Published in final edited form as:

J Scleroderma Relat Disord. 2020 March ; 5(2 Suppl): 48–60. doi:10.1177/2397198320904178.

Patient-reported outcome measures in systemic sclerosis–related interstitial lung disease for clinical practice and clinical trials

Lesley Ann Saketkoo[1,2,3,4], **Mary Beth Scholand**[5], **Matthew R. Lammi**[1,2,3], **Anne-Marie Russell**[6,7]

[1]New Orleans Scleroderma and Sarcoidosis Patient Care and Research Center, New Orleans, LA, USA

[2]Interstitial Lung Disease Clinic Programs, Comprehensive Pulmonary Hypertension Center, University Medical Center, New Orleans, LA, USA

[3]Division of Pulmonary Diseases, School of Medicine, Louisiana State University, New Orleans, LA, USA

[4]School of Medicine, Tulane University, New Orleans, LA, USA

[5]Division of Pulmonary Medicine, University of Utah, Salt Lake City, UT, USA

[6]National Heart and Lung Institute, Imperial College London, London, UK

[7]Imperial College Healthcare NHS Trust, London, UK

Abstract

Systemic sclerosis (SSc) is a progressive vasculopathic, fibrosing autoimmune condition, portending significant mortality; wherein interstitial lung disease (ILD) is the leading cause of death. Although lacking a definitive cure, therapeutics for (SSc-ILD) that stave progression exist with further promising primary and adjuvant compounds in development, as well as interventions to reduce symptom burden and increase quality of life. To date, there has been a significant but varied history related to systemic sclerosis–related interstitial lung disease trial design and endpoint designation. This is especially true of endpoints measuring patient-reported perceptions of efficacy and tolerability. This article describes the underpinnings and complexity of the science, methodology, and current state of patient-reported outcome measures used in (SSc-ILD) systemic sclerosis–related interstitial lung disease in clinical practice and trials.

Corresponding author: Lesley Ann Saketkoo, New Orleans Scleroderma and Sarcoidosis Patient Care and Research Center, Division of Pulmonary Diseases, Louisiana State University and Tulane Schools of Medicine, Interstitial Lung Disease Clinic Programs, Comprehensive Pulmonary Hypertension Center, University Medical Center, 2000 Canal Street, New Orleans, LA 70112, USA. lsaketk@tulane.edu.

Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship and/or publication of this article: The content is based solely on the unbiased research and expertise of the authors, they have received no financial incentives from any party to produce this publication. A.M.R., M.B.S. and M.R.L. report no conflicts of interest. A.M.R. and L.A.S. have provided consultation, research and education for Boehringer Ingelheim pharmaceuticals who are the sponsors for the edition of the *Journal of Scleroderma and Related Disorders*.

Keywords

Systemic sclerosis; scleroderma; pulmonary fibrosis; interstitial lung disease; health-related quality of life; dyspnea; cough; patient-reported; outcomes

Introduction

Systemic sclerosis (SSc) is a complex multi-organ disease of vascular injury, endothelial dysfunction, destruction and disrepair with ensuing inflammation, immune activation, and dysregulated fibroblastic proliferation with extracellular matrix deposition resulting in clinical fibrosis. SSc is highly heterogeneous in degree and type of organ involvement, clinical manifestations, severity, rate of progression, and survival. Interstitial lung disease (ILD), followed by pulmonary hypertension (PH), is the leading cause of SSc-related death.[1,2]

SSc lacks a definitive cure; however, treatment development in SSc, and systemic sclerosis–related interstitial lung disease (SSc-ILD) especially, is rapidly advancing. Increasing numbers of SSc-ILD studies annually yield meaningful information on trial design and endpoints for how best to measure therapeutic responsiveness. Beyond traditional markers such as computed tomography (CT) scanning and pulmonary function testing (PFTs), growing recognition and importance is placed on the group of outcome measures that register patients' perceptions of therapeutic responsiveness and medication tolerability: patient-reported outcome measures (PROMs).

PROMs are intended to reflect the positive and negative impact of changes in a health condition as experienced by patients in the context of their everyday lives as they try to sustain activities of work, social, and family life. This includes treatment side effects and disease-related symptoms such as dyspnea, cough, fatigue, and exercise tolerance as well as impact on factors such as health-related quality of life (HRQoL), physical function, and psychological distress. Patient experience can be described under two overarching and overlapping concepts: *symptomatology* and *HRQoL*. This article endeavors to impart the utility, science, methodologic assurances, challenges, and current state of PROMs in SSc-ILD for use in clinical trials and practice.

Symptomatology of interest to measure in SSc-ILD

Symptoms of a physical health condition emerge as structural and/or physiological stresses are sufficiently extensive to disrupt a patient's ease of engagement with life activities. The behavioral qualities of symptom-related physical impairment are as follows:

1. *Presence* which can be fleeting and completely resolve—as with pain resulting from transient headache or a minor cut when using a kitchen knife; or can be constant and/or episodic as with chronic or flaring diseases;
2. *Intensity* of sensation or impairment which may increase, stabilize, fluctuate, or remit;

3. *Trajectory* whereby symptoms/impairment may either intensify over time perhaps driven by reversible disease activity escalating or irreversible progression of damage, remit with disease reversibility through auto-remission or treatment, or stabilize with cessation of progression with residual permanent damage.

In SSc, inflammation drives acute, reversible disease activity that if not quelled results in evolving and then end-stage irreversible fibrosis. Quieting early inflammatory disease may reverse symptomatology, while later disease stages with irreversible accrual of damage results in more challenging symptom management (Figure 1). In the wake of irreversible biophysical damage, improvement of symptom burden may still be possible through non-disease modifying symptom management strategies.[3–12]

Based on our prior work, primary lung-related biophysical symptoms of interest to patients with SSc and connective tissue disease–related interstitial lung disease (CTD-ILDs) are *cough, dyspnea, exercise tolerance*, and *fatigue*. In these studies, patients voiced an inability to discern whether fatigue was predominantly influenced by underlying ILD or the primary CTD itself. Patients communicated these symptom-related impairments via contextualization of important life aspects rather than discrete terms such as “breathless.” This highlights how the accuracy of symptom-related reporting may be influenced by attention to language as well as physical and psychological context.[13–17]

The multi-organ system involvement of SSc, as well as medication side effects, creates diverse and overlapping coincident manifestations and symptoms in a person living with SSc. These complicate perceptions of anticipated pulmonary symptoms. For example, dyspnea, exercise intolerance, and fatigue in a person with SSc-ILD can arise from effects of ILD; but also other experiences common to SSc, such as anemia, cardiopulmonary microvascular disease and pulmonary hypertension, infection, myopathy, physical deconditioning, multifactorial pain, reflux, systemic inflammation, and common general comorbidities such as coronary artery disease, insomnia, obstructive lung disease, and sleep apnea.

Symptoms may be imperceptible to the clinician and also unrecognized by the patient consistent with the human-drive for adaptability. This adaptability underscores the need for careful clinician questioning for patients at risk. Adaptability of *pace, intensity*, or *corporeal positioning* of activities accommodate for changing physiological circumstance. Patients with restrictive lung diseases may, respectively, slow pace or take breaks, add less groceries (weight) to shopping bags, change routes to avoid inclines, or forgo optional activities; and/or propping chest upright or reducing depth of inspiration to avoid aggravating symptoms. Beyond somatic stress, symptoms are multi-layered experiences influenced by prognostic uncertainty, cultural overlays and inferences, psychological distress, clinician communication, and a patients’ potential for self-efficacy; all of which may contribute to patient report.[18–20] However, despite the complexity of symptoms in SSc, patient-reported “yes/no” on symptom worsening correlates clinically meaningful changes in disease activity in SSc patients.[21]

Cough has high prevalence in SSc-ILD. Patients report cough being a central ILD experience reflecting disease worsening [ref], with adverse impact on HRQoL and disease progression in pulmonary fibrosis correlating with radiological extent of fibrosis, decreased diffusion capacity, vital capacity, and more dyspnea.[14,15,17,22–24] Cough adversely impacts HRQoL[22–24] whereby physical symptoms like chest pain, sleep disturbance, emesis, and incontinence potentially develop[25] with complex associations of embarrassment, anxiety, depression, and severe disruption of life activities.[26] Cough frequency in SSc-ILD decreases significantly with immunosuppressive treatment[22,23,27] and returns to baseline 1 year after withdrawal of treatment.[23,27–29] Suggesting that cough may be an independent variable in assessing both ongoing fibrosis and therapeutic response in SSc-ILD.[23] However, cough frequency in SSc-ILD associated with Gastroesophageal reflux disease (GERD) is also common and correlates with the presence, severity, and treatment of GERD in SSc-ILD.

HRQoL

Over past decades, growing availability of chemotherapy catalyzed contrasting discussions in clinician and patient communities regarding experiences of tolerability/toxicity balanced against days of life gained in oncological disease. The medical community's perspectives on survival evolved accounting for HRQoL. "*Life at all costs*" slowly made way for "*informed decision-making*" whereby patients and families were counseled on side effects and projected length of post-treatment survival; more recently, "*goal-oriented care*" incorporates patients' priorities of daily living in treatment decisions. This conceptual trajectory increasingly illuminated survival, from regulatory and practice standpoints, as an important but not an exclusive indication of successful treatment. Recently, as an example, qualitative studies in idiopathic pulmonary fibrosis (IPF), a fatal progressive disease, highlighted patient concerns that disease-slowing medication side effects may severely limit meaningful life interactions and need to be balanced as a trade off against survival.[8,30,31]

Broad concepts of preventing/reducing disability and augmenting function in the context of survival[32,33] under-pin the utility of PROMs. Traditional disease measures characterize fluctuations of biophysiological processes; from which clinicians can only imagine how poorly or well a patient is able to live their lives. HRQoL PROMs, if well-developed, provide scientific quantification of the impact of a health condition, reliably capturing health condition-related changes in "*the stuff that life is made of*" including symptoms, function, mental health, energy, workability, and coping, in other words the quality of one's life lived. HRQoL PROMs may have discrete components for physical function, energy/fatigue, psychological distress (anxiety, depression), as well as biophysical symptomatology.

Physical function relates directly to physical limitations resulting from a health condition. It is often captured in HRQoL but can be a discrete measure, such as the Health Assessment Question-Disability Index (HAQ-DI), used in addition to a generic (see below) global HRQoL measure. The HAQ-DI is the basis for the SHAQ (Scleroderma Health Assessment Questionnaire)[34,35] that includes six visual analogue scale (VAS) items assessing the perceived severity of pain, Raynaud's phenomenon, digital ulcers, breathing, gastrointestinal, and overall health related to SSc.

Importantly for clinical trials, interventions may improve HRQoL and physical function but not coincide with changes in disease activity. This can occur with palliating medication side effects, treating depression/anxiety, improved access to care or clinician communication, providing education or assistive or mobility devices, and work accommodations.[3,4,11,12,32]

Outcome measures

We expect outcome measures to correlate with disease activity and trajectory. Clinical trial design in SSc-ILD accommodate for therapies that perform at different levels of efficacy: to improve lung function,[28,29,36–39] or prevent or delay worsening of progressive functional loss.[40] This helps differentiate between potential powerful first-line and adjuvant supportive therapies in this enigmatic disease.

In regard to terminology, "*PROM*" is the scientifically appropriate abbreviation. If "*PRO*" is used, a qualifying noun, such as "*measure*," "*instrument*," "*questionnaire*," and "*endpoint*" is required.[41]

Historical perceptions of objectivity

Over the centuries, Western medical science preferred credence in traditional "objective" measures. However, upon reflection, few measures are truly objective, that is, resulted without fallible inference. For instance, physical examination components or clinical investigations, such as imaging, remain largely unquestioned as "objective" but in fact these rely upon human observation, examination and, then, interpretation; each phase inherently subjectively influenced by psychological, educational, environmental, and belief tendencies. Despite this, the label "objective," suggestive of higher scientific quality and value, continues to apply to clinician-interpreted outcome measures. However, given adequate opportunity, patient-reported history provides the greatest proportion of diagnostic information, honed then by physical examination findings, with diagnostic testing for confirmation.[42] Henceforth, we refrain from using "objective" or "subjective" rather referring to measures as "diagnostic" or "traditional" and "patient-reported."

Outcome measure qualities

Whether an instrument be patient-reported or traditional, in order to demonstrate its value as a measure, performance testing under multiple criteria increases the quality of an outcome measure. The more validation strategies and observations an instrument undergoes leads to increasing confidence in that measure. Albeit, there are many instruments that are under-validated in use both in practice and research, but fulfill a yet unmet need by a more substantial measure. One such example is *forced vital capacity* (*FVC*). Despite being a flawed measure vulnerable to influences by comorbidities, administration and equipment variance, and dependence on factors influencing patient performance, it currently is the best traditional measure to mark the trajectory of restrictive lung disease.[15]

The *validity* of a measure is determined by its *content* (relevance, comprehensiveness, purity of intent; for example, red = actual red, all shades of red, all things that are red), *construct* (conceptually sound in performance, for example, red = red if red exists, not yellow = red), and *criterion-related* (correlates with an external instrument measuring the same construct,

for example, identifies between red and not-red in accord with external instrument). "*Unified construct validity*" theory accounts for each of these validity concepts as different facets of a single validation effort.

Beyond validation strategies, essential performance observations establish a measure's confidence in utility. These important benchmarks to demonstrate are as follows:

- *Reliability*, regardless of scale the measure will perform consistently every time (even if measuring the wrong construct, it will do so every time).
- *Discrimination*, ability to distinguish from a similar but not accurate other construct.
- *Responsiveness*, ability to register change in the construct over time with disease worsening or improvement. Not necessary for tools with a single time-point assessments such as for screening, staging, or diagnosis.
- *Practicability*, speaks to issues of ease of use, access, and safety:
  - *Interpretability*, the degree to which instrument's results are easily interpreted.
  - *Access*, the extent to which the instrument is commonly available.
  - *Cost*, financial practicalities of implementation.
  - *Time-intensiveness*, sometimes related to cost of administration, interpretation; but also the burden of time on patients. Sometimes related to fatigue for technician or patient in active performance.
  - *Safety*, the degree of risk for technician or patient.

These concepts are foundational in discerning what constitutes an ideal measure, and what could improve an adequate measure. Furthermore, PROMs may have an established minimal clinically important difference (MCID) in clinical use; however, study results may demonstrate statistically significant change that does not correlate to MCID benchmark of that PROM and vice versa.[43–45] Often, this disparity is not considered as poor PROM performance but rather reported factually in the study results leaving readers or regulatory agencies to interpret.

A PROM may be either *generic* (used across health conditions) or may be disease- or symptom-*specific*. An example of such is the *Short Form-36* for generic HRQoL and *St George's Respiratory Questionnaire* as respiratory-specific HRQoL. There are benefits to each type; generic measures are often performance validated across many diseases, and generic measures also confer generalizability for disease comparisons. Generic measures can act as benchmark measures for construct validity of under-validated measures. Although anticipated to correlate with generic PROMs, disease-specific PROMS provide distinctive disease-related information hopefully conferring greater reliability, discrimination and sensitivity to change.

Path to PROM development

Food and Drug Administration (FDA) and European Medicines Agency (EMA) mandate that patient-reported data collection and PROM development be implemented with sound scientific methodology essentially informed by the experience and priorities directly from patients through *qualitative research*.[41,46] Qualitative research for scientific and health intervention purposes, whether for trial end-points, educational tools, or other purposes, represents the patient's voice; ensuring accuracy relies on exacting and exhaustive methods, with careful attention to context and language.

Although not a precise reproducible science, qualitative research is qualified by standards reflecting validity (credibility), objectivity (confirmability), and generalizability (transferability). FDA/EMA patient-centered regulations established new demands on therapeutic research. However, distinguishing quick, dilute methods sufficient for marketing research purposes from stringent methodology driving scientific research for patient well-being and therapeutic advancement is essential. The most accepted quality assessment for journal publication, the COREQ (consolidated criteria for reporting qualitative research), [47,48] is a 32-item checklist ensuring the rigor of the qualitative research. The COREQ examines qualitative rigor with several questions under each of three domains: *study team, study design*, and *analysis and findings*.

We present seven overarching steps required to occur in robust PROM development (Table 1):

- *Team*. Convening a qualified team which should include patient research partners, clinician expertise, qualitative researchers, and psychometricians.
- *Question*. Clarifying the need and goals of PROM research, for example, *screening versus monitoring, generic versus disease/symptom specific*, and *physical versus inclusion of psychosocial aspects*, which may include literature review and preliminary qualitative investigations.
- *Design*. Designing the research to ensure saturation of relevant concepts resulting in a conceptual framework that supports a wide and accurate item collection while assuring the required level of diversity within the disease of interest.
- *Implementation and qualitative analysis*. Focus groups with interval iterative analyses for assessment of concept saturation, framework revision, and distillation of discrete concepts.
- *Field-testing*. Field-testing of concepts and initial questionnaire items, including field-testing items for relevance of concept, language, phrasing, and preferred response format, can be accomplished with qualitative investigations or in combination with quantitative surveying for larger response.
- *Item reduction and fit*. Quantitative field-testing may support item reduction but psychometric testing is essential using one or more approaches from Rasch analysis, factor analysis, and/or item response theory to reduce items and test item fit.[49–51]

- *Validation*. Validation is a multi-step process, initial testing with test–retest validation, with unified validation taking place in the context generally of registries or as an exploratory endpoint in clinical trials.

Developmental deviations

Deviations from FDA/EMA PROM guidance exist, the most common being utilization of *older, pre-guidance, performance-validated PROMs* developed without patient-centered qualitative inquiry but rather via clinician conjecture and observation, e.g. HAQ-DI, SF-36.[52] *Borrowed PROMs* developed for other diseases like chronic obstructive pulmonary disease (COPD), for example, SGRQ (St. George's Respiratory Questionnaire), Borg, or MTDI (Mahler Transitional Dyspnea Index), when used in SSc-ILD, can provide a needed measure when a disease-specific measure is lacking.[28,29,40] *Adapted PROMs*, using qualitative field-testing to amend PROMs to fit a new disease, such as ILD versions of SGRQ and Dyspnea-12 or to establish compliance with FDA/EMA guidance,[53,54] eliminates foundational steps and may not be sufficient to capture patient priorities and relevance to the particular disease experience. While PROMs with abbreviated development have utility in filling a care and research gap, optimal endpoint performance with robust methodology for patient content[16,55] remain an important concern especially for diseases like SSc-ILD that present assessment challenges.

PROMs of interest in SSc-ILD

Measuring patient-reported changes in SSc-ILD and other CTD-ILDs where there may be multiple organ involvement and where the further burden of medication side effects is complex. This section describes symptom and HRQoL PROMS used in SSc-ILD trials (Tables 2–4).

Dyspnea.

Respiratory-specific HRQoL instruments, such as the SGRQ, reflect the burden of respiratory symptoms, whereas dyspnea-specific instruments have demonstrated utility in assessing symptom response to treatment.[65,66]

SHAQ Breathing VAS is the simplest, most direct, and robustly validated dyspnea measure in SSc-ILD. It is a single-item scale and correlates with FVC, high-resolution computed tomography (HRCT), and HRQoL measures.[28,29] The Borg scale, another single-item 10-point scale validated across several pulmonary conditions,[67,68] assesses pre/post exercise dyspnea, for example, 6-min walk test (6MWT).[69] In SSc-ILD, rest/post-6MWT Borg scores demonstrated reproducibility and correlated to distance.[70,71]

The Medical Research Council Breathless Scale (MRC) queries dyspnea on five scaled statements of dyspnea in relation to life activities, akin to the NYHA Classification, has strong utility in stratifying severity but with threshold limits in serial change. The MRC, validated in IPF but not SSc-ILD,[72] demonstrated responsiveness, reproducibility, and construct validity[73,74] and independently predicted anxiety and depression in ILD.[75]

Mahler indexes, developed for COPD,[76] consist of three abstract questions staff-administered examining changes in “functional impairment,” “magnitude of task,” and “magnitude of effort.” The implementation and results are administrator-dependent, but they demonstrated partial validation with ability to predict treatment responsiveness in SSc-ILD and distinguish between treatment-related improved FVC versus placebo-related worsening.[28,77]

The UCSD-SOBQ, a quick 24-item instrument with short question stems of 1–3 words and a 6-point response scale, has demonstrated validity in ILDs including SSc-ILD[78,79] and IPF.[65] Its psychometric soundness presumes the ease of comprehension and the contextualization demonstrated to be important in patient communication of dyspnea[14–17]; however, it is unknown if contextualization of physical activities in the UCSD-SOBQ is confounded by systemic and musculoskeletal symptoms. Contextualization, though useful, can be highly personal. The UCSD-SOBQ has only two items querying psychological aspects with both relating to fear.

The Dyspnea-12, 12 simple statements of which six focus on the breathing act and six on affective aspects of breathing using a 4-point scale, demonstrated reliability in ILDs including SSc-ILD[53,74,80] and correlates with 6MWD.[74] Dyspnea-12 targets breath perception and avoids contextualizing activities which may be a strategy to help the respondent separate breathing from systemic symptoms, for this reason D-12 was accepted as a potential measure for future studies.[15]

The Functional Assessment of Chronic Illness Therapy–Dyspnea (FACIT-D), a 10-item activity-based questionnaire very similar to the HAQ-DI with responses focused on breathing in relation to daily self-care, therefore contextualizing with essential daily activities was validated in an SSc-ILD cohort.[81]

Cough.

Several instruments measure cough severity, frequency, intensity, and/or impact on HRQoL.[82] The Leicester Cough Questionnaire (LCQ), though not developed for restrictive lung disease, has been the most commonly used cough instrument in SSc-ILD and other ILDs as both a measure of symptom severity and cough-related HRQoL.[15,29,83] The Cough Quality-of-Life Questionnaire (CQLQ)[83,84] is a cough-specific symptom and HRQoL instrument with good internal reliability and total score responsiveness in IPF, but performed less sensitively to extreme physical and psychological symptoms[26,85–87] with ceiling effect possibly related to a static 4-point response scale.

The modified cough index (MCI)[28,29,85–89] assessed frequency, severity, and phlegm production on a 3-point scale, despite SSc-ILD’s association with predominantly dry, inspiratory cough, MCI demonstrated responsiveness. VAS are a simpler serial assessment of cough severity and or frequency, and easy to include as study construct measure. Neither the LCQ nor CQLQ contain direct questioning on cough severity or frequency as do the VASs.

HRQoL.

In SSc-ILD, beyond being a composite of disease-related impact, HRQoL often reflects trade-offs between treatment efficacy and side effects that also introduce physical or mental fatigue, gastrointestinal symptoms, or recurrent infection. The SF-36,[52] Patient Global Assessment (PtGA), SHAQ,[35] and SGRQ[54,90] are used most extensively across ILD studies.

The SF-36, a generic measure, can be sub-divided into independently scored domains addressing mental health, physical health and function, fatigue, and pain, allowing for following these domains and also analyzing the largest contributor to the total SF-36 score. In the SLS I trial, *MTDI* and *VAS Breathing* correlated highly with each other; however, SF-36 was able to differentiate patients with worse dyspnea, FVC, and diffusion capacity of the lung for carbon monoxide (DLCO),[27] emphasizing the utility of SF-36 in SSc-ILD. PtGA VAS is a generic HRQoL measure whose wording can be altered to specify a health condition, is validated across rheumatic and non-rheumatic diseases, correlates with dyspnea, and recommended for use in CTD-ILD trials.[15,35,80]

The HAQ-DI is widely validated across many health conditions and addresses a component of HRQoL, physical function, which is anticipated to improve with treatment efficacy. The SHAQ is the HAQ-DI with five VAS items specific to scleroderma manifestations (Raynaud's phenomenon, digital ulcers, gastrointestinal, breathing, pain), the sixth VAS being a PtGA.

At present, no concise ILD or SSc-ILD-specific HRQoL PROM assesses the experiential benefits or detriments of treatments. This lack has led to uncertainty and difficulties in confidently interpreting PROM results.[40] The SGRQ demonstrated acceptable validity and reliability in ILD, responsive in CTD-ILD[31,91,92] and may be an independent predictor of disease progression in ILD.[31]

Fatigue.

Fatigue, a multi-dimensional phenomenon, struggles to be understood in the context of disease both phenomenologically as a living experience and also as discrete measurable components for clinical trials and practice.[93] Rheumatoid arthritis was the first autoimmune disease establishing patient-reported fatigue as a reliable, reproducible outcome that is sensitive to change over time.[94] Multi-factorial etiology may contribute to distinct fatigue types (mental/cognitive, motivational, physical, etc.); in autoimmune diseases, overlapping etiologies complicate CTDs with CTD-ILD patients reporting inability to distinguish between CTD versus ILD fatigue.[15,17] Beyond the mental and physical dysregulation that inflammatory milieu confer on hypothalamus, brain, muscle, bone and constitution, resultant pain, sedentary susceptibility, and logistical burden confer yet other discrete fatigue sources. The presence of ILD confers yet different potential sources of fatigue related to the systemic sub-acute impact of lung impairment on muscle, brain, and other tissues.[93] Despite this complexity, fatigue remains an indiscriminant phenomenon reported as a general measure sometimes under the reverse phenomenon of *energy* or *vitality*. To date, it is unknown which fatigue scale is optimal in SSc-ILD; the FACIT-Fatigue and the SF-36s Vitality profile have

both demonstrated feasibility and correlate with changes in FVC; however, this may be true of simpler scales such as fatigue/energy VAS.

Coping/self-efficacy.

A number of scales[95–97] have been developed to control for improvement in patient management and coping of disease that may impact other patient-reported variables such as HRQoL, physical function, and fatigue despite lack of improvement in underlying biological disease processes. For clinical practice, we discuss this concept further under “Future directions.”

Future directions

Improved dyspnea and cough measures

None of the above instruments were developed specifically for ILD or developed by current accepted methodology; while current instruments may suffice, an argument can be made for more reliable and sensitive patient-reported data collection. Preliminary guidelines for future CTD-ILD studies supported development of new questionnaires.[15] Although highly individualized, contextualization is an important aspect of dyspnea communication. Many examples persist, within the question concept of lawn-mowing: some patients may have no lawn, and some patients simply stop mowing making that contextualization defunct. Whereas some personal and high-priority contexts are not captured in current PROMs, for example, “reading a bedtime story at night” or “giving the children a bath.”[14,17] Novel PROMs where patients set their own context for priority items warrant investigation. Furthermore, simple response anchors reflecting “decline-no change-improvement” may better capture serial change and lessen recall bias.[21,98]

Patients perceive that cough is central to ILD behavior and supported by traditional measures. The cough in ILD is a different experience than that of COPD;[13,14,17,22] furthermore, SSc-ILD cough requires discrimination between other entities such as reflux. Although the LCQ demonstrated sufficient performance in SSc-ILD, measures from the patient-centered disease experience may enhance accuracy therapeutic trials and clinical practice.

Screening tools for ILD

Given that the largest volume of lung is lost in the first 2 years of SSc-ILD,[99] confident screening mechanisms are essential for early detection and appropriate treatment of this high risk, deadly manifestation. Beyond traditional screening procedures, a careful inquiry approach can reveal changes in patterns of breathing, exercise tolerance, and inspiratory dry cough. Current respiratory system PROMs are not developed for early detection but rather monitoring an existing recognised symptom. An SSc-ILD screening tool comprised content, language, and contextual characteristics bearing sufficient sensitivity to disclose adaptation to symptoms, but specificity to discern between ILD and other causes of similar symptoms could provide great utility in clinical practice.

Concurrent response data collection

An interesting question is the tandem data collection on sentinel *extra*-pulmonary manifestations in multi-organ diseases such as SSc-ILD as a service to treatment advancement in SSc. For example, digital ulcer, musculo-skeletal, or gastrointestinal response to treatment intended for SSc-ILD. There are no guidelines or recommendations for such, but there is potential value of an SSc-ILD trial to provide significant insight into other systemic manifestations in this enigmatic disease and future trials should consider this design augmentation.

Patient-reported experience measures (PREMs)

Patient-reported experience measures (PREMs) complement PROMs and are used to gather information from patients on their experiences of health services and/or systems in order to improve those services for a specified health condition. A PREM in rheumatoid arthritis already exists and a PREM for IPF is in development. PREMs create roadmaps that guide development of clinical services and test innovations in eight validated domains: *Respect for patient-centered values, preferences, and expressed needs; Co-ordination and integration of care; Information, communication, and education; Physical comfort; Emotio nal support; Involvement of family and friends; Transition and continuity; and Access to care*.[100] Applied results build on detected system strengths while addressing weaknesses with innovations often developed with qualitative methods, and changes then retested with the PREM.

Patient activation measures (PAMs)

Patient activation measures (PAMs) assess patient engagement and knowledge of their health condition. The PAM originated as a 22-item measure and subsequently modified to a short form 13-item instrument,[101,102] although there are several other similar well-validated instruments.[95–97,103] Although it can be used for group-based research or clinical trials to control for the influence of increasing self-efficacy, the primary use of PAMs is to detect and address areas of patient self-management strengths and weaknesses with the goal of improved health outcomes. PAMs in ILD are currently being examined by the authors.

Conclusion

PROMs are an essential reflection of a patient's experience of disease for clinical practice and clinical trials. Accurate assessment of patient experiences of SSc-ILD is complicated by overlapping manifestations creating similar symptoms, treatment side effects, and disease-related psychosocial burden. The value of treatment is balanced between survival and long-term tolerability. Development of PROMs and other patient-centered material require care and rigor to accurately reflect patient voice and priorities, all undertakings should be reported with scientific quality benchmarks. Future work is warranted in PROM development for ILD screening, disease activity, symptomatology, healthcare service experience as well as patient engagement.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This work was supported in part by 1 U54 GM104940 from the National Institute of General Medical

Sciences of the National Institutes of Health (NIH) in the United States, which funds the Louisiana Clinical and Translational Science Center (M.R.L.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This work was also supported in part by COBRE Pilot Program grant number 1P306M106392-01A1 (M.R.L.) in the United States and supported in part by National Institutes of Health Research in the United Kingdom.

References

1. Steen VD and Medsger TA. Changes in causes of death in systemic sclerosis, 1972–2002. Ann Rheum Dis 2007; 66(7): 940–944. [PubMed: 17329309]
2. Tyndall AJ, Bannert B, Vonk M, et al. Causes and risk factors for death in systemic sclerosis: a study from the EULAR Scleroderma Trials and Research (EUSTAR) database. Ann Rheum Dis 2010; 69(10): 1809–1815. [PubMed: 20551155]
3. Bakitas M, Lyons KD, Hegel MT, et al. Effects of a palliative care intervention on clinical outcomes in patients with advanced cancer: the Project ENABLE II randomized controlled trial. JAMA 2009; 302(7): 741–749. [PubMed: 19690306]
4. Giese-Davis J, Collie K, Rancourt KMS, et al. Decrease in depression symptoms is associated with longer survival in patients with metastatic breast cancer: a secondary analysis. J Clin Oncol 2011; 29(4): 413–420. [PubMed: 21149651]
5. Huscher D, Saketkoo LA, Pittrow D, et al. Development of clinical trial assessments for the study of interstitial lung disease in patients who have connective tissue diseases-methodological considerations. Curr Rheumatol Rev 2010; 6(2): 145–150. [PubMed: 20676224]
6. Jaeger VK, Distler O, Maurer B, et al. Functional disability and its predictors in systemic sclerosis: a study from the DeSScipher project within the EUSTAR group. Rheumatology 2018; 57(3): 441–450. [PubMed: 28499034]
7. Kanwal F, Gralnek IM, Hays RD, et al. Health-related quality of life predicts mortality in patients with advanced chronic liver disease. Clin Gastroenterol Hepatol 2009; 7(7): 793–799. [PubMed: 19306949]
8. Russell A-M, Scholand MB, Snyder EA, et al. Impact, survival, symptoms and management: US and UK patient perceptions of living with idiopathic pulmonary fibrosis In: C102 strategies to understand ILD: registries, prognostic indicators and more, vol. 193, 17 5 2016, p. A6222 American Thoracic Society, https://www.atsjournals.org/doi/abs/10.1164/ajrccm-conference.2016.193.1_MeetingAbstracts.A6222
9. Liang JW, Cheung YK, Willey JZ, et al. Quality of life independently predicts long-term mortality but not vascular events: the Northern Manhattan Study. Qual Life Res 2017; 26(8): 2219–2228. [PubMed: 28357682]
10. Strookappe B, Saketkoo LA, Elfferich M, et al. Physical activity and training in sarcoidosis: review and experience-based recommendations. Expert Rev Respir Med 2016; 10(10): 1057–1068. [PubMed: 27552344]
11. Irwin KE, Greer JA, Khatib J, et al. Early palliative care and metastatic non-small cell lung cancer: potential mechanisms of prolonged survival. Chron Respir Dis 2013; 10(1): 35–47. [PubMed: 23355404]
12. Rokach A, Romem A, Arish N, et al. The effect of pulmonary rehabilitation on non-chronic obstructive pulmonary disease patients. Isr Med Assoc J 2019; 5(21): 326–329. [PubMed: 31140224]
13. Saketkoo LA, Matteson EL, Brown KK, et al. Developing disease activity and response criteria in connective tissue disease-related interstitial lung disease. J Rheumatol 2011; 38(7): 1514–1518. [PubMed: 21724725]
14. Saketkoo LA, Mittoo S, Frankel S, et al. Reconciling healthcare professional and patient perspectives in the development of disease activity and response criteria in connective tissue disease-related interstitial lung diseases. J Rheumatol 2014; 41(4): 792–798. [PubMed: 24488412]
15. Saketkoo LA, Mittoo S, Huscher D, et al. Connective tissue disease related interstitial lung diseases and idiopathic pulmonary fibrosis: provisional core sets of domains and instruments for use in clinical trials. Thorax 2014; 69(5): 428–436. [PubMed: 24368713]

16. Saketkoo LA and Pauling JD. Qualitative methods to advance care, diagnosis, and therapy in rheumatic diseases. Rheum Dis Clin North Am 2018; 44(2): 267–284. [PubMed: 29622294]
17. Mittoo S, Frankel S, LeSage D, et al. Patient perspectives in OMERACT provide an anchor for future metric development and improved approaches to healthcare delivery in connective tissue disease related interstitial lung disease (CTD-ILD). Curr Respir Med Rev 2015; 11(2): 175–183. [PubMed: 26568747]
18. Carel H Breathlessness: the rift between objective measurement and subjective experience. Lancet Respir Med 2018; 6(5): 332–333. [PubMed: 29523434]
19. Carel H, Macnaughton J and Dodd J. Invisible suffering: breathlessness in and beyond the clinic. Lancet Respir Med 2015; 3(4): 278–279. [PubMed: 25890648]
20. Macnaughton J, Oxley R, Rose A, et al. Chronic breath-lessness: re-thinking the symptom. Eur Respir J 2018; 51(1): 1702331. [PubMed: 29371391]
21. Ross L, Stevens W, Wilson M, et al. Can patient-reported symptoms be used to measure disease activity in systemic sclerosis. Arthritis Care Res. 10.1002/acr.24053.
22. Tashkin DP, Volkmann ER, Tseng C-H, et al. Improved cough and cough-specific quality of life in patients treated for scleroderma-related interstitial lung disease: results of Scleroderma Lung Study II. Chest 2017; 151(4): 813–820. [PubMed: 28012804]
23. Theodore AC, Tseng C-H, Li N, et al. Correlation of cough with disease activity and treatment with cyclophosphamide in scleroderma interstitial lung disease: findings from the Scleroderma Lung Study. Chest 2012; 142(3): 614–621. [PubMed: 22156609]
24. Ryerson CJ, Abbritti M, Ley B, et al. Cough predicts prognosis in idiopathic pulmonary fibrosis. Respirology 2011; 16(6): 969–975. [PubMed: 21615619]
25. Brignall K, Jayaraman B and Birring SS. Quality of life and psychosocial aspects of cough. Lung 2008; 186(Suppl. 1): S55–S58. [PubMed: 17939003]
26. Raj AA, Pavord DI and Birring SS. Clinical cough IV: what is the minimal important difference for the Leicester Cough Questionnaire. Handb Exp Pharmacol 2009; 187: 311–320.
27. Khanna D, Clements PJ, Furst DE, et al. Correlation of the degree of dyspnea with health-related quality of life, functional abilities, and diffusing capacity for carbon monoxide in patients with systemic sclerosis and active alveolitis: results from the Scleroderma Lung Study. Arthritis Rheum 2005; 52(2): 592–600. [PubMed: 15692967]
28. Tashkin DP, Elashoff R, Clements PJ, et al. Cyclophosphamide versus placebo in scleroderma lung disease. N Engl J Med 2006; 354(25): 2655–2666. [PubMed: 16790698]
29. Tashkin DP, Roth MD, Clements PJ, et al. Mycophenolate mofetil versus oral cyclophosphamide in scleroderma-related interstitial lung disease (SLS II): a randomised controlled, double-blind, parallel group trial. Lancet Respir Med 2016; 4(9): 708–719. [PubMed: 27469583]
30. Russell A-M, Scholand MB, Snyder EA, et al. Trajectory of symptom burden, impact and survival in an idiopathic pulmonary fibrosis population In: C102 strategies to understand ILD: registries, prognostic indicators and more, 17 5 2016, p. A6221 American Thoracic Society, https://www.atsjournals.org/doi/abs/10.1164/ajrccm-conference.2016.193.1_MeetingAbstracts.A6221
31. Russell A-M, Fraser U, Molyneaux P, et al. Quality of life measures in patients with idiopathic pulmonary fibrosis. Eur Respir J 2013; 42(Suppl. 57): P3368 https://erj.ersjournals.com/content/42/Suppl_57/P3368
32. Saketkoo LA. Wildflowers abundant in the garden of systemic sclerosis research, while hopeful exotics will one day bloom. Rheumatology 2018; 57(3): 410–413. [PubMed: 29272533]
33. Saketkoo LA, Escorpizo R, Keen KJ, et al. International Classification of Functioning, Disability and Health Core Set construction in systemic sclerosis and other rheumatic diseases: a EUSTAR initiative. Rheumatology 2012; 51(12): 2170–2176. [PubMed: 22919048]
34. Poole JL and Steen VD. The use of the Health Assessment Questionnaire (HAQ) to determine physical disability in systemic sclerosis. Arthritis Care Res 1991; 4(1): 27–31. [PubMed: 11188583]
35. Steen VD and Medsger TA Jr. The value of the Health Assessment Questionnaire and special patient-generated scales to demonstrate change in systemic sclerosis patients over time. Arthritis Rheum 1997; 40(11): 1984–1991. [PubMed: 9365087]

36. Sullivan KM, Goldmuntz EA, Keyes-Elstein L, et al. Myeloablative autologous stem-cell transplantation for severe scleroderma. N Engl J Med 2018; 378(1): 35–47. [PubMed: 29298160]
37. Van Laar JM, Farge D, Sont JK, et al. Autologous hematopoietic stem cell transplantation vs intravenous pulse cyclophosphamide in diffuse cutaneous systemic sclerosis: a randomized clinical trial. JAMA 2014; 311(24): 2490–2498. [PubMed: 25058083]
38. Volkmann ER, Tashkin DP, Li N, et al. Mycophenolate mofetil versus placebo for systemic sclerosis-related interstitial lung disease: an analysis of scleroderma lung studies I and II. Arthritis Rheumatol 2017; 69(7): 1451–1460. [PubMed: 28376288]
39. Volkmann ER, Tashkin DP, Sim M, et al. Cyclophosphamide for systemic sclerosis-related interstitial lung disease: a comparison of scleroderma lung study I and II. J Rheumatol 2019; 46(10): 1316–1325. [PubMed: 30770517]
40. Distler O, Highland KB, Gahlemann M, et al. Nintedanib for systemic sclerosis-associated interstitial lung disease. N Engl J Med 2019; 380(26): 2518–2528. [PubMed: 31112379]
41. U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. Health Qual Life Outcomes 2006; 4: 79. [PubMed: 17034633]
42. Jauhar S. The demise of the physical exam. N Engl J Med 2006; 354(6): 548–551. [PubMed: 16467540]
43. Khanna D, Saggar R, Mayes MD, et al. A one-year, phase I/IIa, open-label pilot trial of imatinib mesylate in the treatment of systemic sclerosis-associated active interstitial lung disease. Arthritis Rheum 2011; 63(11): 3540–3546. [PubMed: 21769849]
44. Khanna D, Denton CP, Jahreis A, et al. Safety and efficacy of subcutaneous tocilizumab in adults with systemic sclerosis (faSScinate): a phase 2, randomised, controlled trial. Lancet 2016; 387(10038): 2630–2640. [PubMed: 27156934]
45. Khanna D, Denton CP, Lin CJF, et al. Safety and efficacy of subcutaneous tocilizumab in systemic sclerosis: results from the open-label period of a phase II randomised controlled trial (faSScinate). Ann Rheum Dis 2018; 77(2): 212–220. [PubMed: 29066464]
46. Bottomley A, Jones D and Claassens L. Patient-reported outcomes: assessment and current perspectives of the guidelines of the Food and Drug Administration and the reflection paper of the European Medicines Agency. Eur J Cancer 2009; 45(3): 347–353. [PubMed: 19013787]
47. Giacomini MK and Cook DJ. Users' guides to the medical literature: XXIII. Qualitative research in health care A. Are the results of the study valid? Evidence-Based Medicine Working Group. JAMA 2000; 284(3): 357–362. [PubMed: 10891968]
48. Tong A, Sainsbury P and Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007; 19(6): 349–357. [PubMed: 17872937]
49. Adcock CJ. Simplified factor analysis. Occup Psychol 1946; 20(4): 188–198. [PubMed: 21003014]
50. Gibbons RD, Clark DC, VonAmmon Cavanaugh S, et al. Application of modern psychometric theory in psychiatric research. J Psychiatr Res 1985; 19(1): 43–55. [PubMed: 3989737]
51. Rasch G. An item analysis which takes individual differences into account. Br J Math Stat Psychol 1966; 19(1): 49–57. [PubMed: 5939145]
52. Ware JE Jr and Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992; 30(6): 473–483. [PubMed: 1593914]
53. Yorke J, Moosavi SH, Shuldham C, et al. Quantification of dyspnoea using descriptors: development and initial testing of the Dyspnoea-12. Thorax 2010; 65(1): 21–26. [PubMed: 19996336]
54. Yorke J, Jones PW and Swigris JJ. Development and validity testing of an IPF-specific version of the St George's Respiratory Questionnaire. Thorax 2010; 65(10): 921–926. [PubMed: 20861296]
55. Chu LF, Utengen A, Kadry B, et al. 'Nothing about us without us' – patient partnership in medical conferences. BMJ 2016; 354: i3883. [PubMed: 27628427]

56. Pakas I, Ioannidis JPA, Malagari K, et al. Cyclophosphamide with low or high dose prednisolone for systemic sclerosis lung disease. J Rheumatol 2002; 29(2): 298–304. [PubMed: 11842824]
57. Vanthuyne M, Blockmans D, Westhovens R, et al. A pilot study of mycophenolate mofetil combined to intravenous methylprednisolone pulses and oral low-dose glucocorticoids in severe early systemic sclerosis. Clin Exp Rheumatol 2007; 25(2): 287–292. [PubMed: 17543155]
58. Fraticelli P, Gabrielli B, Pomponio G, et al. Low-dose oral imatinib in the treatment of systemic sclerosis interstitial lung disease unresponsive to cyclophosphamide: a phase II pilot study. Arthritis Res Ther 2014; 16(4): R144. [PubMed: 25007944]
59. Spiera RF, Gordon JK, Mersten JN, et al. Imatinib mesylate (Gleevec) in the treatment of diffuse cutaneous systemic sclerosis: results of a 1-year, phase IIa, single-arm, open-label clinical trial. Ann Rheum Dis 2011; 70(6): 1003–1009. [PubMed: 21398330]
60. Seibold JR, Denton CP, Furst DE, et al. Randomized, prospective, placebo-controlled trial of bosentan in interstitial lung disease secondary to systemic sclerosis. Arthritis Rheum 2010; 62(7): 2101–2108. [PubMed: 20506355]
61. Daoussis D, Liossis Tsamandas AC, Kalogeropoulou C, et al. Experience with rituximab in scleroderma: results from a 1-year, proof-of-principle study. Rheumatology 2010; 49(2): 271–280. [PubMed: 19447770]
62. Khanna D, Albera C, Fischer A, et al. An open-label, phase II study of the safety and tolerability of pirfenidone in patients with scleroderma-associated interstitial lung disease: the LOTUSS trial. J Rheumatol 2016; 43(9): 1672–1679. [PubMed: 27370878]
63. Burt RK, Shah SJ, Dill K, et al. Autologous non-myeloablative haemopoietic stem-cell transplantation compared with pulse cyclophosphamide once per month for systemic sclerosis (ASSIST): an open-label, randomised phase 2 trial. Lancet 2011; 378(9790): 498–506. [PubMed: 21777972]
64. Payakachat N, Ali MM and Tilford JM. Can the EQ-5D detect meaningful change? A systematic review. Pharmacoeconomics 2015; 33(11): 1137–1154. [PubMed: 26040242]
65. King TE Jr, Bradford WZ, Castro-Bernardini S, et al. A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. N Engl J Med 2014; 370(22): 2083–2092. [PubMed: 24836312]
66. Richeldi L, du Bois RM, Raghu G, et al. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. N Engl J Med 2014; 370(22): 2071–2082. [PubMed: 24836310]
67. Wilson RC and Jones PW. A comparison of the visual analogue scale and modified Borg scale for the measurement of dyspnoea during exercise. Clin Sci 1989; 76(3): 277–282. [PubMed: 2924519]
68. Wilson RC and Jones PW. Long-term reproducibility of Borg scale estimates of breathlessness during exercise. Clin Sci 1991; 80(4): 309–312. [PubMed: 1851065]
69. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002; 166(1): 111–117. [PubMed: 12091180]
70. Buch MH, Denton CP, Furst DE, et al. Submaximal exercise testing in the assessment of interstitial lung disease secondary to systemic sclerosis: reproducibility and correlations of the 6-min walk test. Ann Rheum Dis 2007; 66(2): 169–173. [PubMed: 16868020]
71. Wilsher M, Good N, Hopkins R, et al. The six-minute walk test using forehead oximetry is reliable in the assessment of scleroderma lung disease. Respirology 2012; 17(4): 647–652. [PubMed: 22256786]
72. Papiris SA, Daniil ZD, Malagari K, et al. The Medical Research Council dyspnea scale in the estimation of disease severity in idiopathic pulmonary fibrosis. Respir Med 2005; 99(6): 755–761. [PubMed: 15878493]
73. Manali ED, Lyberopoulos P, Triantafillidou C, et al. MRC chronic Dyspnea Scale: relationships with cardiopulmonary exercise testing and 6-minute walk test in idiopathic pulmonary fibrosis patients: a prospective study. BMC Pulm Med 2010; 10: 32. [PubMed: 20509928]
74. Yorke J, Swigris J, Russell A-M, et al. Dyspnea-12 is a valid and reliable measure of breathlessness in patients with interstitial lung disease. Chest 2011; 139(1): 159–164. [PubMed: 20595454]
75. Holland AE, Fiore JF Jr, Bell EC, et al. Dyspnoea and comorbidity contribute to anxiety and depression in interstitial lung disease. Respirology 2014; 19(8): 1215–1221. [PubMed: 25112470]

76. Mahler DA, Weinberg DH, Wells CK, et al. The measurement of dyspnea: contents, interobserver agreement, and physiologic correlates of two new clinical indexes. Chest 1984; 85(6): 751–758. [PubMed: 6723384]
77. Roth MD, Tseng C-H, Clements PJ, et al. Predicting treatment outcomes and responder subsets in scleroderma-related interstitial lung disease. Arthritis Rheum 2011; 63(9): 2797–2808. [PubMed: 21547897]
78. Swigris JJ, Han M, Vij R, et al. The UCSD shortness of breath questionnaire has longitudinal construct validity in idiopathic pulmonary fibrosis. Respir Med 2012; 106(10): 1447–1455. [PubMed: 22801586]
79. Eakin EG, Resnikoff PM, Prewitt LM, et al. Validation of a new dyspnea measure: the UCSD Shortness of Breath Questionnaire. University of California, San Diego. Chest 1998; 113(3): 619–624. [PubMed: 9515834]
80. Swigris JJ, Yorke J, Sprunger DB, et al. Assessing dyspnea and its impact on patients with connective tissue disease-related interstitial lung disease. Respir Med 2010; 104(9): 1350–1355. [PubMed: 20471238]
81. Hinchcliff M, Beaumont JL, Thavarajah K, et al. Validity of two new patient-reported outcome measures in systemic sclerosis: Patient-Reported Outcomes Measurement Information System 29-item Health Profile and Functional Assessment of Chronic Illness Therapy-Dyspnea short form. Arthritis Care Res 2011; 63(11): 1620–1628.
82. Birring SS. Developing antitussives: the ideal clinical trial. Pulm Pharmacol Ther 2009; 22(2): 155–158. [PubMed: 19041729]
83. Birring SS, Prudon B, Carr AJ, et al. Development of a symptom specific health status measure for patients with chronic cough: Leicester Cough Questionnaire (LCQ). Thorax 2003; 58(4): 339–343. [PubMed: 12668799]
84. French CT, Irwin RS, Fletcher KE, et al. Evaluation of a cough-specific quality-of-life questionnaire. Chest 2002; 121(4): 1123–1131. [PubMed: 11948042]
85. Lechtzin N, Hilliard ME and Horton MR. Validation of the Cough Quality-of-Life Questionnaire in patients with idiopathic pulmonary fibrosis. Chest 2013; 143(6): 1745–1749. [PubMed: 23519393]
86. Fletcher KE, French CT, Irwin RS, et al. A prospective global measure, the Punum Ladder, provides more valid assessments of quality of life than a retrospective transition measure. J Clin Epidemiol 2010; 63(10): 1123–1131. [PubMed: 20303709]
87. Horton MR, Santopietro V, Mathew L, et al. Thalidomide for the treatment of cough in idiopathic pulmonary fibrosis: a randomized trial. Ann Intern Med 2012; 157(6): 398–406. [PubMed: 22986377]
88. Raj AA, Pavord DI and Birring SS. Clinical cough IV: what is the minimal important difference for the Leicester Cough Questionnaire? In: Chung KF and Widdicombe J (eds) Pharmacology and therapeutics of cough. Berlin; Heidelberg: Springer, 2009, pp. 311–320.
89. Petty TL. The National Mucolytic Study. Results of a randomized, double-blind, placebo-controlled study of iodinated glycerol in chronic obstructive bronchitis. Chest 1990; 97(1): 75–83. [PubMed: 2403903]
90. Swigris JJ, Brown KK, Behr J, et al. The SF-36 and SGRQ: validity and first look at minimum important differences in IPF. Respir Med 2010; 104(2): 296–304. [PubMed: 19815403]
91. Chang JA, Curtis JR, Patrick DL, et al. Assessment of health-related quality of life in patients with interstitial lung disease. Chest 1999; 116(5): 1175–1182. [PubMed: 10559073]
92. Patel AS, Siegert RJ, Brignall K, et al. The development and validation of the King's Brief Interstitial Lung Disease (K-BILD) health status questionnaire. Thorax 2012; 67(9): 804–810. [PubMed: 22555278]
93. O'Higgins CM, Brady B, O'Connor B, et al. The patho-physiology of cancer-related fatigue: current controversies. Support Care Cancer 2018; 26(10): 3353–3364. [PubMed: 29961146]
94. Minnock P, Kirwan J and Bresnihan B. Fatigue is a reliable, sensitive and unique outcome measure in rheumatoid arthritis. Rheumatology 2009; 48(12): 1533–1536. [PubMed: 19773406]
95. Al-Janabi H, Flynn TN and Coast J. Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. Qual Life Res 2012; 21(1): 167–176. [PubMed: 21598064]

96. Broadbent E, Petrie KJ, Main J, et al. The brief illness perception questionnaire. J Psychosom Res 2006; 60(6): 631–637. [PubMed: 16731240]

97. Osborne RH, Batterham RW, Elsworth GR, et al. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). BMC Public Health 2013; 13: 658. [PubMed: 23855504]

98. Arnold MB, Khanna D, Denton CP, et al. Patient acceptable symptom state in scleroderma: results from the tocilizumab compared with placebo trial in active diffuse cutaneous systemic sclerosis. Rheumatology 2018; 57(1): 152–157. [PubMed: 29077900]

99. Steen VD, Conte C, Owens GR, et al. Severe restrictive lung disease in systemic sclerosis. Arthritis Rheum 1994; 37(9): 1283–1289. [PubMed: 7945490]

100. Picker Institute Europe. Key domains of the experince of hospital outpatients. Discussion paper 2, 2010, https://www.picker.org/wp-content/uploads/2014/10/Discussion-paper-…-hospital-outpatients.pdf

101. Hibbard JH, Mahoney ER, Stockard J, et al. Development and testing of a short form of the patient activation measure. Health Serv Res 2005; 40(6 Pt. 1): 1918–1930. [PubMed: 16336556]

102. Hibbard JH, Stockard J, Mahoney ER, et al. Development of the patient activation measure (PAM): conceptualizing and measuring activation in patients and consumers. Health Serv Res 2004; 39(4 Pt. 1): 1005–1026. [PubMed: 15230939]

103. Poon BY, Shortell SM and Rodriguez HP. Patient activation as a pathway to shared decision-making for adults with diabetes or cardiovascular disease. J Gen Intern Med. 10.1007/s11606-019-05351-6.

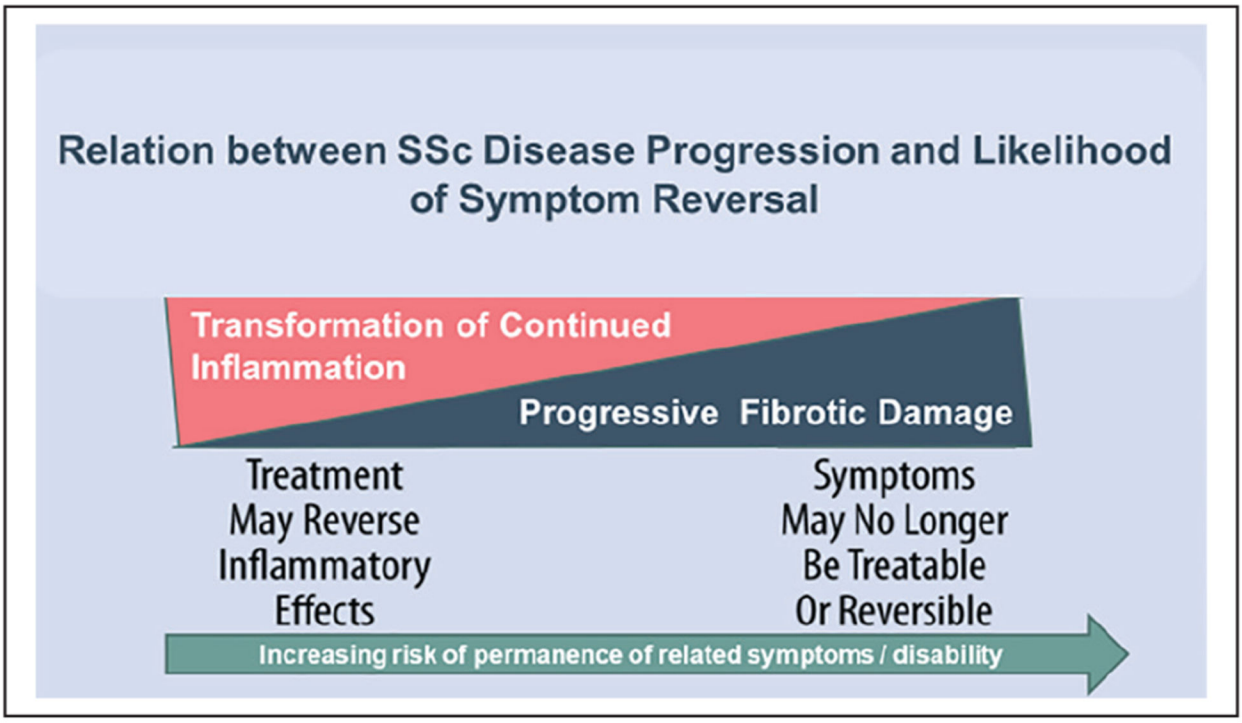

Figure 1.
Relation between SSc disease progression and likelihood of symptom reversal.

Table 1.

Proposed overarching concepts of PROM planning and development.

Team
Convening a qualified team
Question
Clarifying the need and goals of PROM research
Design
Designing the research to ensure saturation of concepts in relevant population
Implementation and qualitative analysis
Data collection, iterative analyses and revision leading to concept saturation and conceptual framework
Field-testing
Testing items for relevance of concept, language, phrasing and response format
Item reduction and fit
Quantitative field-testing for item reduction and psychometric testing
Validation
Multi-step process examining performance

PROM: patient-reported outcome measures.

Table 2.

2012 consensus items for CTD-ILD versus IPF.[15]

	Instrument	CTD-ILD	IPF
Dyspnea	MRC Chronic Dyspnea Scale	YES	YES
	Dyspnea 12	YES	YES
	UCSD-SOBQ	N/A	YES
Cough	Leicester Cough Questionnaire	YES	YES
HRQoL	Short Form-36	YES	YES
	SGRQ	YES	YES
	VAS-PtGA	YES	N/A

CTD-ILD: connective tissue disease–related interstitial lung disease; IPF: idiopathic pulmonary fibrosis; MRC: Medical Research Council Dyspnea Scale; YES: consensus was met; UCSD-SOBQ: University California San Diego Shortness of Breath Questionnaire; N/A: consensus was not met; HRQoL: health-related quality of life; SGRQ: St. George's Respiratory Questionnaire; VAS: Visual Analogue Scale; PROM: patient-reported outcome measures; SSc-ILD: systemic sclerosis–related interstitial lung disease; PtGA: Patient Global Assessment.

It is important to recognize that at this time of the recommendations several PROM instruments currently used today had not yet been developed and that for this study SSc-ILD was the prototype CTD-ILD.

Table 3.

Available prospective treatment studies comparing FVC and PROM performance.

Intervention	Study	Time	FVC	Physical function	HRQoL generic	HRQoL lung specific	Dyspnea	Cough	Fatigue/energy
CYC (SLS 1)	Tashkin et al.[28] Khanna et al.[27]	12 months	+	HAQ-DI+	SF-36+	n/a	MTDI+ SHAQ VAS+	Modified Cough Index+	SF-36 vitality domain +
CYC	Pakas et al.[56]	12 months	+				VAS+		
CYC versus MMF (SLS II)	Tashkin et al.[29]	24 months	+	HAQ-DI+	SF-36+	SGRQ	MTDI+ SHAQ VAS+	LCQ+	
MMF	Vanthuyne et al.[57]	12 months	+	HAQ-DI+	SHAQ VAS−		SHAQ VAS−		
Imitinab	Fraticelli et al.[58]	6 months	+	HAQ-DI−	SF-36−				
Imitinab	Khanna et al.[43]	12 months	+[a]	HAQ-DI−			MTDI[b]		
Imitinab	Spiera et al.[59]	12 months	+	HAQ-DI−	SF-36 PC− SF-36 MC+		SHAQ VAS+		
Bosentan	Seibold et al.[60]	12 months	−	HAQ-DI−			Borg− SHAQ VAS−		
Tocilizumab RCT	Khanna et al.[44]	48 weeks	+[a]	HAQ-DI+	Pt Global VAS+		SHAQ VAS+		FACIT-F+
Extension study	Khanna et al[45]	+48 weeks	+[a]	HAQ-DI+	Pt Global VAS+		SHAQ VAS+		FACIT-F+
Rituximab	Daoussis et al.[61]	12 months	+	HAQ-DI+					
Pirfenidone Open label Phase II	Khanna et al.[62]	16 weeks	−	HAQ-DI−	Pt Global VAS−		MTDI+		
HSCT	Van Laar et al.[37]	24 months	+	HAQ-DI+	SF-36+ MC− PC+ EQ5D+				
HSCT	Sullivan et al.[36]	54 months	+	HAQ-DI+	SF-36+				
HSCT	Burt et al.[63]	12 months	+		SF-36+				
Nintedinab	Distler et al[40]	52 weeks	+	HAQ-DI+		SGRQ+	FACIT D+ SHAQ VAS+		

FVC: forced vital capacity; PROM: patient-reported outcome measures; HRQoL: health-related quality of life; CYC: cyclophosphamide; SLS: Scleroderma Lung Study; +: significant improvement; HAQDI: Health Assessment Questionnaire–Disability Index; SF-36: short form-36; –: non-significant change among prospective studies examining SSc-ILD; MTDI: Mahler Transitional Dyspnea Index; SHAQ: Scleroderma Health Assessment Questionnaire; VAS: Visual Analogue Scale; MMF: mycophenolate mofetil; SGRQ: St. George's Respiratory Questionnaire; LCQ: Leicester Cough Questionnaire; PC: physical component; MC: mental component; RCT randomised controlled trial; Pt-GA: Patient Global Assessment; FACIT-F: Functional Assessment of Chronic Illness Therapy–Fatigue Scale; HSCT: hematopoietic stem cell transplant; EQ-5D: EurQoL 5 Dimensions; SSc-ILD: systemic sclerosis–related interstitial lung disease.

[a]Clinically but not statistically significant improvement.

[b] Statistically but not clinically significant

Table 4.

Validation parameters of PROMs considered for SSc-ILD organized by content, responsiveness and utility.

Instruments		**Dyspnea**		**SHAQ Breathing VAS**	**UCSD SOBQ**	**MTDI generic**	**Cough**	**HRQoL**					
								Physical function					**SHAQ SSc**
		D-12 generic	**MRC generic**				**LCQ generic**	**SF-36 generic**	**EQ-SD generic**	**SGRQ respiratory**	**Pt-GA[a] generic**	**HAQ-DI generic**	
Content	Face validity	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Content validity	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Construct validity	Y	Y	Y	Y	?	Y	Y	Y	Y	NT	Y	Y
	Criterion validity	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT
Response	Discriminatory	Y	Y	Y	Y	UC	Y	Y	UC[b]	Y	NT	Y	Y
	Reliable	Y	Y	Y	NT	UC	Y	Y	Y	NT	NT	Y	Y
	Reproducible	NT	NT	Y	NT	Y	NT	NT	Y	NT	NT	Y	Y
	Sensitive to change	Y	Y	Y	Y	Y	Y	Y	UC[b]	NT	NT	Y	Y
Utility	Cost effective	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Interpretability	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Readily available	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Safety	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Time efficient	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Discrete components								Physical mental vitality					SSc Sx Pt-GA
Developed for ILD		N	N	Y[c]	N	N	N	N	N	N	N	N	Y
Patient-derived content		Y	No	No	No	No	No	No	No	No	No	No	No

PROM: patient-reported outcome measures; SSc-ILD: systemic sclerosis–related interstitial lung disease; D-12: Dyspnea-12; MRC: Medical Research Council Dyspnea Scale; SHAQ: Scleroderma Health Assessment Questionnaire; VAS: Visual Analogue Scale; UCSD SOBQ: University of San Diego Shortness of Breath Questionnaire; MTDI: Mahler Transitional Dyspnea Index; LCQ: Leicester Cough Questionnaire; HRQoL: health-related quality of life; SF-36: short form-36; EQ-5D: EurQoL 5 Dimensions; SGRQ: St. George's Respiratory Questionnaire; Pt-GA: Patient Global Assessment; HAQ-DI: Health Assessment Questionnaire-Disability Index; SSc: systemic sclerosis; Y: yes; NT: not yet tested; UC: unclear; ILD: interstitial lung disease.

[a] Pt-GA is an independent instrument categorized under HRQoL.

[b] EQ-5D has variable performance depending on condition.[64]

$c_{SHAQ\ VAS.}$