Considerations for a combined index for limited cutaneous systemic sclerosis to support drug development and improve outcomes

Alain Lescoat\textsuperscript{1,2}, Susan L. Murphy\textsuperscript{3,4}, David Roofeh\textsuperscript{5}, John D Pauling\textsuperscript{6}, Michael Hughes\textsuperscript{7}, Robert Sandler\textsuperscript{1}, François Zimmermann\textsuperscript{2}, Rachel Wessel\textsuperscript{3}, Whitney Townsend\textsuperscript{8}, Lorinda Chung\textsuperscript{9,10}, Christopher P Denton\textsuperscript{11}, Peter A Merkel\textsuperscript{12}, Virginia Steen\textsuperscript{13}, Yannick Allanore\textsuperscript{14}, Francesco Del Galdo\textsuperscript{15}, Dominique Godard\textsuperscript{16}, Annelise Roennow\textsuperscript{17}, Maya H Buch\textsuperscript{18,19}, Dinesh Khanna\textsuperscript{7}

1-Univ Rennes, CHU Rennes, Inserm, EHESP, Isret (Institut de Recherche en Santé, Environnement et Travail) - UMR_S 1085, Rennes, France.
2-Department of Internal Medicine and Clinical Immunology, Rennes University Hospital, Rennes, France.
3-Department of Physical Medicine and Rehabilitation, University of Michigan, 24 Frank Lloyd Wright Drive, Lobby M, Suite 3100, Ann Arbor, MI, 48105, USA.
4-Research Health Science Specialist, VA Ann Arbor Healthcare System, GRECC, Ann Arbor, USA.
5-Department of Internal Medicine, Division of Rheumatology, Scleroderma Program, University of Michigan, Ann Arbor, Michigan, USA.
6-Royal National Hospital for Rheumatic Diseases (at Royal United Hospitals), Bath, UK.
7-Department of Rheumatology, Royal Hallamshire Hospital, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK.
8-Taubman Health Sciences Library, University of Michigan. Ann Arbor, Michigan, USA
9-Division of Immunology and Rheumatology, Stanford University School of Medicine, Palo Alto, CA, USA.
10-Division of Immunology and Rheumatology, VA Palo Alto Health Care System, Palo Alto, CA, USA.
11-Centre for Rheumatology and Connective Tissue Diseases, Royal Free Hospital Campus, University College London Medical School, London, UK.
12- Division of Rheumatology, Department of Medicine, Division of Clinical Epidemiology, Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania, Philadelphia, PA, USA.
13- Division of Rheumatology, Georgetown University Medical Center, Washington, DC, USA.
14- Rheumatology A department, Cochin Hospital, APHP, Paris Descartes University, Paris, France.
15- Division of Rheumatic and Musculoskeletal Diseases, Leeds Institute of Molecular Medicine, University of Leeds, Leeds LS2 9JT, UK.
16- Association des Sclérodermie de France, Auxerre, France.
17- Federation of European Scleroderma Associations (FESCA), Tournai, Belgium.
18-Centre for Musculoskeletal Research, School of Biological Sciences, Faculty of Biology, Medicine & Health, University of Manchester, Manchester, UK.
19-NIHR Manchester Biomedical Research Centre, Manchester Academic Health Science Centre, Manchester University Foundation Trust, Manchester, UK.

Corresponding author:
Dinesh Khanna, MD, MS
Professor of Medicine
Division of Rheumatology
Department of Internal Medicine
University of Michigan
Ann Arbor, Michigan
Tel.: 734-764-7606
Fax: 734-763-4151
E-mail: khannad@med.umich.edu
Conflict of interest:
DK is a consultant to Acceleron, Abbvie, Actelion, Amgen, Bayer, BMS, Boehringer
Ingelheim, CSL Behring, Corbus,
Galapagos, Genentech/Roche, GSK, Horizon, MitsubishiTanabe Pharma, Sanofi-Aventis, and
United Therapeutics. He has stock options in Eicos Sciences, Inc.
AL has no conflict of interest
JP has received speaker’s honoraria and research grant support (> $10,000) from Actelion
pharmaceuticals. JP has undertaken consultancy work for Actelion pharmaceuticals, Sojournix
Pharma and Boehringer Ingelheim.
CPD reports personal fees from Actelion, Bayer, Boehringer Ingelheim, Corbus, Roche, Sanofi,
CSL Behring, GlaxoSmithKline and Inventiva.

Funding statement:
The project is funded by a grant by SRUK/WSF (UH&UHR1). Dr Khanna was supported by
NIH/NIAMS K24AR063120.

Running head: Combined response index for lcSSc

Word count: 1 Figure; 3090 words
Summary:

Systemic sclerosis (SSc; systemic scleroderma) is a heterogeneous disease characterized by a diverse range of clinical manifestations. Leroy subclassified SSc into limited cutaneous SSc (lcSSc) and diffuse cutaneous subset (dcSSc) based on the extent of skin involvement. As dcSSc is the most severe subset, randomized controlled trials have mainly focused on this subgroup, considering in particular that the presence of extended skin involvement makes its measurement easier, which is critical for evaluating a therapeutic intervention. Nonetheless, the lcSSc subset is also associated with significant morbidity and detrimental impact on health-related quality of life. The lack of interventional studies on lcSSc is partly due to a lack of relevant outcome measures to evaluate this subset. Combining several clinically meaningful outcomes reviewed and selected specifically for lcSSc may therefore improve representativeness and sensitivity to change in randomized controlled trials. A composite index dedicated to lcSSc combining such relevant outcomes would provide the unique opportunity of propelling clinical trials for lcSSc to test and select candidate drugs that could act as disease-modifying treatments for this highly frequent but neglected subset of SSc. This proposed index would include items selected by expert physicians and patients with lcSSc, across domains grounded in the lived experience of lcSSc. This narrative reviews the reasons why lcSSc has been largely overlooked so far, discusses current state of outcome measures, and identifies the challenges and proposes a roadmap for a combined response index dedicated to lcSSc.

Key words: Systemic sclerosis, scleroderma, limited cutaneous systemic sclerosis, classifications, combined response index, composite score, quality of life
Introduction

Systemic sclerosis (SSc; systemic scleroderma) is a heterogeneous disease characterized by a diverse range of clinical manifestations (1). By broadening the previous 1980 ACR classification criteria (2), the 2013 EULAR/ACR classification criteria for SSc ensures the inclusion of patients with earlier and milder disease (3), which constitutes an important step forward for SSc studies. The updated classification criteria of SSc have contributed to foster translational and therapeutic research in SSc (4). Leroy subclassified SSC into limited cutaneous SSc (lcSSc) and diffuse cutaneous subset (dcSSc) based on the extent of skin involvement (5) and later revised this classification (6). Although this dichotomic approach of SSc classification has some limitations (7), it is a clear and simple sub-grouping that has influenced clinical trial design and provides a meaningful clinical prognosis. The lcSSc subset is the most prevalent form of the disease, regardless of the considered geographical regions with a prevalence ranging from 57 to 77% (8–14). In this article, we consider: a) the reasons why lcSSc has been largely overlooked so far, b) discusses current state of outcome measures, and c) identifies the challenges and proposes a road map for a combined response index (CRI).

lcSSc : an overlooked subset

Of all rheumatic diseases, SSc has the highest case-specific mortality (1), which is largely driven by dcSSc. However, lcSSc is also associated with significant morbidity and detrimental impact on health-related quality of life. With regards to visceral involvement and overall disease burden, lcSSc may have been largely overlooked in most studies and may represent “an unfairly neglected subset” (12). A recent analysis from the EUSTAR cohort has proposed a comprehensive view on this issue (15). In this cross-sectional and longitudinal study, more than 8,000 patients with lcSSc were compared with almost 5,000 dcSSc patients. This study highlighted that lcSSc patients experienced multi-system involvement, as suggested by the presence of esophageal symptoms in 62%, 35% had interstitial lung disease (ILD), digital ulcers in 37%, cardiac diastolic dysfunction in 20% and joint synovitis in 13%. Interestingly, 33% of dcSSc were on immunosuppressive therapy 23.7% of lcSSc were on immunosuppressive drugs (15). This cohort highlights that lcSSc patients deserve the same attention for visceral involvement as dcSSc patients, especially the high rate of SSc-associated mortality due to ILD and cardiac involvement (16). The importance of considering lcSSc-associated ILD is also supported by numerous ILD trials that include this subset (17–20). Previous observational studies that have evaluated pulmonary arterial hypertension (PAH) have shown that the lcSSc subset is strikingly overrepresented in comparison with dcSSc patients (21). For example, in the DETECT study, 71% had lcSSc (21). In a meta-analysis specifically addressing the issue of PAH in SSc, more than 80% of PAH patients had lcSSc, with no major differences in survival between the two subsets (22). Quality of life (QoL) in the lcSSc population may also be neglected. In a study addressing QoL impairment in SSc patients, assessed through a pre-defined questionnaire, the main SSc-related manifestations that impacted daily life were Raynaud’s phenomenon, gastro-intestinal and musculoskeletal manifestations (23–28). Nonetheless, the precise definitions and mapping of the domains and related outcome measures may influence these results (15). The overall impact of dcSSc on QoL may be higher than on lcSSc, but the main clinical manifestations responsible for this QoL impairment may be quite similar between the two subsets (29). Considering the increased frequency of lcSSc, much of the SSC-associated morbidity experienced by our cohorts is specifically due to lcSSc. Our systematic emphasis on the diffuse subset within clinical trial programmes of SSc has inadvertently excluded over half of the patients we manage, and in doing so limited the therapeutic options available to this important subset of patients.

Drug development, trials and evaluation studies of QoL, have largely been focused on the dcSSc, and/or have overrepresented the proportion of the dcSSc subset in comparison with lcSSc. There are, on the contrary, very few studies dedicated to the lcSSc subset, especially clinical trials.
A recent analysis of outcomes used in scleroderma trials have highlighted that among the 97 published trials, 53 included both lcSSc and dcSSc patients, 22 dcSSc only, and 4 trials were dedicated to lcSSc (among them only 1 for the entire 2011-2018 period) (30). This could be explained, primarily due to higher impact of dcSSc on mortality, although a more rapid progression of some SSc-associated manifestations in dcSSc also facilitates shorter clinical trials. Moreover, many clinical trials in SSc have been focused on skin involvement, using mRSS as their primary outcome (31–33). It is now well established that mRSS may not be a relevant outcome measure for the assessment of skin involvement in lcSSc, especially considering sensitivity to change in this subset (34). Although digital ulcers and pulmonary related outcome measures may represent shared assessment tools between dcSSc and lcSSc, the lack of interventional studies in lcSSc may be due, in part, to a paucity of relevant outcome measures to effectively evaluate this specific subset. In contrast, clinical trials in dcSSc have recently benefited from the creation and endorsement of a CRI dedicated to dcSSc. This ACR CRISS (Combined Response Index for Systemic Sclerosis) was designed to capture the global improvement of dcSSc based on the selection of domains and items in accordance with the OMERACT (Outcome Measure in Rheumatology) strategy (35). ACR CRISS has been shown to differentiate active therapy vs. placebo in recent trials (31,36) and has been endorsed by the FDA as an acceptable endpoint for registration trials. This CRI is based on a two-step evaluation. The first step evaluates if there has been a new or worsening of the underlying cardiac function (ejection fraction of ≤45% requiring treatment), lung function (loss of FVC% predicted of at least relative 15% in documented ILD, or new onset of PAH), or the occurrence of scleroderma renal crisis during the considered period of time. If such a major event has not occurred, then a second step based on 5 variables [FVC% predicted, mRSS, patient and physician global assessments and disability (health assessment questionnaire disability index)] is used to measure the overall probability of improvement during this same period. Although some items included in this index share relevant outcome measures between dcSSc and lcSSc [FVC%, patient and physician global assessments and disability (health assessment questionnaire disability index), new onset of PAH], other items used in the CRISS are not equally relevant for lcSSc such as the onset of scleroderma renal crisis and mRSS. Most importantly, the CRISS was designed as an assessment tool for dcSSc and was validated using data from randomized clinical trials (RCTs) of patients with early dcSSc and paper patients from longitudinal cohorts of dcSSc (37). Nonetheless, the ACR CRISS and its use in recent RCTs serves as proof of concept that a global assessment of SSc is possible, and potential candidates of disease-modifying drugs can be evaluated using such a tool, even in short term trials with limited sample sizes (36,38). This is especially true considering that in some cases the ACR CRISS can successfully differentiate active therapy vs. placebo, when the primary outcome measure fails to do so, such as in the abatacept trial (31). Thus, an equivalent CRI, that could similarly capture the impact of therapeutic measures on lcSSc may help to act as a timely lever to ensure that this poorly considered subset gets the attention it deserves. This is the overall objective of the C.R.I.S.T.A.L project (Combined Response Index for Scleroderma Trials Assessing LeSSc) as it is dedicated to the creation of a CRI for lcSSc.

**Challenges ahead for a combined index dedicated to lcSSc**

Value-based health care, patient-reported outcomes (PRO) and treatment satisfaction are core values for new reference standards and quality metrics in patient management and drug approval by regulatory agencies (39,40). The creation of a new assessment tool for a complex rheumatological disease, like lcSSc, requires the incorporation of the patients’ perspective, especially for the identification of the most important domains for this CRI. Comprehensive identification of outcome measures, including PRO, will therefore require highlighting the most relevant items within the domains that are considered the most bothersome by patients with lcSSc. The patient perspective on the most bothersome symptoms of lcSSc has been largely overlooked. Including specific PRO directly in candidate combined indices could help to involve the patients in the evaluation process. The majority of PRO development specifically used to assess SSc have not involved the target patients, especially in lcSSc (41). Patient involvement in PRO is now required to satisfy regulatory agencies, such as the FDA, for validity of labelling claims. With
Regarding PRO for lcSSc, the NIH PROMIS® initiative offers a broad range of tools. Some of them, such as the NIH PROMIS® Gastrointestinal Symptoms Scales, are relevant to (but not dedicated to) SSC patients, as SSC patients were involved in their development (42). The NIH PROMIS initiative also provides a large variety of formats, including short forms and computerized adaptive tests (CAT), that allows for customization of the assessment tools that are relevant for the domains of interest, including physical health, mental health and social health domains (43,44).

Concerning this issue of functioning and QoL, an international qualitative analysis of SSC patients’ responses has highlighted that fatigue and pain were among the shared patients’ priorities in all evaluated countries (45). Although this study involved both dcSSc and lcSSc patients, it highlighted that including patients’ perspective and evaluation of QoL would need to include the evaluation of general symptoms, which have been largely overlooked so far (46). This is a challenging issue as nonspecific manifestations, like pain and fatigue, may not be directly impacted by specific therapies for SSC. An effective disease-modifying drug for lcSSc may positively impact these symptoms. Moreover, some specific SSC-associated features, such as physical appearance, change with subsequent impairment of social interactions, and the risk of depression may directly impact functioning and precipitate the development of fatigue (47–49). Similarly, digital ulcers, joint and skin involvement could directly impact pain and the perception of pain. Deciphering the interactions between specific features of the disease and the onset of general symptoms could help to determine the most relevant items to be included in a CRI for lcSSc. Including assessment tools based on modern psychometric and/or item response theories may help to capture important subjective feelings linked to QoL within this new index (43), but we need to keep in mind that the final goal is the creation of an index useful for specific drug evaluation. Achieving the proper balance between the evaluation of lcSSc-related manifestations and the inclusion of considerations on functioning and QoL based on the patients’ perspective is one of the main challenges ahead.

Including input from experts in SSC trials and the careful evaluation of the candidate items for final selection according to the OMERACT filter shall also be vital. The sensitivity to change in lcSSc may constitute specific challenges. For example, the rate of progression in lcSSc is generally slower than in dcSSc, and manifestations such as digital ulcers and PAH may occur much later in the natural history of lcSSc (15). The inclusion of items based on the time to treatment failure within candidate indices may help to tackle this issue, and the combination of multiple items in a same index may also help to increase the power and sensitivity to change.

Another issue is the question of defining the overall treatment goal: improvement or stabilization? There is still an ongoing debate concerning this question in SSC in general, and this decision will greatly impact the selection of the items and domains for lcSSc. This question also highlights that this combined index will not be an activity or severity index, and it will need to be designed with the constant concern of its relevance in clinical trials (50). The creation of different candidate indices, based on relevant items using the OMERACT filter and exploring domains identified as “bothersome” by lcSSc patients will be a necessary step. The comparison of their relevance will need to be tested against existing cohorts, paper patient evaluations by experts, and may require the creation of a dedicated cohort to validate the most promising indices.

A proposed roadmap for the development of the index and limits of such an approach

In Figure 1 we have presented a possible roadmap to guide the development of a CRI for lcSSc. The first stage of the project would be the identification of key domains and related outcome measures to inform these domains. As the cornerstone of this project is the inclusion of the patients’ perspective, the first step will involve a qualitative approach based on e-focus groups, including only patients with lcSSc, to highlight the key domains they considered as the most bothersome. This e-focus group approach would allow identification of items and domains without a priori, and without preconceived or pre-determined clinician-oriented questionnaires. This would ensure that the clinicians’ perspective would have limited impact on early data collection. The identified domains will be informed by a systematic scoping review of the literature that will provide a comprehensive overview of outcome measures used so far to assess lcSSc patients in
observational and interventional studies (30). Analyzing the frequency of use of the outcome measures and the domains identified during the literature review, may allow us to identify the gap between researcher’s/clinician’s concerns and lcSSc patients’ perspectives as identified in the e-focus groups. The next steps for the identification of the core set of items would be the conduction of Delphi exercises for experts and patients to enrich the list of items previously identified allowing for the ranking of the most relevant items from the patients’ and experts’ perspectives. Based on the results of these Delphi exercises the selection of the core set could be proposed though nominal group technique (NGT) exercises, involving patients and experts, with the goal of achieving a consensus for a short list of items that would include the most relevant outcomes for RCTs according to experts (based on the OMERACT filter) and items and/or domains identified as bothersome by the patient participants.

Having defined this core set of items, the psychometric properties of all items should be tested in longitudinal cohorts to appreciate their feasibility, reliability and validity (including responsiveness to change). These longitudinal results will serve as the next steps by providing data for patient profiles and will help to finalize a revised core set of items. Using a dedicated cohort could help to include the self-reported status of the patients, as improved, stabilized or worsened. A similar rating by expert clinicians would also be proposed based on patient profiles. Next, the final selection of the items to be included in the candidate indices would be determined by evaluating their association with the identified goal (improvement/stabilization), testing their redundancy, and determining the helpfulness of the items for predicting evolution. Based on these results, the candidate indices with the most relevant associations, as defined by a steering committee including expert clinicians and patient partners, with different weights for each item would be proposed. The last step will be the inclusion of these candidate indices as secondary endpoints in clinical trials to select the most efficient index for differentiating groups (37).

One could argue that individual organs are more often affected in lcSSc requiring a subsequent need for specific assessment tools dedicated for the involved organs and this should prevail on the creation of a CRI. The development of new PRO and/or outcome measures for specific domains is not in contradiction with the creation of a CRI. These initiatives would be complementary as more tools adapted to assess specific domains in lcSSc could be included in a combined index, which could benefit from better outcome measures for the identified domains (50,51). LeSSc has a more prolonged disease course than deSSc, but this very issue, in combination with the statement that organs are more often affected individually in leSSc, supports the necessity of a combined index to more efficiently capture changes is disease status. It has been recently highlighted that a promising pathway for scleroderma trials would be the increasingly frequent evaluation of combination therapies (52), as the discovery of a single disease-modifying drug is uncertain. Combination therapies could be a viable strategy that may help to manage various SSC-associated domains, and with this in mind, a CRI could constitute a relevant endpoint. The use of a CRI could lead to increased standard deviations with subsequent increases in the required sample size (52). The issue of overtly restrictive inclusion criteria is a major concern for RCTs evaluating early deSSc, but it would likely be a less critical barrier in lcSSc since limiting inclusion based on a maximum disease duration may be of less importance for leSSc than deSSc. Another concern about CRI is the clinical relevance of the differences between groups. This is also true for single item primary outcomes as illustrated with the recent debate concerning the clinically-relevant decline of FVC in SSc-associated ILD (18,53). The involvement of patients at each stage of the collaboration process of a lcSSc focused CRI would strengthen the clinical relevance of the index and ensure that this index would be adequately grounded in and responsive to the lived experience of lcSSc.

Conclusion

The traditional dichotomy of SSc, separating leSSc and deSSc, has recently been challenged as authors have highlighted there are more than two subgroups at stake and a more nuanced classification with various subsets may more accurately reflect the heterogeneity of the disease (54). Interestingly, in previous attempts to develop new classification strategies, a frequent subset
characterized by modest cutaneous evolution, a high prevalence of gastro-intestinal manifestations, low mortality rate, and a high proportion of anti-centromere antibodies positive patients, has typically emerged (55–57). These new classification approaches highlight the limitations of the binary approach to disease sub-setting currently deployed (7), although a major subset of patients still responds to the “classical” image of Leroy’s lcSSc. This subset is still comparatively neglected, specifically in terms of their contribution to clinical trial programmes and availability of targeted therapeutic strategies. A project of a CRI dedicated to lcSSc, that would properly capture relevant key domains, based on the patients’ perspectives, and would include patient partners at each step of its conception in collaboration with expert clinicians, may help to drive more attention to lcSSc. Identifying and defining the domains and relevant outcome measures to be included in such a CRI is a necessary first step for the development of this index. Selecting uniform, patient-informed, and clinically meaningful outcome measures could lead to the design of clinical trials with strong potential to achieve regulatory agency approval and propel drug development in lcSSc.

References


