Wichtig International
PAH was defined by a mean pulmonary arterial pressure of 25 mmHg or higher at rest and a pulmonary wedge pressure of 15 mmHg or less on right heart catheterization (24). Right ventricular enlargement was determined based on expert opinion from echocardiography.

Renal domain

Damage in the renal domain was defined as presence of scleroderma renal crisis. This was defined as having two of three criteria: new onset systemic hypertension, microangiopathic hemolytic anemia and elevated creatinine with an estimated glomerular rate (eGFR) <30 ml/min/1.73 m².

We intended to determine damage accrual over time. The definitions of organ damage were selected as they were the definitions used in the Australian Scleroderma Cohort database at the time of the analysis.

Statistical analysis

An individual item of organ damage was indicated as present or absent, and was accrued for each domain per annual follow-up following enrollment in the study. Once an item of organ damage was accrued during the follow-up duration of the study, the same item could not be used to accrue further damage. Data are presented as means (standard deviation [SD]) for all continuous variables and numbers (percentages) for all categorical variables. Analysis was undertaken to compare organ damage in diffuse and limited disease, and to assess the number (%) of patients with one or more organ damage at year two since disease onset where most of the damage was seen. Statistical analysis was undertaken using STATA 14 (Statacorp, College Station, TX, USA).

Results

Among the 1696 patients enrolled in the ASCS with SSc at the time of analysis, 278 patients had SSc of less than two years’ duration. The majority of these were female (79.8%) and Caucasian (89.5%) (Tab. I). The mean (SD) age at onset of SSc was 52.6 (13.6) years and the mean (SD) disease duration at recruitment was 1.0 (SD) year. A high proportion of these patients with early SSc had diffuse SSc (37.8%). Autoantibody profile included scl-70 (topoisomerase-1) antibody (20.9%), centromere antibody (31.6%) and anti-RNA III polymerase antibody (18.2%). More than half of the patients were on or had been treated with some form of immunosuppressive therapy, including corticosteroids.

During the study period, there were 100 (36%) patients without follow-up reviews for more than 24 months. Over a 10-year period, 33 (11.9%) patients died, 12 (4.3%) withdrew from the study and 10 (3.6%) could not be contacted for follow-up.

The primary SSc-related causes of deaths were: PAH (11), pulmonary fibrosis (6), cardiac (5), gastrointestinal (5), renal (2) and SSc-related ischemia complicated by infection (2). Non-SSc-related causes of death were malignancy (5), sepsis (3), cardiogenic shock (1), coronary artery disease (1) and hemorrhagic stroke (1).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>N (%)</th>
<th>or mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>56</td>
<td>(20.2%)</td>
</tr>
<tr>
<td>Female</td>
<td>221</td>
<td>(79.8%)</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aboriginal-Islander</td>
<td>3</td>
<td>(1.1%)</td>
</tr>
<tr>
<td>Caucasian</td>
<td>239</td>
<td>(89.5%)</td>
</tr>
<tr>
<td>Asian</td>
<td>18</td>
<td>(6.7%)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1</td>
<td>(0.4%)</td>
</tr>
<tr>
<td>Other</td>
<td>6</td>
<td>(2.3%)</td>
</tr>
<tr>
<td><strong>Age at onset of systemic sclerosis (y)</strong></td>
<td>52.6</td>
<td>(±13.6)</td>
</tr>
<tr>
<td><strong>Disease duration at recruitment</strong></td>
<td>1.0</td>
<td>(±0.5)</td>
</tr>
<tr>
<td><strong>Length of follow-up (y)</strong></td>
<td>4.7</td>
<td>(±2.6)</td>
</tr>
<tr>
<td><strong>Disease subtype</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limited</td>
<td>171</td>
<td>(61.5%)</td>
</tr>
<tr>
<td>Diffuse</td>
<td>105</td>
<td>(37.8%)</td>
</tr>
<tr>
<td><strong>Antibodies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scl-70 (Topoisomerase-I)</td>
<td>54</td>
<td>(20.9%)</td>
</tr>
<tr>
<td>Centromere</td>
<td>83</td>
<td>(31.6%)</td>
</tr>
<tr>
<td>Anti-RNA polymerase III</td>
<td>33</td>
<td>(18.2%)</td>
</tr>
<tr>
<td>Corticosteroids</td>
<td>131</td>
<td>(47.1%)</td>
</tr>
<tr>
<td>Immunosuppressives</td>
<td>148</td>
<td>(53.2%)</td>
</tr>
<tr>
<td>Cyclophosphamide</td>
<td>34</td>
<td>(12.2%)</td>
</tr>
<tr>
<td>Methotrexate</td>
<td>97</td>
<td>(34.9%)</td>
</tr>
<tr>
<td>Mycophenolate</td>
<td>45</td>
<td>(16.2%)</td>
</tr>
<tr>
<td>Azathioprine</td>
<td>17</td>
<td>(6.1%)</td>
</tr>
<tr>
<td>Rituximab</td>
<td>2</td>
<td>(0.7%)</td>
</tr>
<tr>
<td>Tocilizumab</td>
<td>3</td>
<td>(1.1%)</td>
</tr>
</tbody>
</table>

*Patients who ever had immunosuppressive treatment(s).

At the end of follow-up, there was more organ damage in the diffuse group compared to the limited group for skin/musculoskeletal (75.2% vs. 25.2%, p<0.001) and respiratory domains (31.4% vs. 18.7%, p = 0.016) (Tab. II). Accrual of organ damage in other domains were similar in the two groups. At year two since disease onset, more than half of this early cohort (53.6%) had accrued at least one form of organ damage and this increased to almost 70% at year seven.

Skin/musculoskeletal damage was high (29.1%) at two years, increasing to 39.6% at four years, then increasing more slowly to 43.9% at seven years (Tab. III). Within this domain, joint contracture was common at two years (22.3%), gradually rising to 29.9% at four years, and plateauing at seven years (32.7%). Proximal muscle weakness was less frequent at two years (8.6%) but increased by more than 50% at four years to 13.3% before it plateaued at seven years (18.0%). In contrast, digital gangrene, digital amputation, and myositis with muscle weakness and muscle atrophy were infrequent at two years (2.5%, 1.8% and 1.1%, respectively) and increased marginally over the next five years (3.2%, 4.0%, 2.2%, respectively).

© 2017 Wichtig International
### TABLE II - Frequency of organ damage according to the disease subtype of systemic sclerosis (SSc) in patients with incident SSc of less than two years’ duration

<table>
<thead>
<tr>
<th>Disease damage</th>
<th>Limited (total = 171)</th>
<th>Diffuse (total = 105)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td></td>
</tr>
<tr>
<td>Skin/musculoskeletal</td>
<td>43 (25.2%)</td>
<td>79 (75.2%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>52 (30.4%)</td>
<td>41 (39.1%)</td>
<td>0.140</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>12 (7.0%)</td>
<td>14 (13.3%)</td>
<td>0.081</td>
</tr>
<tr>
<td>Respiratory</td>
<td>32 (18.7%)</td>
<td>33 (31.4%)</td>
<td>0.016</td>
</tr>
<tr>
<td>Renal</td>
<td>2 (1.2%)</td>
<td>5 (4.8%)</td>
<td>0.065</td>
</tr>
<tr>
<td>Overall</td>
<td>101 (59.1%)</td>
<td>92 (87.6%)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

### TABLE III - Accrual of organ damage over seven years since the onset of the first non-Raynaud’s symptom in systemic sclerosis (n = 278)

<table>
<thead>
<tr>
<th>Disease damage</th>
<th>2 years n (%)</th>
<th>3 years n (%)</th>
<th>4 years n (%)</th>
<th>5 years n (%)</th>
<th>6 years n (%)</th>
<th>7 years n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin/musculoskeletal</td>
<td>81 (29.1%)</td>
<td>104 (37.4%)</td>
<td>110 (39.6%)</td>
<td>115 (41.4%)</td>
<td>120 (43.2%)</td>
<td>122 (43.9%)</td>
</tr>
<tr>
<td>Digital gangrene</td>
<td>7 (2.5%)</td>
<td>7 (2.5%)</td>
<td>8 (2.9%)</td>
<td>8 (2.9%)</td>
<td>8 (2.9%)</td>
<td>9 (3.2%)</td>
</tr>
<tr>
<td>Digital amputation</td>
<td>5 (1.8%)</td>
<td>8 (2.9%)</td>
<td>9 (3.2%)</td>
<td>9 (3.2%)</td>
<td>10 (3.6%)</td>
<td>11 (4.0%)</td>
</tr>
<tr>
<td>Joint contractures</td>
<td>62 (22.3%)</td>
<td>76 (27.3%)</td>
<td>83 (29.9%)</td>
<td>86 (30.9%)</td>
<td>90 (32.4%)</td>
<td>91 (32.7%)</td>
</tr>
<tr>
<td>Myositis with muscle weakness and muscle atrophy</td>
<td>3 (1.1%)</td>
<td>3 (1.1%)</td>
<td>5 (1.8%)</td>
<td>5 (1.8%)</td>
<td>6 (2.2%)</td>
<td>6 (2.2%)</td>
</tr>
<tr>
<td>Proximal muscle weakness</td>
<td>24 (8.6%)</td>
<td>36 (12.9%)</td>
<td>37 (13.3%)</td>
<td>43 (15.5%)</td>
<td>47 (16.9%)</td>
<td>50 (18.0%)</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>52 (18.7%)</td>
<td>66 (23.7%)</td>
<td>79 (28.4%)</td>
<td>85 (30.6%)</td>
<td>91 (32.7%)</td>
<td>91 (32.7%)</td>
</tr>
<tr>
<td>Esophageal stricture</td>
<td>15 (5.4%)</td>
<td>17 (6.1%)</td>
<td>20 (7.2%)</td>
<td>21 (7.6%)</td>
<td>22 (7.9%)</td>
<td>23 (8.3%)</td>
</tr>
<tr>
<td>Gastrointestinal dysmotility</td>
<td>32 (11.5%)</td>
<td>35 (12.6%)</td>
<td>36 (12.9%)</td>
<td>37 (13.3%)</td>
<td>37 (13.3%)</td>
<td>37 (13.3%)</td>
</tr>
<tr>
<td>Pseudo-obstruction</td>
<td>4 (1.4%)</td>
<td>8 (2.9%)</td>
<td>9 (3.2%)</td>
<td>9 (3.2%)</td>
<td>9 (3.2%)</td>
<td>9 (3.2%)</td>
</tr>
<tr>
<td>Fecal incontinence</td>
<td>12 (4.3%)</td>
<td>21 (7.6%)</td>
<td>31 (11.2%)</td>
<td>38 (13.7%)</td>
<td>43 (15.5%)</td>
<td>43 (15.5%)</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>17 (6.1%)</td>
<td>18 (6.5%)</td>
<td>20 (7.2%)</td>
<td>22 (7.9%)</td>
<td>24 (8.6%)</td>
<td>26 (9.4%)</td>
</tr>
<tr>
<td>Myocardial disease and either conduction defect or left ventricular dysfunction</td>
<td>6 (2.2%)</td>
<td>7 (2.5%)</td>
<td>8 (2.9%)</td>
<td>8 (2.9%)</td>
<td>8 (2.9%)</td>
<td>8 (2.9%)</td>
</tr>
<tr>
<td>Conduction abnormality</td>
<td>10 (3.6%)</td>
<td>10 (3.6%)</td>
<td>11 (4.0%)</td>
<td>12 (4.3%)</td>
<td>13 (4.7%)</td>
<td>15 (5.4%)</td>
</tr>
<tr>
<td>Left ventricular ejection fraction &lt;50%</td>
<td>4 (1.4%)</td>
<td>4 (1.4%)</td>
<td>5 (1.8%)</td>
<td>6 (2.2%)</td>
<td>8 (2.9%)</td>
<td>8 (2.9%)</td>
</tr>
<tr>
<td>Respiratory</td>
<td>40 (14.4%)</td>
<td>54 (19.4%)</td>
<td>56 (20.1%)</td>
<td>59 (21.2%)</td>
<td>62 (22.3%)</td>
<td>64 (23.0%)</td>
</tr>
<tr>
<td>Pulmonary fibrosis with FVC &lt;70% predicted</td>
<td>27 (9.7%)</td>
<td>39 (14.0%)</td>
<td>42 (15.1%)</td>
<td>43 (15.5%)</td>
<td>45 (16.2%)</td>
<td>48 (17.3%)</td>
</tr>
<tr>
<td>Moderate to severe extent pulmonary fibrosis on HRCT</td>
<td>21 (7.6%)</td>
<td>27 (9.7%)</td>
<td>27 (9.7%)</td>
<td>27 (9.7%)</td>
<td>27 (9.7%)</td>
<td>27 (9.7%)</td>
</tr>
<tr>
<td>Pulmonary arterial hypertension and right ventricular dysfunction or dilation</td>
<td>14 (5.0%)</td>
<td>18 (6.5%)</td>
<td>18 (6.5%)</td>
<td>22 (7.9%)</td>
<td>24 (8.6%)</td>
<td>24 (8.6%)</td>
</tr>
<tr>
<td>Renal: Renal crisis and eGFR &lt;60 mL/sec</td>
<td>3 (1.1%)</td>
<td>3 (1.1%)</td>
<td>3 (1.1%)</td>
<td>7 (2.5%)</td>
<td>7 (2.5%)</td>
<td>7 (2.5%)</td>
</tr>
<tr>
<td>Overall (damage in one or more organ systems)</td>
<td>149 (53.6%)</td>
<td>168 (60.4%)</td>
<td>176 (63.3%)</td>
<td>186 (66.9%)</td>
<td>191 (68.7%)</td>
<td>193 (69.4%)</td>
</tr>
</tbody>
</table>

PFC = forced vital capacity; HRCT = high resolution computed tomography of the chest; eGFR = estimated glomerular filtration rate.

Gastrointestinal damage was common at two years (18.7%) and increased to 28.4% at four years before plateauing at seven years (32.7%) (Tab. III). The most frequent form of gastrointestinal damage was gastrointestinal dysmotility which occurred early in the disease: 11.5% at two years, increasing to 12.9% at four years then remaining stable at 13.3% from year five onwards. Fecal incontinence was infrequent at two years (4.3%) but almost tripled at four years (11.2%) and quadrupled at seven years (15.5%) before it plateaued thereafter. Esophageal stricture was also infrequent in early SSc, 5.4% at two
years, but increased marginally to 7.2% at four years before it plateaued at seven years (8.3%). Pseudo-obstruction was uncommon in early SSc (1.4% at two years), increasing slowly to 3.2% at four years and remaining unchanged thereafter.

Cardiovascular disease was uncommon at two years (6.1%) and increased gradually to 7.2% at four years before the prevalence stabilized at seven years (9.4%) (Tab. III). Conduction defect and left ventricular (LV) dysfunction (LV ejection fraction <50%) were infrequent at two years and increased marginally over the next five years (3.6% compared with 5.4%, and 1.4% compared with 2.9%, respectively, at seven years).

ILD with forced vital capacity (FVC) <70% predicted was relatively frequent at two years (9.7%), increased to 15.1% at four years and plateaued at 17.3% at seven years (Tab. III). PAH with right ventricular dysfunction or dilatation was uncommon at two years (5.0%) and increased gradually to 8.6% at seven years (Tab. III). In this cohort, 56 patients had right heart catheterization.

Scleroderma renal crisis was infrequent in the first four years of the disease (1.1%) but showed a small spike at five years (2.5%) and remained stable thereafter (Tab. III).

The rate of damage accrual varied depending on the organ involved. The highest rates of organ damage accrual in the five years subsequent to the first two years of disease were seen in fecal incontinence (four-fold), followed by proximal muscle weakness (two-fold) and pulmonary fibrosis (70% increase) (Tab. III).

The rate of accrual of organ damage in most domains, especially in skin/musculoskeletal, gastrointestinal and respiratory systems, was most steep in the two years following year two (Fig. 1). Accrual of organ damage continued, albeit at a slower pace, between years five and seven.

Discussion

In this cohort of early SSc, we found accrual of damage in multiple organs, particularly in the first four years from the onset of the first non-Raynaud symptom, and the rate of accrual declined thereafter to year seven. Furthermore, damage in the skin/musculoskeletal and respiratory domains was more frequent in patients in the diffuse disease subset. At year two, more than half of our cohort (53.6%) accrued at least one form of organ damage and this increased to almost 70% at year seven. The highest prevalence of organ damage in early SSc was seen in the skin/musculoskeletal (29.1%), gastrointestinal (18.7%) and lung (ILD) (9.7%) domains at two years after disease onset. By year seven, we found a high accrual of organ-specific damage due to joint contracture (32.7%), proximal muscle weakness (18.0%), ILD (17.3%), fecal incontinence (15.5%) and gastrointestinal dysmotility (13.3%). The highest accrual rate of organ-specific damage from years two to seven was seen in fecal incontinence followed by proximal muscle weakness and ILD.

Almost one-third of our cohort of early SSc developed joint contracture over seven years from disease onset. Most of the damage due to joint contracture occurred in the first two years (22.3%). This is similar to the overall prevalence of joint contracture at 31% reported in the longitudinal EUSTAR study (n = 7286) (13). Furthermore, these authors reported a prevalence of 33% in the first five years of the disease which increased to 43% in patients with SSc for more than 10 years. Joint contracture was also more prevalent in diffuse compared to limited disease. In addition, joint contracture was associated with disease activity in the musculoskeletal domain more broadly (synovitis, tendon friction rub), elevated acute phase reactants and muscle weakness.

A significant proportion of our cohort (18%) developed proximal muscle weakness in the first seven years of disease. In the prospective EUSTAR study (n = 7655) enriched for diffuse SSc (37.1%) and with long-standing disease, Meier et al reported a prevalence of 25% for muscle weakness, which was higher in the diffuse compared to the limited group (33.5% vs. 18.9%, p<0.001) (25). The authors also reported a relatively high prevalence (12.2%) of muscle atrophy in the same cohort, with a higher prevalence in the diffuse group compared to the limited group (18.0% vs. 8.1%, p<0.001). Although not directly comparable, the EUSTAR prevalence is much higher than the accrual of 2.2% for myositis with
muscle weakness and muscle atrophy at seven years in our study. This is likely due to the fact that this report of the EUSTAR study included patients with long-standing disease in contrast to our cohort of patients with early disease.

In a retrospective study from the Pittsburgh Scleroderma cohort, Follansbee et al identified 183 of 1095 (17%) SSc patients with skeletal myopathy (26). The average onset of skeletal myopathy was 2.6 years from the onset of SSc symptoms. However, Ranque et al reported that the median time from diagnosis of SSc to the onset of myopathy was five years in 35 French patients, 75% of whom had diffuse disease (27).

In another retrospective study of 42 SSc patients with proximal muscle weakness who had a muscle biopsy from the John Hopkins Scleroderma cohort (n = 2830), Paik et al demonstrated that proximal muscle weakness developed four years after the first non-Raynaud's phenomenon symptom in SSc (28).

Our cohort had a low accrual of digital gangrene and digital amputation, 3.2% and 4.7%, respectively, at seven years. Although not directly comparable, other cross-sectional studies have reported the prevalence of digital gangrene and digital amputation in SSc. Allanore et al reported that the prevalence of patients who had ever had digital gangrene and current digital gangrene was 18% and 5.6%, respectively, in the prospective Digital Ulcer Outcome Registry (n = 4494), which is enriched for patients with long-standing disease (14). In this study, the mean time for the development of digital ulcer from the first non-Raynaud’s phenomenon symptom was seven years for patients with current or those who ever had digital gangrene. In contrast, Nihityanova et al reported a lower prevalence of SSc patients with digital gangrene (1.4%) and auto-amputation or required surgical amputation (0.9%) in a retrospective cross-sectional study from the Royal Free Hospital cohort (n = 1168) (15). In another retrospective cross-sectional study on 188 SSc patients, Caramaschi et al reported a prevalence of 4.8% for partial or total surgical digital amputation (16).

Our study identified a relatively high incidence and accrual of gastrointestinal damage in both early diffuse and limited SSc. Gastrointestinal dysmotility was found in almost one-fifth of our cohort within two years following disease onset and that increased to involve one-third of the cohort over the next five years. The accrual of fecal incontinence continued over the first seven years following disease onset to affect 15% of the cohort. In contrast, the accrual of esophageal stricture slowed after the first two years since disease onset, likely due to the availability of effective treatments for esophageal stricture as well as for gastro-esophageal reflux. To the best of our knowledge, we are not aware of any study assessing the accrual of gastrointestinal damage in SSc. Previous studies have reported high gastrointestinal involvement in SSc: esophageal problems (80%) of whom 30%-40% are symptomatic (29), stomach (60%) (29), small intestine (40%) (29), colon, anorectum (50%-70%) with 20% of patients reporting fecal incontinence (30).

Our cohort demonstrated slow accrual of cardiovascular damage in SSc: 6.1% at two years, increasing to 9.4% at seven years. Conduction defects and left ventricular dysfunction were infrequent at two years and increased marginally to 5.4% at seven years. In the prospective EUSTAR study (n = 7655), Meier et al reported a prevalence of conduction blocks of 11.0% and of diastolic dysfunction of 17.4% in a cohort with long-standing disease (25). In a study of 183 SSc patients who had ECG and Holter monitoring, Kostis et al reported a high prevalence of supraventricular and ventricular arrhythmias in SSc: supraventricular ectopic beats (61%), supraventricular tachycardia (21%), ventricular ectopic activity (67%) and ventricular tachycardia (7%) (31). However, conduction defect was less common. This cohort has a high proportion of patients with diffuse (57%) and early SSc with a disease duration of three years or less. More than half (57%) were symptomatic. Similar findings of high prevalence of cardiac arrhythmias in SSc were reported in two other studies with a high proportion of diffuse disease (31, 32). Although not directly comparable to our study, these cross-sectional studies reported a higher prevalence of cardiac arrhythmias and conduction defects in SSc. This is most likely due to the fact that we used studies performed ECG and Holter monitoring in all patients, half of whom were symptomatic. In contrast, our study only undertook routine ECG screening in all patients, most of whom were asymptomatic. Further investigations, including Holter monitoring, were done only if clinically indicated. This would lead to an under-detection of asymptomatic cardiac arrhythmias and conduction defects in our cohort and hence an under-estimation of the accrual rate of cardiovascular damage in SSc.

Our data also showed an early accrual of significant ILD from year two (9.7%) to year seven (17.4%) following disease onset. In the Pittsburgh University Scleroderma cohort (n = 890), Steen et al reported an annualized decline in FVC of 32% in the first two years, 12% in year two to year four, and 3% in year four to six following disease onset suggesting that lung disease occurs early in the disease (19). Similar findings were reported in the Scleroderma Lung Study (n = 158) (20).

Advanced PAH with subsequent right-ventricular dysfunction or dilatation is mainly seen later in the course of the disease and hence this would explain the low accrual in our cohort with early SSc.

The prevalence and rate of accrual of scleroderma renal crisis was also low in our cohort of early SSc. In the recent cross-sectional EUSTAR study (n = 7655), Meier et al reported a low prevalence of 2.1% (25). This is lower than the findings reported in studies prior to 1970s, which showed a prevalence of up to 12% in diffuse disease and 2% in limited disease (33). This may reflect close surveillance for renal crisis in early disease and the use of angiotensin-converting enzyme inhibitors to treat renal crisis.

Our study has several limitations. This is a relatively small sample size with limited duration of follow-up and some loss of follow-up four years following disease onset. Furthermore, we used ad hoc definitions of organ damage in SSc defined by a small group of experts as there is no existing definition of organ damage in SSc. The use of ad hoc definitions reflects more severe organ damage and is not sensitive to detect subclinical organ damage in SSc. We deliberately excluded treatment-related damage from the analyses. The Scleroderma Clinical Trials Consortium Damage Working Group of experts have agreed that treatment-related damage should not be included in a SSc disease damage index and that all damage items should be due to the disease itself (34). In addition, we did not assign weighting for the various items of organ dam-
age based on the differential impact on mortality and morbidity in SSC. We have also not included data using disease onset following the initial Raynaud's phenomenon, which was used to define disease duration of SSC in some studies. We also do not have complete data, such as EMG, MRI and muscle biopsy, to correlate with proximal muscle weakness and muscle atrophy as these tests are not routinely undertaken in the absence of clinical indication. Similarly, we do not have complete data for specific cardiac investigations, such as Holter monitoring, electrophysiological study, and right heart catheterization, to correlate with ECG and echocardiography findings as these tests are not routinely undertaken in our observational study without specific clinical indications.

Conclusions

We have demonstrated an early accrual of organ damage in SSC, particularly in the first four years of the disease which remained stable until year seven. The accrual of organ damage in SSC reflects past and on-going disease activity. There is a lag between the accrual of organ damage and past disease activity, which is the highest in the first four years of diffuse disease and up to seven years for limited disease (35). More than half of our cohort accrued at least one form of organ damage at year two and this increased to almost 70% at year seven. We also showed that organ damage in SSC was the highest in the skin/musculoskeletal system followed by the gastrointestinal and lung domains. Specific SSC-related organ damage included joint contracture, gastrointestinal dysmotility and proximal muscle weakness. Furthermore, the rate of accrual was partly influenced by the availability of effective treatments for the organ damage, such as esophageal stricture. Our study highlights the need for a comprehensive definition of SSC-related organ damage. As irreversible organ damage is the hallmark of SSC, a composite index of damage in multiple organs would be a useful measure of the consequences of this disease.

Disclosures

Financial support: Australian Scleroderma Interest Group (ASIG)/Scleroderma Australia fellowship (TT, NF), NHMRC Clinical Research Fellowship APP1071735 (MN) and University of Melbourne David Bickart Fellowship (MN). The Australian Scleroderma Interest Group receives research support from Actelion, Bayer, CSL, Pfizer, GSK, St. Vincent's Research Endowment Fund, Scleroderma Australia and Arthritis Australia. Conflict of interest: None of the authors has financial interest related to this study to disclose.

References


Chapter Three

Methodology for Systematic Review and Appraisal of the Psychometric Properties of Measures of Disease Status in Systemic Sclerosis
1.1 OMERACT filter

The OMERACT filter is a set of validation criteria based on truth (face, content, construct, and criterion), discrimination capacity (reliability and sensitivity to change), and feasibility, to improve the development and selection of outcome measures in Rheumatology (Table 1). The OMERACT filter has been used to develop outcome measures in rheumatoid arthritis, ankylosing spondylitis, SSc, pulmonary arterial hypertension, systemic lupus erythematosus, osteoporosis, and osteoarthritis.

1.2 Methodology for the Systematic Review

We undertook a search using the Medline (EBSCO) database (1966–December 2015), the EMBASE (EMBASE.com) database (1974–December 2015), and the Cochrane Library (inception–December 2015) with the assistance of two senior librarians at The Carl de Gruchy Library at St Vincent’s Hospital Melbourne. We searched using the following keywords: (scleroderma, systemic OR systemic sclerosis OR scleroderma), (disease activity), (severity scale OR disease severity OR severity of illness index), (outcome assessment (healthcare) OR outcome measure OR response measure OR measure of response), (disease status OR disease assessment indices), and (disease damage OR damage index). Inclusion and exclusion criteria. We limited our search to include (1) English language and (2) “human-only” studies between the inception date for the respective databases and December 2015. The search included review articles relating to outcome measures, systematic reviews, randomised controlled studies, cohort, and case–control studies. We excluded publications relating to subjective measures of health status, which are predominantly based on patient-reported instruments. The bibliographies of relevant articles identified using our search strategy were also reviewed for additional studies. Two researchers (myself and NF) independently screened the search results using the inclusion and exclusion criteria. Data abstraction was performed using a customised form (Supplementary Table 1). If there was any uncertainty regarding the inclusion of a publication, this was adjudicated by two senior researchers (MN and WS). We accepted any title if there was uncertainty about how to classify, thus being inclusive rather than exclusive.
then reviewed the content of each measure, focussing on whether it measured mostly activity, damage and/or severity and whether it had been validated according to the OMERACT filter summarised below as per our customised data abstraction form (1). The manuscript was prepared in accordance with the 2015 revised PRISMA-P guidelines (2).

Table 1: Definitions of OMERACT terms

Validity

1. **Face validity** is defined as the measure of how representative a measure is ‘at face value’.

2. **Content validity** refers to the extent to which a measure represents the range of possibilities associated with the disease state.

3. **Criterion validity** is the extent to which a measure compares to a gold standard.

4. **Construct validity** defines the degree to which a measure captures what it purports to be measuring.

Discrimination capacity

1. **Reliability** is defined as the overall consistency of a measure.

2. **Sensitivity to change** reflects if the measure discriminates between states at different times.

Feasibility refers to the time, cost and the ease of measuring the outcome.

References


Chapter Four

Systematic review and assessment of the psychometric properties of Measures of Disease Status in Systemic Sclerosis

This chapter is comprised of the following original research article:

Supplementary Table 1. Data abstraction form for Systemic Sclerosis Damage Index

Tracking number □□□□

Reviewer Name: ____________________________

Title: ______________________________________

Key words: __________________________________

Authors: ____________________________________

Year published: ____________________________

1. **Source:**
   - Medline
   - EMBASE
   - Cochrane Central
   - Bibliography (specify: _____________________________________________)

2. **Study type**
   - Published article
   - Abstract / presentation
   - Technical report
   - Book / book chapter
   - Unpublished dissertation / thesis
   - Others (specify: _____________________________________________)

3. **Study Design**
   - Systematic review
   - Meta-analysis
   - Randomized trial
   - Cohort study  □ Prospective  □ Retrospective
   - Case-control study □ Prospective  □ Retrospective
   - Review
   - Others (specify: _____________________________________________)

4. **Outcome measure index**
   - Composite (multi-organ) index (specify: _____________________________________________)
   - Organ-specific index (specify: _____________________________________________)
   - Disease activity
   - Disease severity
   - Disease damage
   - Response criteria
   - Weighted
   - Global assessment score
   - Others (specify: _____________________________________________)

5. **Organ system(s)**
   - Musculoskeletal
     - Arthritis
     - Myopathy / myositis
     - Others (specify: _____________________________________________)
   - Skin
   - Vascular (specify)
     - Raynaud’s phenomenon
- Nailfold capillaroscopy
- Digital ulcers / digital amputation
- Erectile dysfunction

- Cardiac
  - Pulmonary arterial hypertension
  - Cardiac arrhythmias
  - Conduction defects
  - Myocardial fibrosis
  - Others (specify: ____________________________)

- Pulmonary
  - Interstitial lung disease
  - Others (specify: ____________________________)

- Gastrointestinal
  - Esophageal
  - Gastric
  - Intestinal
  - Malnutrition
  - Anal incontinence
  - Others (specify: ____________________________)

- Renal
  - Renal crisis
  - Others (specify: ____________________________)

  - Others (specify: ____________________________)

6. Objective measure(s)
- Laboratory (specify: ____________________________)
- Imaging (specify: ____________________________)
- Others (specify: ____________________________)

7. Patient reported outcome(s)
- Organ-specific (specify: ____________________________)
- General health (specify: ____________________________)

8. OMERACT filter
- Validity
  - Face validity
  - Content validity
  - Criterion validity
  - Construct validity

- Discrimination capacity
  - Reliability
  - Sensitivity to change

- Feasibility
  - Validation
  - Internal validation
  - External validation
Supplementary Table 2: Completed checklist based on the 2015 revised PRISMA-P guidelines – adapted from Moher et al. [1]

I. Administrative information

1. Title
   a. Identification: identified the report as a systematic review
   b. Update: no previous systematic review of health status in systemic sclerosis

2. Registration: not registered

3. Authors
   a. The contact details of all authors and the corresponding author have been provided
      (page 1)
   b. Contributions: TT and NF undertook the systematic review and prepared the
      manuscript. MB, WS, MH and SP designed the study and contributed to the
      preparation of the manuscript. MN designed the study, supervised and assisted TT
      and NF in undertaking the systematic review, and prepared the manuscript. All
      authors have read and approved the final manuscript. Dr Mandana Nikpour is the
      guarantor of the systematic review.

4. Amendments: not applicable

5. Support
   a. Sources and sponsor: sources of financial support have been listed on page 1
   b. Role of sponsor / funder: sponsors and funders played no role in this study

II. Introduction

1. Rationale:
   a. There is no gold standard instrument for measuring disease activity in systemic
      sclerosis
   b. Disease severity indices cannot distinguish between activity (potentially reversible)
      and damage (irreversible) in systemic sclerosis
c. There is lack of a disease damage index to quantify organ damage in systemic sclerosis

2. **Objective**: To identify and appraise measures of disease status in systemic sclerosis, especially in relation to validation according to the OMERACT filter.

III. **Methods**

1. **Eligibility criteria**: stated in the Methods section (page 2)

2. **Information sources**: stated in the Methods section (page 2)

3. **Search strategy**: the search strategy for Medline will be provided on request

4. **Study records**:
   
   a. **Data management**: The articles identified from the literature search were imported into EndNote X7. Data abstraction forms have been stored.

   b. **Selection process**: The selection process has been stated in the Methods section (page 2)

   c. **Data collection process**: articles identified using data abstraction form (Supplementary Table 1)

5. **Data items**: data were grouped according to methodology, composite or organ-specific indices, subjective or objective measures and the OMERACT filter (Tables 2 to 6 and Supplementary Tables 3 to 8)

6. **Outcomes and prioritization**: outcomes were based on the OMERACT filter

7. **Risk of bias in individual studies**: potential sources of bias and the limitations of each of the studies was addressed

8. **Data synthesis**: data were analyzed and described qualitatively

Reference

Supplementary Table 3: Summary of measures of musculoskeletal manifestations in systemic sclerosis (SSc) according to the OMERACT filter (adapted from Clements et al. [19])

<table>
<thead>
<tr>
<th>Musculoskeletal manifestations</th>
<th>Disease status</th>
<th>Face</th>
<th>Content</th>
<th>Criterion</th>
<th>Construct</th>
<th>Validity</th>
<th>Discrimination capacity</th>
<th>Feasibility</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joint/synovial/tendon disease activity&lt;sup&gt;[19]&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arthritis</td>
<td>A</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>PV</td>
<td>ND</td>
<td>PV</td>
<td>Y</td>
<td>?</td>
</tr>
<tr>
<td>Swollen joint count</td>
<td>A</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>PV</td>
<td>ND</td>
<td>PV</td>
<td>Y</td>
<td>?</td>
</tr>
<tr>
<td>Tender joint count</td>
<td>A</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>PV</td>
<td>ND</td>
<td>PV</td>
<td>Y</td>
<td>?</td>
</tr>
<tr>
<td>Tendon friction rub</td>
<td>A</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>PV</td>
<td>ND</td>
<td>PV</td>
<td>Y</td>
<td>?</td>
</tr>
<tr>
<td>Imaging to detect SSc-related arthritis&lt;sup&gt;[19]&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiography</td>
<td>D</td>
<td>Y</td>
<td>PV</td>
<td>Y</td>
<td>PV</td>
<td>?</td>
<td>?Y</td>
<td>Y</td>
<td>?</td>
</tr>
<tr>
<td>Ultrasound</td>
<td>A, D</td>
<td>Y</td>
<td>PV</td>
<td>Y</td>
<td>PV</td>
<td>?</td>
<td>ND</td>
<td>Y</td>
<td>?</td>
</tr>
<tr>
<td>MRI</td>
<td>A, D</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>PV</td>
<td>?</td>
<td>ND</td>
<td>PV</td>
<td>?</td>
</tr>
<tr>
<td>Radiographic evidence of joint disease&lt;sup&gt;[19]&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bony erosions</td>
<td>D</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>Y</td>
<td>?</td>
</tr>
<tr>
<td>Calcinosis</td>
<td>D</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>?Y</td>
<td>Y</td>
<td>?</td>
</tr>
<tr>
<td>Acro-osteolysis</td>
<td>D</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>Y</td>
<td>?</td>
</tr>
<tr>
<td>Joint contracture&lt;sup&gt;[19, 31]&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small joints</td>
<td>D, S</td>
<td>Y</td>
<td>PV</td>
<td>ND</td>
<td>Y</td>
<td>ND</td>
<td>Y</td>
<td>Y</td>
<td>?</td>
</tr>
<tr>
<td>Large joints&lt;sup&gt;[32, 33]&lt;/sup&gt;</td>
<td>D, S</td>
<td>Y</td>
<td>PV</td>
<td>ND</td>
<td>ND</td>
<td>PV</td>
<td>Y</td>
<td>Y</td>
<td>?</td>
</tr>
<tr>
<td>Finger range of movement&lt;sup&gt;[19]&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finger-to-palm distance&lt;sup&gt;[3]&lt;/sup&gt;</td>
<td>S</td>
<td>PV</td>
<td>Y</td>
<td>Y</td>
<td>PV</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Delta finger-to-palm distance&lt;sup&gt;[30]&lt;/sup&gt;</td>
<td>S</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>ND</td>
<td>Y</td>
<td>ND</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Hand function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fist closure&lt;sup&gt;[33]&lt;/sup&gt;</td>
<td>A, S, D</td>
<td>Y</td>
<td>PV</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>PV</td>
<td>Y</td>
<td>PV</td>
</tr>
<tr>
<td>Modified hand anatomic index&lt;sup&gt;[27]&lt;/sup&gt;</td>
<td>A, S, D</td>
<td>Y</td>
<td>Y</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Hand Mobility in Scleroderma&lt;sup&gt;[36]&lt;/sup&gt;</td>
<td>A, S, D</td>
<td>Y</td>
<td>Y</td>
<td>PV</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>PV</td>
</tr>
<tr>
<td>Muscle weakness</td>
<td>A, S, D</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>---------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Myopathy[8]</td>
<td>A, S</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Myositis[8]</td>
<td>A, S</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Muscle atrophy</td>
<td>S, D</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>Y</td>
<td>?</td>
</tr>
<tr>
<td>Serum creatine kinase[10]</td>
<td>A</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Serum aldolase[10]</td>
<td>A</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>?</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

A: disease activity; D: disease damage; S: disease severity; PV: partially validated; ND: no data.

Reference

Supplementary Table 4: Summary of measures of vascular manifestations in SSc according to the OMERACT filter

<table>
<thead>
<tr>
<th>Vascular manifestations</th>
<th>Disease status</th>
<th>Face</th>
<th>Content</th>
<th>Validity Criterion</th>
<th>Construct</th>
<th>Discrimination capacity Reliability</th>
<th>Sensitivity to change</th>
<th>Feasibility Internal</th>
<th>Validation Internal</th>
<th>Validation External</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raynaud phenomenon</td>
<td>A, S</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Digital pitting scar</td>
<td>S, D</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>Y</td>
<td>ND</td>
<td>?</td>
<td>Y</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Digital ulcer</td>
<td>A, S, D</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Digital gangrene/necrosis</td>
<td>D</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Digital amputation</td>
<td>D</td>
<td>Y</td>
<td>Y</td>
<td>?</td>
<td>ND</td>
<td>NA</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Dilated nailfold capillaries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nailfold capillaroscopy (microscopy count)</td>
<td>A, S, D</td>
<td>Y</td>
<td>Y</td>
<td>?</td>
<td>Y</td>
<td>Y</td>
<td>PV</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Nailfold videocapillaroscopy (NVC)[42]</td>
<td>A, S, D</td>
<td>Y</td>
<td>Y</td>
<td>?</td>
<td>Y</td>
<td>Y</td>
<td>PV</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Nailfold capillaroscopy simple scoring system</td>
<td>A, S, D</td>
<td>Y</td>
<td>Y</td>
<td>PV</td>
<td>Y</td>
<td>Y</td>
<td>PV</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Modified NEMO[45]</td>
<td>A, S, ?D</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
</tr>
<tr>
<td>Erectile dysfunction</td>
<td>A, S, D</td>
<td>Y</td>
<td>Y</td>
<td>?Y</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>Y</td>
<td>PV</td>
<td>PV</td>
</tr>
</tbody>
</table>

A: disease activity; S: disease severity; D: disease damage; ND: no data; PV: partially validated; NA: not applicable; NEMO: number of micro-hemorrhages or micro-thrombosis
### Supplementary Table 5: Summary of measures of skin involvement in SSc according to the OMERACT filter

<table>
<thead>
<tr>
<th>Measures of skin involvement</th>
<th>Face</th>
<th>Content</th>
<th>Validity</th>
<th>Criterion</th>
<th>Construct</th>
<th>Reliability</th>
<th>Discrimination capacity</th>
<th>Feasibility</th>
<th>Internal</th>
<th>External</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rodnan's skin score[^48]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>modified Rodnan's skin score[^49]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Geirsson et al.[^5]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>PV</td>
<td>Y</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>Furst et al.[^8]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Silman et al.[^50]</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>Y</td>
<td>PV</td>
<td>ND</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>Kahaleh et al.[^51]</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>Y</td>
<td>PV</td>
<td>ND</td>
<td>PV</td>
<td>Y</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>Photopheresis trial[^52]</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>Y</td>
<td>ND</td>
<td>ND</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>% body surface area (BSA)[^52]</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>Y</td>
<td>ND</td>
<td>ND</td>
<td>Y^^</td>
<td>Y</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>Gruber et al.[^53]</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>Y</td>
<td>ND</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>Localized scleroderma skin damage index[^54]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>Other forms of objective skin assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durometer[^40]</td>
<td>Y</td>
<td>Y</td>
<td>?</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Ultrasound[^60]</td>
<td>?</td>
<td>Y</td>
<td>PV</td>
<td>PV</td>
<td>PV</td>
<td>PV</td>
<td>PV</td>
<td>PV</td>
<td>PV</td>
<td>PV</td>
</tr>
<tr>
<td>Skin biopsy[^61]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>PV</td>
<td>N</td>
<td>PV</td>
<td>?</td>
<td></td>
</tr>
</tbody>
</table>

ND: no data; PV: partially validated; ^^ degree of skin involvement recorded on a body diagram then %BSA determined using a digital planimeter.
<table>
<thead>
<tr>
<th>Measures of health status for cardiac and pulmonary manifestations in SSC</th>
<th>Disease status</th>
<th>Validity</th>
<th>Discrimination capacity</th>
<th>Feasibility</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiac-specific outcome measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiac score</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Cardiac index</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Pulmonary arterial hypertension items</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 minute walk test</td>
<td>A, S, D</td>
<td>Y</td>
<td>NV</td>
<td>ND</td>
<td>PV</td>
</tr>
<tr>
<td>Oxygen desaturation</td>
<td>A, S, D</td>
<td>Y</td>
<td>PV</td>
<td>Y</td>
<td>PV</td>
</tr>
<tr>
<td>Echocardiography</td>
<td>S, D</td>
<td>Y</td>
<td>NV</td>
<td>PV</td>
<td>PV</td>
</tr>
<tr>
<td>Right heart catheterization</td>
<td>S, D</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>VO₂ max</td>
<td>S, D</td>
<td>Y</td>
<td>PV</td>
<td>ND</td>
<td>PV</td>
</tr>
<tr>
<td>NYHA/WHO functional class</td>
<td>A, S, D</td>
<td>Y</td>
<td>PV</td>
<td>Y</td>
<td>PV</td>
</tr>
<tr>
<td>Respiratory items</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interstitial lung disease specific outcome measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Warrick et al.</td>
<td>S, D</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Goldin et al.</td>
<td>S, D</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
</tr>
<tr>
<td>Goh et al.</td>
<td>S, D</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
</tr>
<tr>
<td>Outcome measures to assess interstitial lung disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High resolution computed tomography</td>
<td>S, D</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>PV</td>
</tr>
<tr>
<td>Forced vital capacity</td>
<td>S, D</td>
<td>Y</td>
<td>NV</td>
<td>Y</td>
<td>PV</td>
</tr>
<tr>
<td>Diffusion capacity for carbon monoxide</td>
<td>S, D</td>
<td>Y</td>
<td>NV</td>
<td>Y</td>
<td>PV</td>
</tr>
</tbody>
</table>

A: disease activity; S: disease severity; D: disease damage; ND: no data, PV: partially validated; NV: not validated.

Reference

Supplementary Table 7: Summary of the scoring methods to assess interstitial lung disease (ILD) related to systemic sclerosis (SSc) (adapted from Assayag et al.\textsuperscript{[74]})

(1) Comparative scoring method

**Wells et al.\textsuperscript{[75]}**

<table>
<thead>
<tr>
<th>Abnormality</th>
<th>Grade</th>
<th>Max score (per lobe)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parenchymal opacification alone</td>
<td>1</td>
<td>0 to 5</td>
</tr>
<tr>
<td>Parenchymal opacification more than reticular pattern</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Parenchymal opacification equals to reticular pattern</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Reticular pattern more than parenchymal opacification</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Reticular pattern alone</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

(2) Simplified scoring method

**Morelli et al.\textsuperscript{[76]}**

<table>
<thead>
<tr>
<th>Abnormality</th>
<th>Score</th>
<th>Scoring done in 3 zones: upper, middle, lower</th>
<th>Max score (summation of scores of all zones): 0 to 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ground-glass, reticular markings, bronchiectasis</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Honeycombing</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Tochimoto et al.\textsuperscript{[77]}**

<table>
<thead>
<tr>
<th>Abnormality</th>
<th>Zones</th>
<th>Score</th>
<th>Max score (summed of all zones)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intensity of interstitial shadow</td>
<td>Upper</td>
<td>Y</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Middle</td>
<td>Y</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Lower</td>
<td>Y</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Lung base</td>
<td>Y</td>
<td>4</td>
</tr>
<tr>
<td>Ground glass opacity</td>
<td></td>
<td>Add 1</td>
<td></td>
</tr>
</tbody>
</table>

(3) Semi-quantitative scoring method

**Warrick et al.\textsuperscript{[78]}**

<p>| Abnormality | Grade | Extent score |
|-------------|-------|--------------|-------------|
|             |       |              |             |</p>
<table>
<thead>
<tr>
<th>Abnormality</th>
<th>% disease extent</th>
<th>Score</th>
<th>Anatomical regions scored</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ground-glass opacity alone</td>
<td>0</td>
<td>0</td>
<td>Each lobe is scored independently</td>
</tr>
<tr>
<td>Mixed ground glass and reticular disease</td>
<td>1 to 25</td>
<td>1</td>
<td>Max score (summed of all scores): 0 to 20</td>
</tr>
<tr>
<td>Reticular fibrosis alone</td>
<td>26 to 50</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Honeycombing</td>
<td>51 to 75</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>76 to 100</td>
<td>4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Goldin et al.** [79]

**Abnormality**

- Pure ground-glass opacity
- Fibrosis (including thickened reticular markings, bronchiectasis, bronchiolectasis)
- Honeycombing

**Pandey et al.** [81]

Scoring for ground glass opacity, reticular abnormalities and honeycombing in all five lobes. Each score is weighted for the relative lobar volume.

**Goh & Wells** [68]

**Global extent of disease**

Visual estimation to the nearest 5%

**Extent of individual HRCT abnormality**

- Reticular disease
- Ground-glass attenuation

**Coarseness of reticulation**

(for the most severe disease in each section)

**Grade**

**Abnormality**

- 0 Ground-glass attenuation alone

**Joi et al.** [80]

**Grading for each abnormality**

<table>
<thead>
<tr>
<th>Bronchopulmonary segments involved</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 to 3</td>
<td>1</td>
</tr>
<tr>
<td>4 to 9</td>
<td>2</td>
</tr>
<tr>
<td>&gt; 9</td>
<td>3</td>
</tr>
</tbody>
</table>

Max extent score (summed of all zones): 0 to 15
1. Fine intralobular fibrosis
2. Microcystic honeycombing
3. Macrocystic honeycombing

Assessed at 5 levels: origin of great vessels, main carina, pulmonary venous confluence, halfway between levels 3 and 5, immediately above right hemidiaphragm. Total coarseness score is the summed score for all five levels (max score: 0 to 15).

(4) Computer-aided quantitative scoring method
Carolti[82] (open-source dicom viewer software, OSIRIX)
Nguyen-Kim[83] (histogram-based quantification of fibrosis in SSc-ILD)
Supplementary Table 8: Summary of measures of health status for renal manifestations in SSc according to the OMERACT filter (adapted from Furst et al.\cite{40,91})

<table>
<thead>
<tr>
<th>Renal manifestations</th>
<th>Disease status</th>
<th>Validity</th>
<th>Discrimination capacity</th>
<th>Feasibility</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Face</td>
<td>Content</td>
<td>Sensitivity to change</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Criterion</td>
<td>Construct</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Reliability</td>
<td>Feasibility</td>
<td>Internal</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Validation</td>
<td>External</td>
</tr>
<tr>
<td>Scleroderma renal crisis</td>
<td>A, S, D</td>
<td>PV</td>
<td>PV</td>
<td>Y</td>
<td>ND</td>
</tr>
<tr>
<td>Blood pressure</td>
<td>A, S</td>
<td>PV</td>
<td>Y</td>
<td>PV</td>
<td>ND</td>
</tr>
<tr>
<td>Proteinuria</td>
<td>A, S, D</td>
<td>Y</td>
<td>ND</td>
<td>PV</td>
<td>ND</td>
</tr>
<tr>
<td>24 hour urine collection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urinalysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urine microscopy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urine dipstick</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persistent renal impairment</td>
<td>A, S, D</td>
<td>Y</td>
<td>Y</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Direct measurement GFR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum creatinine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eGFR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A: disease activity; S: disease severity; D: disease damage; ND: no data; PV: partially validated; GFR: glomerular filtration rate; eGFR: estimated glomerular filtration rate.
Measures of disease status in systemic sclerosis: A systematic review☆

Tien Tay, MBBS, MMed, FRACP, Nava Ferdowsi, MBBS, FRACP, Murray Baron, MD, FRCPC, Wendy Stevens, MBBS, FRACP, Marie Hudson, MD, MPH, FRCPC, Susanna M. Proudman, MBBS, FRACP, Mandana Nikpour, MBBS, FRACP, FRCPA, PhD, Damage Index Working Group of the Scleroderma Clinical Trials Consortium

A R T I C L E   I N F O

Keywords:
Systemic sclerosis
Disease activity
Severity scale
Outcome measure
Response measure
Damage index

A B S T R A C T

Objectives: To identify and appraise measures of disease status in systemic sclerosis (SSc).

Methods: A systematic review of Medline (1966–2015), EMBASE (1974–2015), and Cochrane Library (inception–2015) was undertaken to identify indices of disease status in SSc. We focused on objective measures and excluded non-English articles. Measures were reviewed for content, whether they measured activity, damage and/or severity and whether they were validated according to the OMERACT filter.

Results: Of the 4558 articles retrieved through the search, we identified 58 articles for review. We found a further 44 articles through a search of the bibliography of relevant articles. We identified the following 10 “composite” (multi-organ) indices: two disease activity indices, six disease severity scales, and two combined response indices. There was no disease damage index for SSc.

Conclusions: We identified a number of composite and organ-specific indices in SSc, incorporating mostly objective measures, developed to quantify disease activity, severity, and response in clinical trials. However, none of the indices was developed to exclusively quantify organ damage. Most of the existing indices require further validation according to the OMERACT filter. There is a need to develop and validate a disease damage index in SSc.

© 2016 Elsevier Inc. All rights reserved.

Introduction

Systemic sclerosis [(SSc) scleroderma] is an autoimmune disease characterized by inflammation, vasculopathy, and fibrosis, resulting in progressive loss of function in various organ systems. The type and severity of organ involvement in SSc varies among patients and the disease phenotype is influenced by autoantibodies, racial, genetic, and environmental factors. Unlike other connective tissue diseases, such as systemic lupus erythematosus (SLE), which have a relapsing–remitting course, the lack of clearly defined episodes of flare and remission in SSc has made it difficult to describe and quantify disease “activity.” Hence, there is no gold standard instrument for measuring disease activity in SSc [1,2]. Disease “severity” captures the total impact of disease on organ function due to disease activity and damage [3–9]. The main limitation of disease severity indices is the lack of distinction between features that are actively progressing (activity), and irreparable organ damage (damage) that may not be reversible with therapy. More recently, outcome measures have been developed for clinical trials in SSc [10,11]. However, these tools are designed to measure response to therapy rather than disease status at any particular point in the course of disease. To the best of our knowledge, we are not aware of any index that measures organ damage per se in SSc. In addition, there is a paucity of published data on the impact of organ damage on morbidity and mortality. Furthermore, existing indices of health status in SSc have not all been fully validated according to the OMERACT filter [12] (Table 1). The OMERACT filter is a set of validation criteria...
based on truth (face, content, construct, and criterion), discrimination capacity (reliability and sensitivity to change), and feasibility to improve the selection of outcome measures in Rheumatology. The OMERACT filter has been used to identify outcome measures in rheumatoid arthritis, ankylosing spondylitis, SSc, pulmonary arterial hypertension (PAH), SLE, osteoporosis, and osteoarthritis.

This systematic review was undertaken to identify and appraise the existing measures of disease status, in SSc, especially in relation to validation according to the OMERACT filter. The purpose of this systematic review was twofold. Firstly, this review was intended as a comprehensive reference article in relation to measures of disease status in SSc. Secondly, this review was intended to identify items for potential inclusion in a disease damage index in SSc that is being developed by the SSc-Damage Index Working Group of the Scleroderma Clinical Trials Consortium.

Methods

We undertook a search using the Medline (EBSCO) database (1966–December 2015), the EMBASE (EMBASE.com) database (1974–December 2015), and the Cochrane Library (inception–December 2015) with the assistance of two senior librarians at The Carl de Gruchy Library at St Vincent’s Hospital Melbourne. We searched using the following keywords: (scleroderma, systemic OR systemic sclerosis OR scleroderma), (disease activity), (severity scale OR disease severity OR severity of illness index), (outcome assessment (health care) OR outcome measure OR response measure OR measure of response), (disease status OR disease assessment indices), and (disease damage OR damage index).

Inclusion and exclusion criteria

We limited our search to include (1) English language and (2) “human-only” studies between the inception date for the respective databases and December 2015. The search included review articles relating to outcome measures, systematic reviews, randomized controlled studies, cohort, and case-control studies. We excluded publications relating to subjective measures of health status, which are predominantly based on patient-reported instruments.

The bibliographies of relevant articles identified using our search strategy were also reviewed for additional studies.

Two researchers (T.T. and N.F.) independently screened the search results using the inclusion and exclusion criteria. Data abstraction was done using a form (Supplementary Table 1). If there was any uncertainty regarding the inclusion of a publication, this was adjudicated by two senior researchers (M.N. and W.S.). We accepted any title if there was uncertainty about how to classify, thus being inclusive rather than exclusive. We reviewed the content of each measure, whether it measured mostly activity, damage and/or severity, and whether it had been validated according to the OMERACT filter.

The article was prepared in accordance with the 2015 revised PRISMA-P guidelines [13]. The completed PRISMA-P checklist has been included as Supplementary Table 2. Outcome measures were assessed using the OMERACT filter [12].

Results

Search results

A total of 4558 articles were identified using the above search (Fig. 1). There were 1327 duplicate articles. A further 3034 articles were excluded as they were not relevant to the systematic review. Of the remaining 198 articles, 58 articles were included in the systematic review. In addition, 44 articles were either identified from the bibliography of relevant articles or from hand search. We assessed the validity of each of these measures by applying the Outcome Measures for Rheumatology (OMERACT) filter [12].

Composite measures of health status in SSc

Composite measures of disease activity in SSc

We identified two indices that measure disease activity in SSc (Table 2) [1,2]. The Valentini disease activity score is a weighted composite multi-organ index consisting of 10 objective and subjective items, which are grouped into 5 organ systems [1]. These items were identified from a prospective study involving 290 patients from 19 European centers and were correlated with the physician global assessment score assigned by three experts. Minier et al. [2] validated the Valentini index in 131 Hungarian patients with SSc (Table 2) and found that lung disease is not adequately represented in the index so they proposed a new 12-point scale to include more variables in the heart and lung domains [change in diffusion capacity (DLCO) and forced vital capacity (FVC)/DLCO ratio].

OMERACT filter

The disease activity indices developed by Valentini et al. [1] and Minier et al. [2] are only partially valid for content as they do not cover gastrointestinal and renal systems (Table 2). The Valentini disease activity score is only partially validated for criterion validity as it has been tested against mortality in a case–control study which demonstrated the prognostic value of antitopo-isomerase-1 antibody in patients with SSc compared with other rheumatic conditions [14]. The Valentini disease activity index was assessed for construct validity in two studies [2,15]. In the first study, Valentini et al. [15] showed a moderate to strong agreement between the ranking by the four SSc experts and the actual disease activity scores of the patients for the whole cohort \( r = 0.428–0.720, p < 0.001–0.018 \). However, there was a low to moderate level of agreement between the ranking and the actual activity scores when patients were stratified into diffuse and limited disease groups (lowest \( r = 0.357, p > 0.05 \)). Subsequently, Minier et al. [2] found that the Valentini disease activity index is associated with several clinical and laboratory markers of disease activity. However, reliability has not been assessed. Sensitivity to change was partially validated as it was only assessed in a small
study reporting the efficacy of rituximab in SSc [16]. Internal and external validation have been undertaken in other studies [1,17].

**Composite measures of disease severity in SSc**

We identified six indices that measure disease severity (i.e., both activity and damage) in SSc [3–9]. Hughes et al. [7] and Casas et al. [9] developed the initial disease severity indices based on a limited number of organ systems, with a cumulative score that represents overall disease severity. In a case–control study to assess T-cell lymphocyte count in 27 diffuse SSc patients and 45 normal controls, Hughes et al. [7] developed a disease severity index based on the major organ involvement with a weighted score for each organ involvement. In a small case series of 12 diffuse SSc patients who were treated with intravenous 5-FU, Casas et al. [9] used a disease severity index based on similar major organ involvement, including Raynaud’s phenomenon with severity graded on a spectrum from mild to severe.

Furst et al. [18] developed an organ-specific summary index to quantify and summarize organ involvement in a case–control study comparing clinical and serological characteristics of 17 diffuse and 17 limited SSc patients. However, the authors acknowledged that the study was underpowered to estimate the weighting for most items apart from skeletal and upper gastrointestinal systems. The composite index was further refined in a randomized controlled trial comparing chlorambucil and placebo involving 90 patients with SSc [8]. Mortality was monitored as part of the 3-year clinical trial (criterion validity). This summary index consists of 33 variables which are grouped into 10 domains representing a composite of disease activity and severity. However, this index lacks a global assessment score and is too complicated for use in routine clinical practice.

The Medsger disease severity scale is the most widely used severity scale used in SSc. In 1999, Medsgar et al. [3] developed the scale based on 78 patients who were prospectively recruited from 14 international centers. The items for the Medsger severity scale were then correlated with mortality (criterion validity) using the University of Pittsburgh databank (n = 579) (Table 2). The authors defined severity as the total effect of disease on organ function, which has both reversible (activity) and irreversible (damage) components. The scale describes disease involvement in nine organ systems (general health, peripheral vascular, skin, joint/tendon, muscle, gastrointestinal tract, lungs, heart, and kidneys). Each organ system is graded from 0 (no involvement) to 4 (end-stage involvement). This index was revised in 2003 to incorporate amendments, such as including hemoglobin, defining pulmonary hypertension in terms of a systolic pulmonary arterial pressure (sPAP) threshold and removing proteinuria from the severity scale.

**Fig. 1.** Flow diagram showing the results of the search strategy and the inclusion of additional studies from bibliographies in the final review.
OMERACT system of 0 (normal) to 4 (very severe) (Table 2) [6].

Several other disease severity indices for SSc have been developed based on earlier work by Medsger et al. [3,4]: Geirsson and co-workers developed a five organ system severity scale based on the Medsger disease severity scale in an observational study with 100 consecutive patients with SSc in Lund, Sweden and correlated with mortality (criterion validity) (Table 2) [5]. This modified scale excluded several items due to missing data on the peripheral vascular system, joints/tendons, muscles, and heart (partially validated content validity). A similar disease severity index for SSc was developed in Japan based on eight organ systems (partially validated content validity). A similar disease severity index for SSc was developed based on earlier work by Medsger et al. [3,4]: Geirsson and co-workers developed a five organ system severity scale based on the Medsger disease severity scale in an observational study with 100 consecutive patients with SSc in Lund, Sweden and correlated with mortality (criterion validity) (Table 2) [5].

Other disease severity indices are yet to be shown to have criterion and content validity. Internal and external validation studies are currently under way. The damage index is being developed based on the World Health Organization International Classification of Functioning, Disability, and Health framework [25].

Outcome measures for use in SSc clinical trials

The Scleroderma Clinical Trials Consortium conducted a three round Delphi exercise to develop a combined response index for use in clinical trials in SSc (CRISS) [10] and a similar method was used for Pulmonary Arterial Hypertension related to Systemic Sclerosis (EPOSS) [11]. These use a combination of patient- and physician-reported measures as well as objective parameters, to measure response to treatment.

OMERACT filter

The outcome measures CRISS and EPOSS have face and content validity. Internal and external validation studies are currently under way.

OMERACT filter

The outcome measures CRISS and EPOSS have face and content validity. Internal and external validation studies are currently under way.
Validation of composite measures of health status in systemic sclerosis according to the OMERACT

<table>
<thead>
<tr>
<th>OMERACT filter</th>
<th>Disease activity indices</th>
<th>Disease severity indices</th>
<th>Outcome measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Valentini et al. [1]</td>
<td>Medsger et al. [3,4]</td>
<td>CRISSt</td>
</tr>
<tr>
<td>Validity</td>
<td>Face validity Y Y Y Y Y Y Y Y</td>
<td>Y Y Y Y Y Y Y Y</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Content validity PVY PVY PVY PVY PVY PVY</td>
<td>Y Y Y Y Y Y Y Y</td>
<td></td>
</tr>
<tr>
<td>Discrimination capacity</td>
<td>Reliability ND ND ND ND ND ND ND ND</td>
<td>ND ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sensitivity to change PVF ND ND ND ND ND ND ND ND</td>
<td>ND ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>External validation Y [17] ND Y ND ? ? ND ND IPR IPR</td>
<td>ND ND</td>
<td></td>
</tr>
</tbody>
</table>

*: insufficient data to evaluate item.
I PR: in progress.

Organ-specific indices

Musculoskeletal system

The disease activity indices developed by Valentini et al. [1] and Minier et al. [2] use “arthritis” and self-reported measures of symptoms and functional status to assess musculoskeletal disease activity in SSc (Table 4). The disease severity indices use a combination of musculoskeletal measures which represent disease activity (such as tender joint count) and disease damage (joint contractures and radiographic evidence of joint damage) [3–9]. The outcome measure developed by the CRISS group uses tender joint count and tendon friction rubs to assess the musculoskeletal system in clinical trials for SSc [10]. Notably, this index does not include a patient-reported outcome measure for the musculoskeletal system.

We found 27 instruments which assess the effects of SSc in the various components of the musculoskeletal system including joints, tendons, bones, and muscles (Supplementary Table 3 [19,26–29]). With respect to joints, Clements et al. [19] published a comprehensive list of outcome measures for patients with arthritis in SSc which were characterized according to the OMERACT filter. However, the authors found that there is no single validated method for assessing tender and swollen joint count to measure joint involvement in SSc.

Contractures in SSc can be related to disease affecting joints and/or tendon, skin, or muscle. The finger-to-palm (FTP) distance has been widely used to assess finger range of movement (ROM) [3,4]. It has criterion and convergent construct validity as a measurement of disease severity but it has significant inter-observer variability [4,19]. In addition, Clements et al. [19] raised concerns regarding its face validity as it underestimates finger contracture in patients with SSc whose fingers are “fixed” in flexion with extremely limited finger ROM. Delta FTP distance has been proposed to avoid potential misclassification [30] but further studies are needed to validate delta FTP as a reliable outcome measure in SSc.

Avouac et al. [31] reported a high prevalence of joint contractures of 31% in the cross-sectional EUSTAR database (n = 7286). The prevalence of large joint contractures in two randomized controlled trials was 46.5% [Scleroderma Lung Study (SLS)] [32] and 55.6% [d-penicillamine (d-pen)] study [33], which recruited patients with early diffuse disease. Both SLS and d-pen studies reported no change in large joint contracture over 24 months in either treatment arm. It is unclear whether large joint contracture is irreversible but it lacks criterion and construct validity as well as discrimination capacity as a measure of organ damage.

Most indices of hand function have not been validated as measures of disease activity and severity according to the OMERACT filter in terms of criterion and construct validity as well as discriminant capacity. The Hand Mobility in Scleroderma (HAMS) instrument has been validated in small studies [34,35] with good to excellent reliability for most of the items [36]. The Cochin hand function scale has validity, discrimination and feasibility and has high correlation to hand-related items in HAQ [Spearman’s rank correlation coefficient (rs) = 0.858] and moderate correlation to the physical summary score in SF-36 (rs = –0.321) [37]. Clements et al. [19] found that only the HAQ-DI, SF-36 and Cochin Hand Function scale possess validity, discrimination and feasibility, and are ready for use in clinical trials. While loss of hand function due to damage is likely to be captured by these scales, they are unlikely to discriminate easily between damage and activity.

Proximal muscle weakness, myositis and myopathy have been included in a disease severity score developed by Furst and colleagues but these measures need further validation before they can be recommended for use in clinical trials [8]. Serum creatine phosphokinase and aldolase have been selected for inclusion in the CRISS response index for use in clinical trials of SSc [10]. However, these biomarkers, which reflect disease activity rather than damage, are only partially validated against alanine transaminase (ALT), aspartate aminotransferase (AST), and C-reactive protein (CRP) for myopathy in SSc [38]. The Scleroderma Clinical Trials Consortium (SCTC) Working Group for Myositis is currently developing an instrument to assess myositis in SSc.

Vascular system

Most of the existing composite indices contain objective and subjective variables to assess vascular manifestations in SSc.
Digital ulcer and self-reported changes in vascular manifestations have been used as markers of disease activity in SSc (Table 5). Interestingly, digital necrosis is included in the indices of disease activity though from a biological perspective, it also represents damage. None of the indices of health status in SSc include a definition of digital ulcer. This may be due to a lack of consensus on the assessment of digital ulcers, which limits the reliability of digital ulcers as an outcome measure in SSc. Furthermore, the existing composite indices have not included digital amputation (auto-amputation or surgical amputation) as an outcome measure despite the significant morbidity associated with this marker of damage.

Table 5
Summary of measures of health status for the vascular system used in composite indices for systemic sclerosis (SSc)

<table>
<thead>
<tr>
<th>Measures of disease status for vascular system</th>
<th>Disease activity</th>
<th>Disease severity</th>
<th>Outcome measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital pitting scar</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Digital ulcer</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Digital gangrene/necrosis</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Subjective measures</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

A: disease activity; D: disease damage; S: disease severity.

* Combined Response Index in Systemic Sclerosis.

b Expert Panel on Outcome Measures in Pulmonary Arterial Hypertension related to Systemic Sclerosis.

c Wrist, elbow, and knee.

d Serum creatine phosphokinase.

e Serum aldolase.

f Self-reported change in “vascular manifestations in the past month.”

(Tables 2 and 5). Digital ulcer and self-reported changes in vascular manifestations have been used as markers of disease activity [1,2] in SSc (Table 5). Interestingly, digital necrosis is included in the indices of disease activity though from a biological perspective, it also represents damage. None of the indices of health status in SSc include a definition of digital ulcer. This may be due to a lack of consensus on the assessment of digital ulcers, which limits the reliability of digital ulcers as an outcome measure in SSc [39,40]. Furthermore, the existing composite indices have not included digital amputation (auto-amputation or surgical amputation) as an outcome measure despite the significant morbidity associated with this marker of damage.
Raynaud’s phenomenon is not included in any of the indices of disease activity but it has been incorporated into a number of disease severity scores as part of the spectrum of SSC-related vascular complications [3,4,6,9]. However, in a recent online survey of 443 patients with primary and secondary Raynaud’s phenomenon (RP), Hughes et al. [41] reported that 58% of the respondents could predict at least 51% of RP attacks. Furthermore, 57% could predict attack severity either fairly well or better.

Cutolo et al. [42] described three nailfold capillary patterns (early, active, and late phases) in SSC, using nailfold videocapillaroscopy (NVC), which reflect progression of microscopic vascular damage (Supplementary Table 4). In a systematic review, Ingegnoli et al. [43] reported that the extent of SSC microangiopathy correlated with the degree and severity of vascular complications such as telangiectasia, digital ulcers, and pulmonary arterial hypertension. Furthermore, Kayser et al. [44] demonstrated that an avascular score > 1.5 on nailfold capillaroscopy was an independent predictor of mortality (criterion validity) in a retrospective study involving 135 patients with SSC. Despite the suggestion that late capillaroscopic changes are an indicator of organ damage in SSC, recent studies have reported an improvement in the microvascular pattern from a late to an active pattern following treatment with bosentan and hemopoetic stem cell transplantation [43]. This suggests that the microvasculature has a capacity for remodeling in response to treatment.

In a study of 107 SSC patients who underwent NVC, Sambataro et al. [45] reported that the number of micro-hemorrhages (MHE) or micro-thrombosis (MT) (collectively known as NEMO), giant capillaries (GC), and reduced capillary score (CS) were associated with Valentini disease activity score and mRSS. In addition, Sambataro et al. proposed a modified NEMO score (mNEMO) that is defined as having ≥ 6 MHE/MT, or 3–5 MHE/MT, and at least three GCs to predict the presence of moderate to high disease activity in SSC (Valentini disease activity score ≥ 3).

Overall, nailfold capillaroscopy to assess microvascular activity and damage in SSC fulfills reliability, reproducibility, and feasibility criteria (Supplementary Table 4) [43]. However, novel NVC scoring systems, including modified NEMO score [45] (number of micro-hemorrhages or micro-thrombosis) requires longitudinal studies to assess its impact on outcomes such as mortality (criterion validity), construct validity, reliability, and sensitivity to change (Supplementary Table 4). Further validation of criterion validity to measure disease damage in SSC is required.

Based on self-reported measures of sexual function, there is a high prevalence of sexual dysfunction in SSC (81% in men and 62% in women) [46,47]. There are limited data on erectile dysfunction (ED), which is thought to be a vascular manifestation. In the EULAR Scleroderma Trial and Research database (n = 130 men), 39% had severe erectile dysfunction and only 17.7% had normal erectile function [46]. Patients with any form of ED were found to have a higher EULAR activity score and more severe skin, muscular, lung, and vascular involvement, compared to those without ED [46]. Furthermore, the authors reported that the median time interval from the onset of the first non-Raynaud’s phenomenon to the onset of ED was only 4.1 years. In a smaller study, Rosato et al. [47] used color Doppler to assess disease severity and microvascular damage in 20 men with moderate to severe ED and either limited or diffuse SSC. The authors found a shorter duration of SSC and Raynaud’s phenomenon (RP) in the low vascular damage group (early or active capillaroscopic pattern with a vascular score on the Medsger severity scale ≤ 2) compared to the high vascular damage group (late capillaroscopic pattern and vascular score ≥ 3) [47]. This finding suggests that microvascular damage accumulates over time.

**Skin**

Table 6 summarizes the various clinical and non-clinical methods of assessing skin involvement in SSC [48–54]. The modified Rodnan skin score (mRSS) correlates with the collagen content of the skin and it remains the gold standard (criterion validity) for assessing skin thickening in SSC [48]. It is feasible, easy to use, reliable and valid, and modestly responsive to change although it may be less valid in late disease as mRSS may not take skin atrophy into account (damage) [55].

Skin thickness progression rate (STPR) and latent linear trajectory model (LTM) have been reported as tools predictive of internal organ involvement and mortality [56–58]. Data from the Pittsburgh cohort showed that rapid STPR (≥ 40 units per year) is an independent predictor of mortality and renal involvement in early diffuse SSC and those with anti-topoisomerase I antibody compared to those with intermediate (15–40 units/y) or slow STPR progression rate.

Table 6

<table>
<thead>
<tr>
<th>Types of clinical assessment of skin involvement in SSC</th>
<th>Anatomic sites</th>
<th>Comments</th>
<th>Scale</th>
<th>Range of total score</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rodnan skin score [48]</td>
<td></td>
<td></td>
<td>0–4</td>
<td>0–104</td>
<td>Y</td>
</tr>
<tr>
<td>Stage 1 (edematous phase)</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 2 (indurative phase)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 3 (atrophic phase)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modified Rodnan skin score [49]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silman et al. [50]</td>
<td>17</td>
<td></td>
<td>0–3</td>
<td>0–51</td>
<td>Y</td>
</tr>
<tr>
<td>Geirsson et al. [5]</td>
<td>17</td>
<td></td>
<td>0–2</td>
<td>0–34</td>
<td>ND</td>
</tr>
<tr>
<td>Furst et al. [8]</td>
<td>24</td>
<td></td>
<td>0–3</td>
<td>0–72</td>
<td>ND</td>
</tr>
<tr>
<td>Kahaleh et al. [51]</td>
<td>10</td>
<td></td>
<td>0–3</td>
<td>0–30</td>
<td>Y</td>
</tr>
<tr>
<td>Modified Rodnan skin score</td>
<td>22</td>
<td></td>
<td>0–3</td>
<td>0–66</td>
<td>ND</td>
</tr>
<tr>
<td>Face</td>
<td>Rest of sites</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCP &amp; PIP</td>
<td>0, 2, and 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distal to PIP</td>
<td>0 or 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Photopheresis trial [52]</td>
<td>15</td>
<td></td>
<td>0–3</td>
<td>0–45</td>
<td>ND</td>
</tr>
<tr>
<td>% body surface area [52]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gruber et al. [53]</td>
<td>19</td>
<td></td>
<td>0–3</td>
<td>0–100</td>
<td>ND</td>
</tr>
<tr>
<td>Localized scleroderma skin damage index [54]</td>
<td></td>
<td></td>
<td>0–3</td>
<td>0–57</td>
<td>ND</td>
</tr>
<tr>
<td>Dermal atrophy</td>
<td></td>
<td></td>
<td>0–3</td>
<td>0–162</td>
<td>Y</td>
</tr>
<tr>
<td>Subcutaneous atrophy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dermal pigmentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MCP: metacarpal joint; PIP: proximal interphalangeal joint; ND: no data.
( < 15 units/y) [56,57]. Furthermore, Shand et al. [58] showed that survival was the lowest in the subgroup of early diffuse patients with a high baseline skin score with little improvement during the 3-year follow-up compared with other subgroups.

The Localized Scleroderma Skin Damage Index (LoSDI) has been included in this review, even though it is designed for localized scleroderma in children, as it is a potential alternative clinical assessment tool to measure the extent of skin involvement in SSC [54]. The LoSDI is a sum of the following three scores for cutaneous features of skin damage: dermal atrophy (DAT), subcutaneous atrophy (SAT), and dyspigmentation (DP), measured at 18 anatomic sites. Disease damage was defined as irreversible changes or persistent changes of the lesion beyond 6 months due to previous active disease or complications of therapy. Based on a study with 30 children with localized scleroderma and 112 lesions, Arkachaisri et al. [54] demonstrated moderate to excellent intra- and inter-rater reliability and content validity and moderate correlation with the physician global assessment for disease damage when applied to SSC-related skin damage by 10 experts. Interestingly, all experts considered skin thickness as either completely reversible (43%) or reflecting activity more than damage. In addition, all experts agreed that DP was more indicative of damage than activity, and SAT and DAT were considered to be completely irreversible reflecting damage (29% and 57%, respectively) more than activity (71% and 43%, respectively) [54]. It is unclear whether these findings are generalizable to limited or diffuse SSC in adults.

Durometry is an objective method of measuring skin hardness, which may be affected by skin thickness, skin density, elasticity, and edema. A review by Furst et al. [59] suggested that durometry has full face and content validity as a measure of disease severity and damage, inter- and intra-reliability and feasibility, but it is only partially validated with regards to discriminant capacity (Supplementary Table 5).

In a systematic review, Ch'ng et al. [60] identified 17 studies which assessed ultrasound as an objective outcome measure of skin changes in SSC. Three of the 17 studies showed excellent intra- and inter-observer variability. There are issues with face validity as most studies did not show any correlation between ultrasound findings and the local mRSS in the region of interest, although a few studies reported a correlation with the global mRSS. Only one study assessed criterion validity against skin biopsy findings, which showed modest sensitivity of 64.6% and specificity of 100%. Ch'ng et al. [60] identified five studies that assessed sensitivity to change using ultrasound. Four of five studies reported change in skin thickness using ultrasound over time or with therapy. However, the authors noted a number of issues relating to the use of ultrasound as a potential outcome measure of skin thickness in SSC and they are as follows: (1) standardization of the acquisition of images (machine settings, number, and sites for assessment), (2) definition of skin thickness measured using ultrasound, and (3) the determination of reference values and variations with age and ethnicity.

Skin biopsy has been suggested as a possible outcome measure for clinical trials to assess whether novel therapies for SSC have the potential to attenuate and/or reverse skin fibrosis [61]. However, skin biopsy is more invasive and may not be feasible in routine clinical care.

Cardiac system

The disease activity indices developed by Valentini and Minier do not incorporate objective measures of cardiac involvement in SSC but rely on self-reported change in “cardiopulmonary manifestations” in the past month (Table 2) [1,2]. In contrast, most of the indices for disease severity include objective measures to assess cardiac-related complications in SSC. The combined response indices developed by the CRISS [10] and EPOSS groups [11] incorporate a number of objective and subjective cardiopulmonary assessments in SSC. In particular, the EPOSS index includes cardiopulmonary measures such as the 6-minute walk test (6MWWT), echocardiography, and right heart catheterization (RHC) [11].

The EPOSS group subsequently reported in a systematic review that the 6MWWT has been partially validated according to the OMERACT filter for measuring disease activity and severity with respect to face validity, discriminant capacity with treatment and feasibility but that it lacks content validity as it is not specific for SSC-PAH [62]. They suggested further studies to validate both criterion and construct validity, and reproducibility, taking into account the co-morbidities which may confound the test. In some clinical contexts, a reduced 6MW distance may be irreversibly decreased due to damage but even in an irreversible complication such as PAH, it can improve with therapy.

In another systematic review by the EPOSS group, Kowal-Bielecka et al. [63] reported that echocardiography has face validity based on expert opinion and the high number of studies using echocardiography to evaluate SSC patients. It was considered to be partially validated with regard to criterion validity based on the correlations between echocardiographic and RHC parameters in those at risk of PAH and pulmonary hypertension overall. However, it was deemed to lack content validity as a measure of disease activity and severity because the assessment of sPAP is not technically possible in some patients. In addition, it is unclear whether echocardiography has the discriminant capacity and responsiveness required to measure change over time. In fact, measures such as sPAP may decrease in late stage disease despite progressive pulmonary vascular dysfunction due to right ventricular failure. The authors suggested further studies to fully validate echocardiography as an outcome measure in SSC. Changes on echocardiogram may be irreversible in many cases but even in PAH, therapy can lead to improvement in right ventricular function due to remodeling [64].

Right heart catheterization (RHC) fulfils face, content, construct and criterion validity but there are insufficient data to suggest that it is sensitive to change. Despite it being the gold standard to diagnose PAH, RHC has not been fully validated as an outcome measure (Supplementary Table 6).

A weighted cardiac score was developed by Clements et al. [22] that is predictive of survival (Supplementary Table 6). The cardiac score consists of two variables: left axis deviation on ECG and moderate to severe pericardial effusion; this was derived from the cardiac score which was developed by Clements et al. [65] more than 20 years ago based on a small cohort of patients with SSC (n = 48).

In a large case series of patients with SSC (n = 183), Kostis et al. [66] found a high prevalence of supraventricular ectopic beats (67%), supraventricular arrhythmias (21%), ventricular ectopic activity (67%), and ventricular tachycardia (7%). Cardiac arrhythmias were strongly associated with mortality [66]. In another case series of SSC patients (n = 50), Roberts et al. [67] reported that 40% had significant supraventricular arrhythmias and 40% had significant ventricular arrhythmias. Furthermore, Roberts et al. showed that 14% had second or third degree conduction defects.

In a small retrospective study involving 25 SSC patients with suspected myocardial involvement, Mueller et al. [21] demonstrated a trend for the occurrence of any cardiovascular event, defined as cardiovascular death, arrhythmic endpoints or rehospitalization due to heart failure, and the degree of inflammation as well as fibrosis detected on endomyocardial biopsy. Larger studies in SSC are needed to show that cardiac inflammation (activity) causes fibrosis (damage) resulting in these cardiac endpoints, and premature mortality.
Pulmonary system

Composite indices for disease severity by Medsger et al. [3,4] and Minier et al. [2] incorporate objective measures of pulmonary function in SSC (Table 2). However, CRISI [10] and EPOSS groups [11] include a number of objective and subjective measures of cardiopulmonary function in SSC as part of the response index for clinical trials.

In SSC-related ILD, loss of lung function in the early stages and extent of fibrosis assessed radiographically at baseline appear to be the most important prognostic variables [68–71]. Forced vital capacity (FVC) and diffusion capacity for carbon monoxide (DLCO) are routinely used to assess lung function, screen for ILD, and to assess for disease progression in SSC. The optimal time interval between serial pulmonary function tests (PFT) in SSC-ILD to assess for trends in lung function is yet to be determined with ranges from three months in the Scleroderma Lung Study to 24 months for stable, chronic disease [69,72]. The generally accepted threshold for a significant and reproducible change in FVC and DLCO compared to the baseline absolute values is 10% and 15%, respectively [73].

High resolution computed tomography (HRCT) has become the gold standard for the diagnosis of SSC-related interstitial lung disease (SSc-ILD) but there is no consensus regarding the optimal scoring system for SSc-ILD [74]. The following four types of scoring system to assess SSc-ILD have been developed: comparative [75], simplified [76,77], semi-quantitative [68,78–81], and quantitative methods [82,83] (Supplementary Table 7).

The comparative method by Wells et al. [75] evaluates one HRCT abnormality (parenchymal disease) relative to another (reticular disease), and these lesions are graded from 1 to 5 and correlated with lung histology. Morelli et al. [76] used a simpler scoring system, dividing the lung into three zones and assigning a score for the HRCT abnormality in each zone added together to produce a global HRCT score. In another simplified scoring system, Tochimoto et al. [77] assigned a score for the extent of honeycombing, bronchiectasis and bronchiolectasis in terms of the intensity of interstitial shadowing. An additional point was allocated if ground glass attenuation is seen.

Several semi-quantitative methods have been developed. Warrick et al. [78] combined the severity of lung involvement on HRCT and the extent, based on the number of bronchopulmonary segments involved. Ooi et al. [80] graded the extent of disease for each abnormality as ground glass opacity (GGO), reticular fibrosis, mixed ground glass and reticular disease, and honeycombing. Each lobe is graded separately and the global score is the summation of all scores. Pandey et al. [81] scored for the presence of ground glass opacification, reticular abnormalities and honeycombing in each of the five lobes and weighted each score for the relative lobar volume using correction factors.

Golchin et al. [79] used a semi-quantitative scoring method for the Scleroderma Lung Study, based on the extent of disease for each HRCT abnormality with follow-up HRCT after completing one year of the study. The global score for each lung is the sum of the grades in all six zones and correlated with mortality (criterion validity). Goh and colleagues proposed a semi-quantitative scoring method based on the total extent of disease, irrespective of appearance (reticular or GGO) [68]. The authors demonstrated that an extent greater than 20% is predictive of mortality (criterion validity). Hence, extensive and limited ILD were categorized as greater than or less than the 20% threshold respectively. Those staged as indeterminate (10–30% extent on HRCT) can be further stratified into extensive and limited disease based on less than or greater than an FVC threshold of 70% predicted, respectively. Goh’s work has been validated by Moore et al. [84] in a multi-center cohort of SSC patients (external validation). Furthermore, Moore et al. [84] showed that changes on serial HRCT did not correlate well with outcome events and no patients had a major change in the raw HRCT score without significant (>30%) decline in PFTs. This finding suggests that serial PFTs may be more appropriate for monitoring changes in SSC-ILD over time than serial HRCT.

Quantitative scoring methods provide a more precise assessment of the extent of ILD and reduce bias associated with visual estimation used in other scoring methods. Carotti et al. [82] recently used open-source dicom viewer software (OSIRIX) to assess ILD in SSc (n = 62). The authors found a high correlation between the radiologists’ visual scoring and the HRCT quantitative assessment of ILD after stratifying by the overall extent and fibrosis. However, they found no correlation between quantitative assessment and DLCO values. Nguyen-Kim et al. [83] used a histogram-based quantification of fibrosis in SSC based on nine of the HRCT slices in 37 patients and reported similar findings. They found a high correlation between visual scoring and the histogram-based quantification method. In addition, Nguyen-Kim et al. [83] showed a correlation between the degree and severity of ILD with FVC and Forced Expiratory Volume (FEV1) but not with Total Lung Capacity (TLC) or DLCO.

There has been a growing interest in lung ultrasound (LUS) to screen, and to assess for progression of SSc-related ILD. LUS detects and quantifies the lung comet tails signs (B-lines) due to thickened inter-lobular septa [85]. The number of B-lines from LUS has been shown to correlate with lung involvement in HRCT in previous studies [86]. Cappelli et al. [85] proposed using PFT (with DLCO) and LUS to screen for ILD every 1–2 years. If there is any suspicion of ILD, a HRCT is done to confirm the diagnosis. In addition, Cappelli et al. [85] suggested that in those with SSc-ILD, PFT (with DLCO) and LUS should be undertaken every 6 months to screen for changes or progression in ILD with the final diagnosis confirmed by HRCT. This algorithm using LUS reduces radiation risk and health care cost. However, longitudinal studies are needed to assess the accuracy of LUS in larger studies and to validate the algorithm. Furthermore, it is unclear if DLCO performs better than FVC as a screening tool for ILD, particularly if there are other confounders for reduced DLCO in SSC, such as pulmonary hypertension.

OMERACT filter

There is no universally accepted definition of damage due to lung disease in SSc. In the Pittsburgh University Scleroderma cohort (n = 890), Steen et al. [70] identified 531 (60%) with no or minimal lung disease (FVC > 75% predicted), 243 (27%) with moderate restrictive lung disease (FVC 50–75% predicted), and 11 (13%) with severe restrictive lung disease (FVC < 50% predicted). In a subset of patients (n = 152) who had moderate to severe lung disease (FVC = 50–75% predicted) at baseline, 55 had lung function testing at baseline and repeated within 2 years. Thirty of the 55 (54.5%) patients had an annualized rate of decline in FVC of 32% in the first two years following disease onset suggesting that lung damage occurs early in the disease [70]. DLCO paralleled the decline in FVC in all three groups of restrictive lung disease. Similar findings were reported in the Scleroderma Lung Study [69].

Khanna et al. [69] showed that those with severe pulmonary fibrosis on baseline HRCT had a greater decline in the FVC % predicted compared to the group with no or moderate fibrosis. Furthermore, the authors showed that this impact was most evident during the first 2 years following disease onset.

In a review of the various HRCT scoring methods, Assayag et al. [74] reported variable inter-observer agreement. In the Scleroderma Lung Study, Golchin et al. [79] reported variable inter-observer agreement for the semi-quantitative method at baseline,
which was higher for ground glass opacification (GGO) compared to “fibrosis” and honeycombing (HC) ($\chi^2 = 0.72, 0.61$ vs $0.39$, respectively). However, the same authors reported higher interobserver agreement for HRCT to detect change over 12 months for fibrosis compared to GGO and HC ($\chi^2 = 0.51, 0.36$ vs $0.16$, respectively) (partially validated sensitivity to change). These measures have the potential to be used as indicators of damage but have not been validated for this purpose.

The Connective Tissue Disease-Associated Interstitial Lung Disease (CTD-ILD) Working Group of Outcome Measures in Rheumatology recently published a preliminary core set of outcome measures consisting of 18 instruments grouped into five domains including dyspnea/cough, health-related quality of life, lung imaging, lung physiology/function, and survival [87]. These outcome measures are designed to assess the response to new therapies in clinical trials and have been developed using a combination of Delphi, Nominal Group Technique and focus group meetings with patients, and cluster analysis. FVC had $100\%$ acceptance as the most reliable serial variable in CTD-ILD.

**Gastrointestinal system (GI)**

All indices for disease severity incorporate a spectrum of objective measures of GI involvement in SSc. Notably, none of the indices for disease activity has a GI component. This is surprising given that this is the most common internal organ involvement in SSc with significant impact on morbidity and/or mortality [88], but presumably reflects the difficulties encountered in measuring GI activity and damage.

Our systematic review did not identify an organ-specific objective assessment tool for gastrointestinal manifestations in SSc. The CRSS group suggested using a validated GI patient-reported outcome measure for clinical trials in SSc [10]. The UCLA 2.0 GIT instrument is a validated self-reported questionnaire with 34 items grouped into seven domains with satisfactory reliability (Intraclass Correlation Coefficient is 0.67) [89]. Items are scored from 0–3 with lower scores indicating better health-related quality of life. It has convergent and divergent validity compared with six self-reported indices of severity of GI disease. However, it is unclear if it correlates well with objective measures of GI manifestations in SSc.

The Malnutrition Universal Screening Tool (MUST) is a validated screening tool for malnutrition which can be used in SSc [90]. The MUST tool has face and content validity, and feasibility to measure disease severity. However, it does not have construct validity as it is not specific for SSc. It remains unclear whether the MUST tool has criterion validity as well as sensitivity to change in measuring disease severity.

**Renal system**

The optimal measure of renal function in SSc remains unknown [59,91]. Supplementary Table 8 shows the variables commonly used to assess renal manifestations in SSc. Only blood pressure has been validated as an outcome measure for scleroderma renal crisis although its face validity may be affected by patients with normotensive renal crisis [91]. Proteinuria is a predictor of mortality and is related to impaired renal function, but like blood pressure, it is not specific for renal disease in SSc (construct validity) and has not been validated as an outcome measure for SSc clinical trials [91].

Calculations of the estimated glomerular filtration rate (eGFR) have been shown to be a reliable measure of renal impairment in SSc. Serum creatinine may be affected by muscle mass while the direct measurement of glomerular filtration rate (GFR) is expensive and not feasible in routine clinical care. The Cockcroft-Gault or the Modification of Diet in Renal Disease equations are relatively accurate in measuring the estimated GFR (eGFR) although neither has been fully validated in SSc [59]. There is no universally accepted definition of renal damage due to SSc although scleroderma renal crisis is likely to indicate an irreversible component. In the University of Pittsburgh database ($n = 807$ with diffuse disease), Steen et al. [92] identified 145 patients who had scleroderma renal crisis; 89 (61%) had good outcomes (55 did not require dialysis, 34 received temporary dialysis), and 28 (19%) required permanent dialysis. Of the 34 patients (23%) who required temporary dialysis, 53% did not require long-term dialysis and the time for discontinuation ranged from 2 to 18 months (mean of 8 months) [92]. Interestingly, survival was reduced compared to those who did not receive dialysis, even among patients who had temporary dialysis.

In a retrospective analysis of the Australian and New Zealand Dialysis and Transplant (ANZDATA) registry containing 40,238 patients who commenced dialysis, Siva et al. [93] identified 127 (0.3%) patients with end-stage kidney disease (ESKD) secondary to SSc. In addition, the authors found that patients with ESKD due to SSc were more likely to recover renal function and most of the recovery occurred in the first 12–18 months following commencement of dialysis [93]. Although, seven of the 13 SSc patients who recovered dialysis-independent renal function died, and four returned to dialysis, Sharmugam et al. [94] suggested postponing renal transplantation until the patient is dependent on dialysis for at least two years given the relatively high rate of renal recovery.

A minority of patients with SSc develop renal abnormalities other than renal crisis. Decreased GFR is associated with other manifestations such as hypertension and cardiac involvement indicating possible pre-renal causes. In a separate study, Steen et al. [95] reported that 49 of 675 (7.3%) patients with diffuse disease had transient elevation in serum creatinine. The serum creatinine returned to 1.2 mg/dL (106.1 μmol/L) within 1 year of the abnormal value. Furthermore, 91 (13.5%) of the cohort developed proteinuria, most of which was due to β-lactam/metalloproteinase toxicity and proteinuria, of which 55/87 (82%) cases resolved within one year. None of the patients with proteinuria developed renal crisis, renal insufficiency, or nephrotic syndrome.

**Discussion**

**Composite measures of health status in SSc**

**Composite measures of disease activity in SSc**

The lack of clearly defined episodes of flare and remission in SSc has made it difficult to describe and quantify disease “activity.” Hence, there is no gold standard instrument for measuring disease activity in SSc.

Currently, there are two indices of disease activity in SSc, of which the Valentini disease activity score is the more commonly used in research studies. However, Hudson et al. [38] highlighted a number of limitations with the Valentini disease activity index. Firstly, two-thirds of the cohort from which it was derived, had long-standing disease. Diffuse SSc, in particular, demonstrates higher disease activity in the initial four years after disease onset [96]. Hence, disease activity is better assessed in a cohort with early disease. Furthermore, Hudson et al. [38] noted a high number of missing values and high variability of frequency of various SSc manifestations among the study centers.

The disease activity index proposed by Minier et al. [2] represents a variation of that by Valentini et al. [1]. Although the addition of lung-related items and the alteration in the weighting of other items seems intuitive, it remains unclear if the revised index provides an advantage over the Valentini disease activity index [1]. Furthermore, both disease activity indices proposed by Valentini et al. [1] and Minier et al. [2] lack renal and
gastrointestinal items. In addition, some of the items, such as digital necrosis, may reflect disease damage as well as disease activity.

**Composite measures of disease severity in SSc**

The Medsger disease severity index is the most commonly used measure of disease severity in SSc [3,4]. However, disease severity does not differentiate between potentially treatable disease activity and irreversible organ damage. The Medsger severity scale has some limitations. Firstly, it is designed to include activity and damage, but excludes some variables that may reflect activity, such as the erythrocyte sedimentation rate (ESR) and tendon friction rubs. Secondly, there is no component quantifying overall damage. In addition, the items included in the scale predict mortality but not morbidity. Some items, such as anal incontinence, which are a common manifestation in SSc were excluded from the Medsger severity scale as this may be because of collinearity with other items, or it could mean that the index may not adequately reflect severity in all organ systems. Furthermore, the Medsger disease severity index requires more studies to assess the construct validity, reliability and sensitivity to change.

Other indices developed by Geirsson et al. [5] and Morita et al. [6] represent variations of the Medsger disease severity index. Furst et al. [8] developed a comprehensive multi-organ severity index for SSc but this has questionable feasibility for use in routine clinical practice. Hence, these indices are not commonly used in research or in clinical practice.

**Measures of disease damage in SSc**

This systematic review highlights the lack of a disease damage index to identify and quantify organ damage in SSc. There are a number of potential applications of a damage index in SSc. It could be used as an outcome measure or end-point in observational studies. The damage index could also enable identification of variables that are associated with earlier and greater accrual of damage. The damage index is also potentially useful for comparison of cohorts, enrichment of studies for SSc patients with more or less severe disease, as well as for measuring burden of disease for epidemiological studies to guide policy-making. In the absence of a gold standard for disease activity in SSc, a damage index in SSc could be used as a potential outcome measure in clinical trials (i.e., did the drug prevent damage accrual?), and to measure interval damage progression in SSc.

The relationship between disease activity and damage in diffuse and limited SSc is illustrated in Figure 2. In diffuse SSc, there is an initial peak in disease activity possibly due to active involvement of skin, musculoskeletal, lung, cardiac, gastrointestinal, and/or renal domains. A subsequent peak might occur 5–10 years later due to active disease in the vascular (digital ulcers, calcinosis, telangiectasia, and pulmonary arterial hypertension), gastrointestinal, and lung domains. In contrast, limited SSc is generally a slow progressive disease. Active disease in limited SSc is seen in the skin and peripheral vascular system first, followed by musculoskeletal and gastrointestinal domains. In the later course of the disease, there may be more active disease in the gastrointestinal and vascular domains causing digital ulcers, calcinosis, telangiectasia, and pulmonary arterial hypertension. In both diffuse and limited SSc, there is a time lag between disease activity and accrued organ damage as damage is the result of past activity.

**Outcome measures for use in SSc clinical trials**

CRISS [10] and EPOSS [11] are combined response indices which are developed for use in clinical trials in SSc and SSc-related PAH, respectively. However, criterion and construct validity, reliability, and sensitivity to change are yet to be assessed for CRiSS and EPOSS. Internal and external validation studies are currently under way.

Activity is defined as disease processes which are potentially reversible spontaneously or with treatment. Damage is defined as permanent and irreversible loss of anatomical structure or physiological function due to the disease itself, or its treatment. Severity is defined as the total impact of disease on organ function due to disease activity and damage (i.e., area under the curve for activity and damage).

**Organ-specific indices**

*Musculoskeletal system*

There is no validated instrument to assess joint involvement in SSc, especially irreversible changes. In a systematic review, Clements et al. [19] recommended the tender and swollen joint count, and the Cochin hand function for inclusion as primary outcome measures for musculoskeletal involvement in SSc clinical trials. In addition, they proposed using synovitis, tendon friction rub, HAMIS hand scale, HAQ-DI, SF-36, physician and patient globalVAS as well as patient pain VAS as secondary outcome measures in clinical trials. However, apart from HAQ-DI, SF-36, and the Cochin hand function scale, none of the other measures has been fully validated for construct and criterion validity, reliability, and sensitivity to change. In contrast, CRiSS proposed using tender joint count, tendon friction rub, serum creatine phosphokinase, and aldolase as the response criteria for clinical trials in SSc [10]. These outcome measures reflect disease activity and severity but do not define disease damage in SSc. Items such as joint
contracture and muscle weakness are indicators of disease damage for musculoskeletal system in SSc. Hand function indices, such as HAMIS [34] and the Cochin hand function scale [97], measure disease severity. However, they do not differentiate activity from damage, or limited from diffuse SSc.

Although MRI and ultrasound are reportedly more sensitive for detecting SSc-related arthritis and tenosynovitis than clinical measures, larger studies are needed to assess if the additional cases of synovitis and tenosynovitis detected on imaging only result in functional disability and impaired health-related quality of life.

Vascular system
There is a spectrum of digit-related vasculopathy complications, ranging from Raynaud’s phenomenon to digital gangrene, which may reflect disease activity, severity, or damage in SSc. Digital ulcer lacks a consensus definition. Furthermore, digital ulcers are often chronic in SSc and reflects a combination of disease activity, damage, and severity. Digital amputation represents disease damage but it has not been used as an outcome measure in SSc.

Nailfold videocapillaroscopy (NVC) is increasingly becoming an important tool to assess disease activity, damage, and severity in SSc with prognostic prediction for subsequent systemic vascular involvement and mortality. More studies are needed to identify and validate accurate NVC scoring systems to measure disease activity, damage, and severity in SSc. Raynaud’s phenomenon (RP), despite its disabling effect, appears to be under-estimated by SSc patients half of the time. Hence, this would limit self-reported RP as a reliable outcome measure of disease activity in SSc. Erectile dysfunction (ED) is common and under-reported in SSc. Further studies are needed to validate the various methods for assessing ED in SSc and its prognostication of morbidity and mortality.

Skin
The mRSS is the most widely used measure of skin involvement in SSc. Potential alternative measures of skin damage in SSc are dermal atrophy, subcutaneous atrophy, and dyspigmentation although these items require validation for use in adult-onset SSc. In addition to mRSS, skin thickness progression rate and latent linear trajectory models are tools predictive of internal organ involvement and mortality in SSc. Objective measures of skin involvement in SSc, such as dermometry and ultrasound, assess skin hardness and skin thickness, respectively. DERMometry is well validated for use in SSc but is not commonly used in clinical practice and in research. Ultrasound, however, still has a number of unresolved technical issues before it can be considered as a potential outcome measure of skin involvement in SSc.

Cardiac system
At present, the 6MWT, echocardiography, and right heart catheterization are the most commonly used measures of cardiac severity in SSc but none of the measures has been fully validated according to the OMERACT filter. However, there is no definition of cardiac damage in SSc. A weighted score developed by Clements et al. [22] based on left axis deviation on ECG and moderate to severe pericardial effusion is predictive of survival. Cardiac arrhythmias and conduction defects are common in SSc and are a significant cause of premature mortality in SSc. However, it remains unclear if these cardiac events arise from cardiac inflammation (activity) with subsequent fibrosis (damage) rather than occurring as a consequence of aging and cardiovascular risk factors. The Scleroderma Clinical Trial Consortium Cardiac Subcommittee has been formed to identify patients at high risk for cardiac arrhythmias and other SSc-related myocardial involvement [98].

Pulmonary system
HRCT and PFTs remain key parameters for the diagnosis, screening, and measurement of disease severity as well as organ damage in SSc-related ILD. However, the optimal scoring system to assess severity and damage in SSc-ILD is yet to be determined. Furthermore, there is an increasing interest in lung ultrasound (LUS) to screen for ILD, and to detect progression of disease in SSc. Confirmation of the diagnosis still requires HRCT. The CTD-ILD outcome measures assess the response to therapies in clinical trials [87]. Further studies are needed to validate the use of LUS and CT-ILD outcome measures in SSc-ILD.

Gastrointestinal system
There is a lack of objective composite and organ-specific indices to measure disease activity and disease damage for SSc-related gastrointestinal complications. In contrast, a number of disease severity indices have included some outcome measures of GI involvement in SSc. The UCLA 2.0 GIT instrument is well validated for use in SSc [89]. However, larger studies are needed to correlate the self-reported GIT score with objective measures of disease activity, severity and damage in SSc. The Malnutrition Universal Screening Tool (MUST) has been partially validated for use as a screening tool for malnutrition in SSc which has significant impact on morbidity and mortality in SSc [90].

Renal system
Blood pressure has been validated to evaluate scleroderma renal crisis in SSc, except in the case of normotensive renal crisis. Estimated GFR is the most reliable measure of renal impairment in SSc but it does not distinguish between disease activity, severity or damage. Established scleroderma renal crisis (SRC) is often irreversible (damage) as seen in histopathological studies [24], while early features leading to SRC are potentially reversible with angiotensin converting enzyme inhibitors and likely represents disease activity.

This study has a number of limitations. The systematic review covers a broad area in SSc and there is lack of specific MESH terms to search for articles relating to measures of health status. Hence, a number of the articles were not identified using the standard search strategy, and required a search of the bibliography of relevant articles as well as hand search. Furthermore, some of the less established measures of health status may not have appeared in the title, abstract or keywords. Hence, it is possible that some indices of health status in SSc may have been omitted.

Conclusions
This systematic review has identified a number of composite and organ-specific indices, which assess disease activity, disease severity, and response criteria for clinical trials in SSc using a combination of objective, or subjective parameters, or both. Notably, this review failed to identify any measure of disease damage in SSc, which may in part relate to the under-recognized importance of distinguishing organ damage accrual in SSc from disease activity. Most of these indices require further validation according to the OMERACT filter. Some of the composite indices for disease severity included investigational tools, which are not routinely available and hence, these indices may not be feasible for use in clinical practice. However, none of these indices measure disease damage per se in SSc. Further work should focus on the measurement of disease damage in individual organs and the impact this has on morbidity and/or mortality. An international collaborative endeavor led by the Scleroderma Clinical Trials Consortium Damage Index Working Group is currently
under way to develop and validate an index to quantify organ damage in SSc.

Acknowledgments

Australian Scleroderma Interest Group/Scleroderma Australia Clinical Research Fellowship (T.T. and N.F.); NHMRC Research Fellowship (APP1071735) and the University of Melbourne David Bickart Clinician Research Fellowship (M.N.); unrestricted grants from Actelion Australia, GSK and Pfizer.

We gratefully acknowledge the input of all members of the Damage Index Working Group of the Scleroderma Clinical Trials Consortium: Prof. Thomas Medsger, Prof. Peter Merkel, Prof. Janet Pope, Prof. Virginia Steen, A/Prof. Lorinda Chung, A/Prof.Damage Index Working Group of the Scleroderma Clinical Trials Consortium: Prof. Thomas Medsger, Prof. Peter Merkel, Prof. Janet Pope, Prof. Virginia Steen, A/Prof. Lorinda Chung, A/Prof. David Fiorentino, A/Prof. Anna Haemel, A/Prof. Ankoor Shah, Dr. Soumya Chatterjee, Dr. Tracy Frech, Dr. Jessica Gordon, Dr. Sindhu Johnson, Prof. Dinesh Khanna, Dr. John Pauling, Dr. Tatiana Rodriguez-Reyna, Dr. Lesley Saketkoo, Dr. Jim Seibold, and Dr. Gemma Strickland.

Appendix A. Supplementary material

Supplementary data are available in the online version of this article at http://dx.doi.org/10.1016/j.semarthrit.2016.07.010

References

In systemic sclerosis, the prevalence of ...


Chapter 5

Implications of the Findings

of this Thesis and the Future Directions
1.1 Implications from the Findings of this Thesis

There is no gold standard to assess disease activity in SSc. The European Scleroderma Study Group activity index was developed in 1999 (1) and revised in 2017 (2). It measures potentially reversible components of SSc. Furthermore, it lacks gastrointestinal and renal items. The Composite Response Index in Systemic Sclerosis (CRISS), on the other hand, contains several items to provide a single summary score to assess the efficacy of novel therapies in randomised controlled trials for diffuse cutaneous SSc (3).

There are a number of disease severity indices in SSc (4-10). However, disease severity is the total impact of disease activity and disease damage. Disease severity can improve with treatment and does not solely describe disease damage per se.

The findings from my study on the accrual of organ damage in early SSc highlight the need for a comprehensive definition of SSc-related organ damage for the development of a disease Damage Index using consensus and data-driven methodology. The early accrual of organ damage in SSc implies that disease modification early in the course of disease is of paramount importance if irreversible organ damage is to be prevented. The rapid accrual of organ damage means that a disease Damage Index could serve as a ‘sensitive’ and ‘responsive’ outcome measure in clinical trials of new therapies in SSc.

My systematic review did not identify an index to measure disease damage in SSc. Furthermore, the systematic review identified items for potential inclusion in a Damage Index. Previous measures of organ damage in other rheumatic diseases, such as Systemic Lupus International Classification Criteria / America College of Rheumatology Damage Index for SLE (11), and Vasculitis Damage Index for systemic vasculitis (12), have demonstrated the
prognostic importance of having a disease damage index for a multi-organ autoimmune disease.

Potential uses of a SCTC-DI include use as an outcome measure or end point in interventional and observational studies (i.e. does a new therapy prevent or delay the accrual of organ damage?), identify predictors of earlier and greater accrual of organ damage, and for comparison of cohorts. A SCTC-DI is also useful for the enrichment of clinical trials for patients with less irreversible organ damage, who are more likely to benefit from therapy. It could also be used to measure the burden of disease in epidemiological studies and to assess the interval progression of disease.

1.2 Future Directions

The SCTC Working Groups for Development of Activity and a Damage Index have been set up to develop and validate activity and damage indices for SSc. My systematic review has formed the basis for item generation for inclusion in a survey of SCTC members. The survey responses of SCTC members has been combined with the Australian Scleroderma Cohort Study database to reduce and weight items for inclusion in SSc-DI. External validation of the SSc-DI is currently underway using Canadian Scleroderma Research Group database.
References


Author/s:
Tay, Tien

Title:
Quantifying organ damage in systemic sclerosis: rationale, methodology and item generation

Date:
2018

Persistent Link:
http://hdl.handle.net/11343/210744

File Description:
Quantifying Organ Damage in Systemic Sclerosis: rationale, methodology and item generation

Terms and Conditions:
Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.