Published in final edited form as:

Curr Respir Med Rev. 2015 ; 11(2): 163–174. doi:10.2174/1573398X11666150619183527.

Outcome Measures for Clinical Trials in Interstitial Lung Diseases

Matthew R. Lammi^{1,2}, Robert P. Baughman³, Surinder S. Birring⁴, Anne-Marie Russell⁵, Jay H. Ryu⁶, Marybeth Scholand⁷, Oliver Distler⁸, Daphne LeSage⁹, Catherine Sarver¹⁰, Katerina Antoniou¹¹, Kristin B. Highland¹², Otylia Kowal-Bielecka¹³, Joseph A. Lasky^{2,14}, Athol U. Wells⁵, and Lesley Ann Saketkoo^{2,14,*}

¹Louisiana State University Health Sciences Center, New Orleans, New Orleans, LA, USA ²New Orleans Scleroderma and Sarcoidosis Patient Care and Research Center; New Orleans, LA, USA ³University of Cincinnati, OH, USA ⁴Division of Asthma, Allergy and Lung Biology, King's College London, UK ⁵Royal Brompton Hospital and National Heart and Lung Institute; London, UK ⁶Mayo Clinic College of Medicine, Rochester, MN, USA ⁷University of Utah, Salt Lake City, UT, USA ⁸Division of Rheumatology, University Hospital Zurich, Switzerland ⁹Patient Research Partner, Office of Public Health, New Orleans, LA, USA ¹⁰Patient Research Partner, Maryland, USA ¹¹University of Crete, Heraklion, Greece ¹²Cleveland Clinic Foundation, Respiratory Institute, Cleveland, OH, USA ¹³Medical University of Bialystok, Bialystok, Poland ¹⁴Tulane University Lung Center; New Orleans, LA, USA

Abstract

The chronic fibrosing idiopathic interstitial pneumonias (IIPs) are a group of heterogeneous pulmonary parenchymal disorders described by radiologic and histological patterns termed usual interstitial pneumonia (UIP) and non-specific interstitial pneumonia (NSIP). These include idiopathic pulmonary fibrosis (IPF) and those related to connective tissue disease (CTD) and are associated with substantial morbidity and mortality. Beyond the importance of establishing an appropriate diagnosis, designing optimal clinical trials for IIPs has been fraught with difficulties in consistency of clinical endpoints making power analyses, and the establishment of efficacy and interpretation of results across trials challenging.

Preliminary recommendations, developed by rigorous consensus methods, proposed a minimum set of outcome measures, a 'core set', to be incorporated into future clinical trials (Saketkoo et al, THORAX. 2014.). This paper sets out to examine the candidate instruments for each domain (Dyspnea, Cough, Health Related Quality of Life, Imaging, Lung Physiology and Function, Mortality). Candidate measures that were not selected as well as measures that were not available for examination at the time of the consensus process will also be discussed.

*Address correspondence to this author at the Tulane University Lung Center, New Orleans Scleroderma and Sarcoidosis Patient Care and Research Center, Department of Medicine, Sections of Rheumatology and Pulmonary Medicine and Critical Care, 1324 Tulane Avenue, New Orleans, LA 70112, USA; Tel: 1 504 988 0743; Fax: 1 504 988 2144; CTD.ILD@gmail.com.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

Keywords

Chronic fibrosing; connective tissue disease related interstitial lung disease; idiopathic interstitial pneumonia; idiopathic pulmonary fibrosis; rheumatoid arthritis; scleroderma; systemic sclerosis

INTRODUCTION

The Idiopathic interstitial pneumonias (IIPs) are a group of heterogeneous pulmonary parenchymal disorders with substantial morbidity and mortality. Of the six IIPs [1] the two IIPs of interest for the purposes of this review are pathologically and radiographically classified as Usual Interstitial Pneumonitis (UIP) and Non-Specific Interstitial Pneumonitis (NSIP) and well-described elsewhere in this special issue [2,3]. UIP and NSIP are categorized as chronic fibrosing IIPs with UIP being prototypically fibrotic on pathology and NSIP containing varying degrees of inflammation and fibrosis. Current clinical trials in chronic fibrosing IIPs investigate radical interventions that target fibrosing processes rather than palliative interventions that address associative symptoms. Placebo-controlled trials in IPF nor CTD-ILD are any longer ethically appropriate given the availability of potentially efficacious treatment - thus heightening the need to address selection of measures that are very sensitivity to change over time while maintain reliability.

Treatment response is widely variable across and within these patterns, thus establishing a specific diagnosis or realm of non-specific but potentially treatable disease such as Interstitial Pulmonary Autoimmune Fibrosis (IPAF) [4] critical in intervention clinical trials. UIP, associated predominantly with idiopathic pulmonary fibrosis (IPF), is a parenchymal disease of fibrosis with no or extremely sparse inflammation on biopsy; it may also be associated with other entities such as rheumatoid arthritis (RA); making careful clinical and serologic assessment essential - as treatment response of IPF is distinct; however both are unfavorable [5,6]. NSIP is sometimes idiopathic but is often associated with connective tissue disease (CTD) such as RA, idiopathic inflammatory myopathies (IIM) and Sjogren's Syndrome (SjS), of which systemic sclerosis (SSc) is the most rigourously studied [7–11] (hence, extensively discussed herein). NSIPs are associated with some responsiveness to immune suppressive agents.

Beyond the importance of establishing an appropriate diagnosis, designing optimal clinical trials for IIPs has been fraught with difficulties in consistency of clinical endpoints making power analyses, and the establishment of efficacy and interpretation of results across trials challenging. A recent effort attempted to address this dilemma via consensus of an international community of IIP experts [12,13] and provide preliminary recommendations for the minimum set of outcome measures to be incorporated into future clinical trials.

The consensus effort was a mixed methodology approach [12,13] and included >250 ILD experts and 45 patient participants. Ultimately, there were 5 agreed upon domains for each IPF and CTD-ILD: *Dyspnea, Cough, Health-Related Quality of Life (HRQoL), Lung Physiology and Function, Imaging and Survival*. Each of these domains was required to contain at least one outcome measure for each IPF and CTD-ILD (See table of measures for details). Other domains such as *Biomarkers* were regarded as promising, but no such

measure for clinical trials in the near future was recognized as sufficiently reliable or specific to characterize a change in disease state over time. This paper sets out to discuss this 'core set' of domain-related instruments in terms of validity and feasibility for use in clinical trials. We also examine some measures that were not selected as well as measures that were not available for examination at the time of the consensus process.

DYSPNEA AS AN OUTCOME MEASURE

Dyspnea is the most common complaint of patients with ILD; however, no prospective studies evaluate changes in dyspnea as a primary end point - and in most clinical trials not even as a secondary end point. HRQoL instruments, such as the St. George's Respiratory Questionnaire (SGRQ) have been utilized as a reflection of the burden of dyspnea, whereas dyspnea-specific instruments have demonstrated utility in assessing response to treatment [14,15]. In the final phases of the consensus effort, four instruments in addition to Borg Dyspnea Index at rest and also pre/post exercise [16] were discussed: the MRC Breathlessness Scale (MRC) [17], the Mahler Baseline and Transitional Dyspnea Index (BDI and TDI) [18], the University of California San Diego Shortness of Breath Questionnaire (UCSD-SBQ) [19], and the Dyspnea 12 index [20]. Dyspnea scales are inexpensive and free of risk to patients.

The Borg scale is a linear point scale from 0–10 with descriptors, used at rest and pre/post exercise, such as with the six minute walk test (6MWT) [21]. It has been validated in healthy controls and a variety of pulmonary conditions [22,23]. In SSc-ILD, both the rest and post exercise Borg scores were found to be reproducible [7]. A weak but significant correlation between the post exercise Borg score and six minute walk distance (6MWD) was demonstrated in one study [7] but was not reproducible [8] while the Borg is a component of the 6MWT, it has not been used as a major endpoint in clinical trials.

The Medical Research Council Breathless Scale (MRC) and the American Thoracic Society Dyspnea scale are almost identical as both require a single response to a five-point scale of dyspnea related to activity/functional descriptors that best describes a patient's perceived level of dyspnea in relation to activities. The MRC has been validated in IPF, but not in SSc [24]. It was found to be reproducible and to correlate with several other parameters, including the post exercise Borg score [25] and 6MWD [26]. MRC score had also been found to be an independent predictor of anxiety and depression in a mixed ILD cohort [27].

While the MRC has not been specifically reported in a SSc-ILD interventional trial, it was included in the preliminary core set recommended for future studies, because of its low cost, feasibility of administration and the broad but specific characterizations of dyspnea in the response scale, which has led to its wide application in many other pulmonary diseases.

Mahler developed a baseline and translational dyspnea index (BDI and TDI respectively) for patients with COPD [28]. The BDI and follow-up TDI consist of three questions administered by a third party that attempts to deconstruct dyspnea-related impairment under three domains: level of 'functional impairment', 'magnitude of task' and 'magnitude of effort' (combination of time and effort to complete task). Because of the abstract nature of the questions, responses can be administrator dependent and cumbersome, but a standard

computerized version is available. This is arguably contrasted with the MRC whereby its single response items each capture the combined effect of these domains barring the change in time intensity. Both have been partially validated in SSc-ILD. The BDI was predictive of response to cyclophosphamide therapy for SSc-ILD patients in one large, randomized trial (with important implications for SSc-ILD cohort enrichment). TDI was able to distinguish between patients with improved lung function in response to cyclophosphamide versus worsening in those receiving placebo [9,10] with calculation of a minimal clinically important difference [29]. Largely due to feasibility issues, TDI was required to undergo further research in SSc-ILD before recommendation as a major endpoint in future trials [12].

The UCSD-SBQ has demonstrated validity in ILDs including SSc-ILD [19,30,31] and was used as a major secondary end point in IPF trials [14]. The current data supported its use in IPF trials, but not yet for S-ILD.

The D-12 is a compilation of 12 simple questions inquiring about the action of breathing (e.g. requires work, does not go in all way) and its affective affects (irritating, exhausting etc), each comprised of an average of five words requiring response on a four-point scale from none to severe. It has demonstrated reliability in ILDs including SSc-ILD [20,26,30] and correlates with 6MWD [26]. D-12 purely targets breath perception avoiding reference activities which may be an important confounder in conditions that have musculoskeletal, restrictive skin or other systemic inflammatory manifestations such as fatigue, which are common in CTD-ILDs; therefore D-12 was accepted as a potential measure for future studies in IIPs with strong support for continued assessment of its performance in trial context.

It is important to remember that none of the above instruments was developed specifically for IIPs. Although the D-12 was built with patient response to expert-collected concepts, none of the dyspnea instruments were developed by current accepted methods of patient-generated language and concepts. Further, these instruments were mostly developed in the context of obstructive pulmonary disease which differs from restrictive lung physiology both in mechanisms for dyspnea, (e.g. dyspnea due to hypoxia more frequent in restrictive disease) and patient-reported qualitative. It is also important to note that current dyspnea scales were developed to describe symptom severity at a single point in time and thus have strong utility in stratifying severity; but contain limitations in assessing serial change. For example, in the MRC or ATS dyspnea scales, change is highly dependent on proximity to patient status to threshold; meaning that minor changes can result in change of category while major changes may result in no category change. Perhaps a scale that is able to avert recall-bias and capture the patient's perception of dyspnea along a spectrum of 'decline-no change-improvement' may better reflect serial change. The preliminary guidelines for future IIP studies supported the development of questionnaires specific for dyspnea in CTD-ILD and IPF patients (9). Beyond, the instruments available at the time of the consensus effort, the functional assessment of chronic illness therapy-dyspnea (FACIT-D) was recently validated in a cohort of SSc-ILD patients [11].

COUGH AS AN OUTCOME MEASURE

Cough is an important symptom of most chronic lung disorders associated with an adverse impact on health-related quality of life (HRQoL) [32]. Physical symptoms associated with cough include chest pain, vomiting, hoarse voice and sleep disturbance [33]. Psychosocial symptoms associated with cough include anxiety and depression, social embarrassment and disruption to work and social activities [34]. Cough is an important symptom of ILD. In IPF, cough is associated with poor HRQoL and disease progression [35]. In CTD-ILD, the prevalence of cough has been reported to be as high as 73% [36] and is associated with a poor quality of life, and worse extent of lung fibrosis [36]. The mechanism of cough in ILD is poorly understood. Cough reflex hypersensitivity, airway inflammation, lung distortion and gastro-esophageal reflux (GERD) are potential mechanisms [37].

Cough can be measured for clinical practice or for clinical trials using instruments that capture severity, frequency, intensity and/or impact on HRQoL [38]. Cough Visual Analogue Scales (VAS) can assess status and changes in cough severity [39]. Examples of cough questionnaires include the Leicester Cough Questionnaire (LCQ) and the Cough-specific Quality of Life Questionnaire (CQLQ) [40,41]. The urge to cough may reflect cough reflex hypersensitivity, and this can be assessed with urge to cough Visual Analogue Scales [42]. Cough monitors afford an objective measure of cough *frequency* [43]. They are comprised of an external microphone that records patient's cough over periods greater than 24 hours. The Leicester Cough Monitor (LCM) and Vitalojak are two validated monitors [44,45]. Although they are now being used as the preferred primary endpoint in clinical trials for anti-tussive medications, they are best utilised in combination with subjective tools such as VAS and HRQoL assessments. It is not currently possible to assess cough *intensity* objectively in clinical trials. The relevance of measuring cough intensity for the evaluation of anti-tussive medications is therefore not known.

The importance of assessing cough in clinical trials of ILD has been investigated recently [12]. Nevertheless, the experts surprisingly reached consensus at an initial stage to exclude cough as an endpoint in clinical trials of ILD [12]. In contrast, a group of patients reached consensus that cough was an important symptom. They reported that cough was a central symptom of ILD and reflected worsening of their condition [12]. The reason why this group of experts initially excluded cough as an outcome measure of ILD is unclear; fortunately cough was reintroduced into the consensus process based on patient feedback. The experts eventually concluded that cough should be assessed in clinical trials, and that the Leicester Cough Questionnaire (LCQ) was an appropriate clinical outcome tool for this purpose [12]. The LCQ is the most widely used of all cough tools in clinical trials with a minimally important difference (MID) of 1.3 [34], and anticipated to capture serial change albeit not specifically validated in IIPs. It has been translated and used in many cultures [47]. Key et al have assessed the LCQ in IPF [48]. It is the only study that has compared cough-specific quality of life measures with objective cough frequency. There was a strong correlation between the LCQ and all its domains with cough frequency.

There are other options to assess cough in ILD. The CQLQ is another validated cough-specific quality of life tool. The CQLQ total score has good internal reliability in IPF, but

this does not apply to some of its domains, such as extreme physical complaints and personal safety fears [49]. Similarly, it has good concurrent validity for its total score but, this did not apply to some domains such as personal safety fears. The minimal important difference (MID) in IPF has been reported for the CQLQ (5–6 units). This is, however, less than that reported for this tool for patients with chronic cough (13 units). The reasons for this are not clear. The MID was assessed with a retrospective anchor scale that assessed change. The authors of the original CQLQ have reported limitations of this methodology when they evaluated the MID of the CQLQ for chronic cough [50]. However, this method is widely used by other investigators [34]. The CQLQ has been used to evaluate the efficacy of Thalidomide in IPF, and was reported to be a responsive outcome measure for this purpose [51].

In conclusion, cough is a common symptom and important to patients with ILD. For now, the LCQ which has demonstrated feasibility and sensitivity to serial change; is an appropriate measure to assess cough in clinical trials.

HEALTH RELATED QUALITY OF LIFE (HRQoL)

Dyspnea and *cough*, as described above are important symptoms that are patient-reported outcome measures (PROMs). This section will highlight features and process of PROM development in the context of HRQoL. PROMs attempt to quantify a person's perception of health status, symptoms, or QoL and can reveal information patients may not disclose spontaneously [52]. A PROM is any report of the status of a patient's health condition provided directly by the patient, without interpretation of the patient's response by a clinician or third party [53]. PROMs are valuable in identifying differences between groups of patients with either IPF and CTD-ILD undetected by physiologic or radiologic measures - as in *cough* and *dyspnea*. A PROM can be described in absolute terms (e.g., severity of a symptom, sign, or state of a disease), which may have utility in cohort enrichment or as a change over time, which is of importance in clinical trials. In any disease, candidate instruments should be chosen carefully with consideration given both to the content of the instrument and the intervention [54]. Historically PROMs were rarely developed in concert with priority content and language from patients. This is no longer acceptable by the Food and Drug Administration [53]. Instead, collection of qualitative data from patients with subsequent investigations in item construction with item response analysis is required.

Disease symptoms often impact a person's ability to carry on a 'normal' life, and changes in this impact are measurable. In aggregate, these multi-faceted areas are referred to as HRQoL and are defined by personal health status referring to aspects of life influenced by mental or physical well-being [55]. An HRQoL PROM may be scored as a broad survey or may have the utility of providing scores for individual components/domains; of which mental health, physical function, and pain are common domains. Other important domains include social engagement and work productivity. In addition to gauging changes in patient impact of disease, treatment interventions for ILDs often carry high side effect profiles. HRQoLs may provide an important signal to ensure that the investigated treatment does not have a more negative impact than the disease itself. This is particularly crucial when treatment is not

curative. A HRQoL PROM may be disease-specific [56] or generic [57] (used broadly across diseases); both are recommended to capture HRQoL signals in treatment trials.

HRQoL is impaired in IPF and CTD-ILD particularly in domains pertaining to symptoms or physical activity [12,51,58–60]. In IPF and CTD-ILD research, HRQoL has only ever been a secondary endpoint. Although physiological outcomes (e.g. Forced Vital Capacity (FVC)) yield information on disease stage and provide prognostic information that is essential to clinical management; PROMs disclose what patients with IPF and CTD-ILD are actually experiencing. In turn, that experience affects any number of life domains and correlates to overall health outcomes. This is essential to effective therapeutic management.

At present there is no concise disease-specific HRQoL PROM to assess the experiential benefits or detriments of treatments in IPF and CTD-ILD. The lack of HRQoL PROMs and longitudinal validation of symptom measures as well as the paucity of research in PROMs in CTD-ILD and IPF has created an environment of uncertainty and has led to difficulties for investigators in confidently interpreting PROM scores. IPF-specific PROMs are now emerging that conform to FDA criteria [53].

The two instruments used most extensively to measure health status in these conditions in outcome studies of ILDs are: SF-36 [57] & SGRQ [56,61]. The SF-36 is a recognised measure for use in cost benefit analyses, but as a generic tool its clinical utility is limited. In IPF studies, the greatest impairments are observed in physical health domains of the SF 36 (except bodily pain) [62].

The SGRQ is reported to have acceptable validity and reliability in ILD [59,63]. Longitudinal data utilising the SGRQ-original in a cohort of patients with IPF indicates that the SGRQ is an independent predictor of disease progression [58]. The original SGRQ is also reported to be responsive in a CTD-ILD cohort [60]. The SGRQ has been subjected to item reduction utilising Rasch analysis in a cohort of patients to produce the SGRQ-ILD [61]. The data set was obtained from the BUILD-1 trial in which the SGRQ was administered at baseline, month 6 and month 12. Baseline and 6 month data were included in the analysis. The SGRQ-ILD has 34 items divided into Symptoms, Activity, and Impact domains. Data supporting the responsiveness of the SGRQ-ILD are needed.

One disease-specific HRQL tool (which was not available in time to be included in the ILD consensus process) has been developed to assess HRQoL in IPF (A Tool to Assess Quality of Life in IPF, ATAQ-IPF) [64]. However, it is lengthy (86 items version 1 and 89 items version 2) and thus cumbersome for routine use or even for use in clinical trials. The ATAQ-IPF has been subjected to an item reduction process utilising RASCH analysis in a mixed US / UK population resulting in the shorter 43 item ATAQ-IPF-CA [65]. Limitations to this study are acknowledged by the authors particularly regarding the absence of pulmonary physiology values in 47% of the subjects.

The K-BILD (The King's Brief ILD Health Status Questionnaire, not available at time of consensus effort), is a 15-item HRQoL questionnaire comprised of three health domains (psychological, breathlessness and activities associated with chest symptoms) with an overall HRQoL score [66]. The K-BILD may ultimately prove to serve as both a dyspnea

and HRQoL. The conceptual framework that informed the development of the K-BILD was developed by a professional multidisciplinary team and generated by 10 one-to-one interviews with patients diagnosed with a range of different ILDs. The questionnaire was tested across a predominantly female patient population (60%) with a mean age of 60 (± 13). The SGRQ-ILD, K-BILD and the ATAQ all require longitudinal evaluation in IPF and CTD-ILD populations.

Given the lack of data to support the longitudinal validity of HRQoL PROMs, the choice of candidate instruments is currently based on what is acceptable for use at this time. There is an urgent need for international collaboration to address this specific area of need in IPF and ILD assessment. A National Institute of Health Research (NIHR) funded study in the UK is developing an IPF-PROM according to FDA criteria. Validation of this instrument will commence in 2015.

Lastly, the Visual Analogue Scale of Patient Global Assessment of Disease Activity (VAS-PtGA) is a 0 – 100 mm VAS scale anchored by “the best I’ve ever felt” and “worst I’ve ever felt” upon which the patient places a mark in response to the question, ‘Considering all the ways your condition affects you, how are you doing today?’ The purpose may not be granular enough to differentiating changes in lung disease status, but rather allows the overall impact of all processes and interventions as perceived by the patient to be gauged. Although there is a paucity of literature examining patient global assessments in ILDs, The VAS-PtGA has been widely validated across many diseases and in most CTDs [67]; and is a standard part of assessment in the Scleroderma Health Assessment Questionnaire (SHAQ). For these reasons, the consensus effort supported its immediate use in trials for CTD-ILD and defined an MCID of 10mm. In regards to IPF, this instrument requires further research on its performance within the context of randomized controlled trials and longitudinal observational studies prior to its adoption.

Recognised challenges of symptom and HRQoL measures as primary endpoints include controlling for the common use of additional therapies that palliate concurrent morbidities such as pain, reflux and chronic infection that influence these HRQoL, dyspnea and cough favourably without exerting effect on ILD progression. These measures however may be easier to justify as primary endpoints in trials of palliative agents rather than targeted anti-fibrotic therapy; and when highly correlative with the primary endpoint in trials of targeted therapy strengthens the reliability of that primary endpoint.

LUNG IMAGING

High-resolution computed tomography (HRCT) is an essential tool in the diagnosis and management of patients with ILDs. In the context of clinical trials, HRCT provides information on diagnosis, prognostication, the likelihood of therapeutic responsiveness and serial change of disease progression [68–88]. HRCT employs a combination of thin collimation (1 to 2 mm slices) and a high spatial reconstruction algorithm to provide optimal assessment of the lung parenchyma. The higher spatial resolution allows for better definition of the pattern, extent, and distribution and an ‘in vivo’ perspective of parenchymal

abnormalities and changes as compared to chest radiography; with the ability to identify evidence of ILD in patients whose chest radiographs appear normal.

There are several radiologic features of UIP and NSIP patterns on HRCT but the most relevant of these in fibrosing IIPs are: 1. Ground glass opacities (GGOs), hazy nonspecific expanses of increased attenuation that in IIPs may represent inflammation or fine fibrosis; 2. Reticular opacities, also a non-specific finding suggesting a diminished gas to parenchyma ratio, in fibrosing IIPs, these are highly suggestive of fibrosis and accompanied by linear septal fibrosis; 3. Honeycombing, well-defined cystic changes of the lung parenchyma that are associated with advanced fibrotic and irreversible disease but can be misinterpreted as emphysema; and 4. Traction bronchiectasis, a consequential but important characteristic of fibrosis in which distortion of the bronchi occur secondary to the tension exerted by areas of fibrosis [81,85]. The first three tend to occur in a peripheral and basilar distribution -and agreement in distinction of these characteristics are challenging even among [78,89] experts. UIP pattern characteristically has minimal GGO and early stages with predominance of reticulation with increasing honeycombing in later stages [81,85]. NSIP may predominate with GGOs and reticulation with subpleural honeycombing being less frequent. It is important to clarify that recently published guidelines (1) advocate that an HRCT pattern absent of honeycombing but supported by less typical features of IPF warrant biopsy for diagnosis [1]. Such less typical HRCT patterns in IPF may impact the utility of HRCT as an outcome measure in IPF. Diagnostic uncertainty of etiology of an NSIP or UIP pattern, depending on the clinical context, may warrant need for a lung biopsy.

Progression in fibrosing IIPs are associated with increasing extent of radiologic abnormalities and decreasing lung volume. With worsening disease, there is an increase and coarsening of reticular opacities which may eventually be replaced by honeycombing, while the extent of GGO may decrease [68,76,77,90]. Thus, the serial HRCT parameter that correlates best with disease progression is the overall extent of these combined parenchymal abnormalities rather any single radiologic textural components [69,71,73,88].

HRCT scoring may be approached by several methods that range from complex to simple and by visual assessment or computer-aided. Most commonly, HRCT abnormalities are quantified by radiologists' visual scoring, this may be a simple three or five-point scale or more typically a percentage to the nearest 5% of involved parenchyma which may evaluate total volume or selected cross-sections. This quantitation can be confounded by intra- and inter-observer variability - and therefore has infrequently been included as a trial primary endpoint. Software tools provide computer-aided quantitative scoring either for global assessment or for specially selected parenchymal lung abnormalities (i.e. textural -one or more of the patterns above) and/or regions. Quantitative scoring provides objectivity, reproducibility [72,74,79,82] and perhaps a greater reliability and sensitivity in quantifying treatment effect over shorter time intervals than visual scoring methods [79]. The optimal time interval between HRCT evaluations is yet unknown.

HRCT is valuable as assessed at a single point in time in clinical trials. HRCT provides essential diagnostic inclusion criteria for fibrosing IIP studies [81,85]. CT fibrosis score portends ability to prognosticate and predict mortality in SSc and IPF

[69,70,71,78,84,87,88,94]; but statistically loses significance when adjusted for lung function suggesting need for either finer techniques for fibrosis quantification or continued efforts in the development of an index comprised of multiple variables (lung, function, HRCT, demographics, co-morbidities) [91–94]. Further, HRCT patterns and scoring have demonstrated great potential in identifying patients likely to respond to an intervention. For example, commonly patients with SSc have minimal non-advancing ILD, and unlikely to respond to a treatment intervention. This was demonstrated in a comparative sub-analysis of patients with a baseline overall extent of disease >25% by visual fibrosis scoring had a greater reduction of FVC over one year [10]. Further, the extent or presence of honeycombing on baseline HRCT in UIP versus UIP-without-honeycombing may predict mortality and the likelihood of treatment response; with consideration that IPF cohort selections enroll patients with UIP pattern but without HRCT feature of honeycombing or use this strategy as a sub-stratification tool [77,95].

Serial assessment of HRCT increasingly holds promise as an endpoint to quantify disease progression. Visual scoring [95] suggests correlation in NSIP and UP; but there is uncertain correlation between overall extent of ILD on HRCT and serial FVCs [95–97]. Quantitative (computer-aided) fibrosis scoring on serial CT, primarily through its ability to numerically quantify subtle serial changes of fibrotic intensity within reticular patterns not easily distinguished by the human eye, provides advanced ability to identify and quantify a treatment response with serial numerical quantification in both cases and cohorts. The SLS trial data analysed by serial computer-aided fibrosis quantification demonstrated a clear difference between treatment arms [75]. Further to serial assessment of treatment response, short-term quantitative fibrosis scoring may provide a level of sensitivity of change at short intervals to be a tool to optimize cohort selection [79].

As yet, however, it is not clear that radiologic evaluation identifies progression that is not evident by other tools, such as pulmonary function testing, that do not pose radiation risks. Furthermore, continuing efforts at reducing the radiation dose involved in CT examinations will make serial HRCT assessment an accurate biomarker used in clinical trials [73,98].

Overall extent of parenchymal disease on HRCT was identified by the consensus effort as the currently most suitable radiologic endpoint with reservations related to slow rate or absence of change despite significant changes in other variables such as FVC [95–97]. Given continued advancement in technology and need for validated studies, HRCT scoring method was intentionally not selected.

LUNG PHYSIOLOGY AND FUNCTION

Interstitial lung disease, both in IPF and CTD-associated lung disease, is characterized by decreases in forced vital capacity (FVC) and diffusion capacity of lung for carbon monoxide (DLCO), as well as impaired exercise tolerance (as measured by the distance walked in six minutes [6MWD]) and exercise-induced oxygen desaturation. The majority of information available about using lung function and physiology as a clinical endpoint in CTD-ILD is derived from patients with SSc. Data from studies including patients with IPF can be carefully extrapolated to patients with CTD-ILD, although there are notable differences that

can limit this interpretation. Measures of lung physiology and function can be affected by comorbid conditions such as pulmonary vascular disease, airways disease, musculoskeletal disease, neuromuscular weakness and reduced/abnormal oral aperture [99].

Interpretation of pulmonary function is usually based on comparisons of data measured in an individual patient or subject with reference (predicted) values based on healthy subjects. These predicted values should be obtained from studies of “normal” or “healthy” subjects with the same anthropometric and, where relevant, ethnic characteristics of the patient being tested.

Evaluation of an individual’s change in lung function following an intervention or over time is even more clinically valuable than a single comparison with external reference (predicted) values. However, it is not always easy to determine whether a measured change reflects a true change in pulmonary status or is only a result of test variability or is statistically versus clinically significant. All lung function measurements tend to be more variable when made weeks to months apart than when repeated at the same test session or even daily. Significant changes from year to year are 10% for lung volume measures and 15% for DLCO. The guidelines do not provide guidance on significant change in ILD although a decline in FVC of greater than 10% has been associated with a higher mortality [81,100]; also the change which protects against variability.

Forced Vital Capacity

The Forced Vital Capacity (FVC) is the volume delivered during expiration made as forcefully and completely as possible starting from full inspiration. This value may be influenced by external restriction (e.g. vertebral compression fracture), neuromuscular weakness or by air-trapping. Serial change from baseline FVC was the measure most favored as an outcome for both CTD-ILD and IPF, with 100% acceptance among the consensus panelists [9]. FVC has been the most commonly used outcome in IPF and CTD-ILD studies, including the recently published IPF studies investigating pirfenidone and nintedanib which demonstrated major reduction in disease progression; as a result these two drugs are now FDA approved [14,15,101]. In a retrospective analysis of a 330 patient study with interferon- γ 1b for IPF, mortality (mean follow-up 58 weeks) increased 2.4-fold in those with the protocol-defined 10% decrease in percent predicted FVC [102]. Another retrospective analysis of 1,156 patients sought to define the MCID of FVC change in IPF, and found it to be between 2 and 6% of baseline [103].

Although used as an endpoint in CTD-ILD, the MCID for change in FVC has not been well defined. In the large Scleroderma Lung Study (SLS), change in FVC was the primary endpoint comparing oral cyclophosphamide to placebo in SSc-ILD [10]. They found a 2.53% difference in FVC change, which was statistically significant. In an analysis of the placebo group in the SLS, the mean decline in FVC over a one year period was $4.2 \pm 12.8\%$ of predicted [29]. In this same study, <15% of patients experienced a 10% decrease from baseline in percent predicted FVC [99]. Though favored, FVC is flawed in several aspects. Finding a large change in FVC with therapy may be difficult, since most of the lung function decline occurs in the first 4–6 years of disease in SSc [104]; making stabilization or lack of worsening a potential alternate paradigm of treatment efficacy. Further, prospective change

in FVC used as a surrogate endpoint has not been shown to be associated with survival over 1 or 3 years [105].

Both primary and secondary endpoints should be reliable, reproducible, sensitive to change over time and feasible in administration, safety and cost. A *primary* endpoint is chosen for its high discrimination and most direct reflection of disease progression with evident correlation to secondary endpoints notably HRQoL, symptom and mortality. This leaves serial assessment of FVC as the most favoured clinical endpoint [10,14,15]. Strategies to strengthen the sensitivity of serial FVC incorporate FVC into a physiologic composite measure with another variable such as a 50 metre decline in 6MWD [14]; with theoretical considerations of future composite measures integrating serial change of HRCT with serial FVC.

Diffusion Capacity

Again, FVC was favored over serial DLCO in both CTD-ILD and IPF by the consensus effort, which had an 87% acceptance rate as an outcome for CTD-ILD [12]. In neither entity, is DLCO specific to the interstitium and at any point in the disease course may be a composite of both interstitial and vascular change.

The ability of DLCO to serve as a surrogate marker in IPF is controversial. In the analysis of the interferon- γ 1b study, change in DLCO did not correlate with mortality [102], although another retrospective study found that change in DLCO was predictive of survival [106]. Even less is known about DLCO as an outcome in CTD-ILD whereby concomitant or preceding vasculopathy exists in variable degrees with inability to accurately isolate vascular contribution, although we expect the FVC:DLCO ratio to remain closer to one in pure ILD [107]. In the SLS, there was no difference in DLCO at 1 year between treatment groups [10]. In another study, change in DLCO at 3 years, but not 1 year, correlated to survival rates in SSc-ILD [105].

Exercise Testing

Six-Minute Walk Distance—6MWD has been used as a clinical outcome measure in COPD and idiopathic pulmonary hypertension; its value in ILD is less certain. It has been found to be reproducible in both IPF [108] and SSc-ILD [9,10]. In data derived from an interferon- γ 1b study, the risk of death was 4-fold higher for patients with a decline in 6MWD of >50m, compared to those with a decline of <25m [109]. The MCID for hospitalization or death in IPF was estimated to be 24m. The minimal clinically important difference in CTD-ILD has not been defined.

In a retrospective study of 48 IPF patients, 33 patients with SSc-ILD and 19 patients with SSc and both ILD and PH, limitations to six minute walk trended toward dyspnea in IPF and lower extremity pain in scleroderma [110]. The percentage predicted FVC and DLCO were more strongly predictive of 6MWD in IPF than in SSc. Correlates of 6MWD differed between SSc and IPF, therefore, the 6MWD is not always reflective of the same physiologic process and should include consideration of vascular, pulmonary, and musculoskeletal exercise limitations.

Exercise-Induced Oxygen Desaturation—Although measures of oxygen desaturation may suffer from poor reproducibility in IPF [108], even mild desaturation has been found to be strongly associated with mortality [109] with a >4-fold increase in mortality with desaturation to <88% on a 6MWD [110]. Exercise oxygen measurement in CTD-ILD, especially SSc, is perceived as problematic due to diminished signal integrity related Raynaud's. However, earlobe and forehead oximetry may produce reliable oxygen values in SSc patients [8,112]. Shorter survival was associated with desaturation to <89%, or a >4% decrease in saturation with cardiopulmonary exercise in SSc-ILD [112]. Despite associated mortality, declines in saturation are not linearly correlated with other measures such as dyspnea or FVC and thus not able to provide a reliable serial metric for IPF and CTD-ILD trials.

In conclusion, though no pulmonary function or physiology endpoint has been validated as a surrogate outcome for mortality in CTD-ILD nor, arguably for IPF [113]; serial FVC rises to most promising over all other current endpoints with potential as both a mortality surrogate and a marker of disease progression. Remaining questions include (a) what is the appropriate duration between serial measurements, (b) should absolute or percent predicted changes be quantified, and (c) should variables be treated as being continuous or as proportion of patients meeting a threshold change [99].

MORTALITY

Even as overall mortality rates for many of the CTDs has declined, pulmonary involvement in CTD is now one of the leading causes of morbidity and mortality [114–116]. Thus, mortality is instinctively the endpoint that seems ideal representing the most rigorous and clinically meaningful endpoint for both CTD-ILD and IPF [113]. A controversial statement published in 2012 (prior to recent studies demonstrating efficacy [14,15]) regarding IPF clinical trials argued that all-cause mortality was the single most important clinically meaningful outcome. This opinion had been based primarily on mortality signals from the IPF-NET trials [117,118] and shortcomings associated with any other endpoint. Specifically, the authors concluded that any endpoint other than mortality used to measure treatment effect in a clinical trial could not be considered a reliable predictor of a clinically meaningful outcome; including FVC, with some patients meeting a mortality endpoint prior to meeting change in FVC endpoint [102].

Despite these assertions, adopting all-cause mortality as the only meaningful endpoint in clinical trials of IPF has not been universally endorsed [119]. Utilization of mortality as the primary endpoint is largely considered infeasible with requirements for prolonged trials, large numbers of subjects with associated high costs [100]. Further, two recent trials [14,15] definitively substantiated earlier studies [120,121] indicating that serial change in FVC can capture major reduction in disease activity, thus strongly challenging mortality as the incomparable outcome measure - except perhaps in end-stage disease. This endpoint would also be viewed as unethical for placebo controlled trials with effective treatments being available. Moreover, mortality in IPF trials spans a wide range, perhaps related to the number of patients enrolled and the study duration; with placebo arms of two separate trials enrolling physiologically similar subjects spanning 2–17% [102,118]. The latest consensus

guidelines on the diagnosis and management of IPF support a 10% change in absolute FVC (due to its lack of variability at this magnitude) may serve as a surrogate for mortality and is evidence of disease progression [81].

Similar concepts in the use of mortality in IPF translate to consideration of use of mortality in the CTD-ILDs. The consensus effort recognized mortality as an essential endpoint in all treatment trials and all-cause mortality was identified as a valid measure of survival in CTD-ILD and IPF. However, despite the indisputable value of mortality, similar debate surrounds this as the only valuable endpoint. FVC may be a valid surrogate marker for mortality, but this is unknown in CTD-ILD. In general, it appears that declines in FVC correlate with increased subsequent mortality, particularly in SSc [9,122], but the direct relationship to time has not been determined and, again, the duration to this endpoint being a clinically meaningful surrogate may also confound trial feasibility.

Clinical trials in CTD-ILD have not historically utilized mortality as an endpoint, this is likely due to similar concerns with IPF surrounding the impractical nature of this endpoint. Notably, mortality was minimal or absent in recent clinical trials of SSc-ILD [10,123]. These low numbers portend too low an event rate to perform a clinical trial in any CTD-ILD that can be feasibly powered for mortality.

Other surrogates for mortality that are suggestive of time to clinical worsening include time to transplantation and pulmonary-specific hospitalization admission as well as a proposed definition of acute exacerbation (124). However, each of these are fraught with confounding elements (113) that are especially pronounced in CTD-ILD such as access to transplant services or infectious pulmonary-related hospital admission caused by immune suppression or by inherent immune dysfunction of CTDs rather than true exacerbation. In IPF, acute exacerbation appears to correlate to mortality, functional status and symptomatology. A standard definition is likely to make acute exacerbation a more reliable measure and predictor of mortality (124).

At the present time, mortality, while indisputably the cleanest and most rigorous endpoint for clinical trials in CTD-ILD, is not practicable clinical trial endpoint. It seems more likely that we will need to better understand how physiologic or CT scan measures can serve as meaningful surrogates for ultimate mortality.

COMPOSITE MEASURES

Composite measures can be developed as prognostication or severity stratification tools as well as assess disease progression over time. The premise of a composite measure is that they incorporate reliable but *different* information from variables that have demonstrated reproducibility, discrimination and sensitivity to change (for serial measures). The end goal is that the combination of these variables provide a sharper assessment tool that is more sensitive to change than any variable alone. Another goal is to avoid dependency on a single domain. In IIPs, these candidate measures have been variable combinations of HRCT to provide radiologic information and FVC, DLCO and 6MWT to provide physiologic information as well as clinical information such as age, gender, respiratory-related hospitalization [70,92,125–127].

Composite measures can be with variable or equal weighting. Equal weighting is considered preferable. However composite endpoints are as strong as their most flawed element, therefore it may be necessary to weight elements differently. Another strategy is to employ variables in a sequential or contingent manner such as that described by Goh et al that follows sequential steps a HRCT fibrosis score of >10% proceeded to inclusion of FVC for further stratification or such as has been proposed for pulmonary sarcoidosis whereby obtaining chest radiograph is contingent upon an indeterminate magnitude of change in FVC [71,128]. Though the consensus effort recognized that composite measures have promise in measuring disease progression in the fibrosing IIPs, it was determined during the early stages of the effort that more work is needed in this area prior to utilization [12].

OUTCOME MEASURES FROM PATIENTS' POINT OF VIEW

The inclusion of patients as advisors in the design of research endeavours is increasingly recognized as valuable. Patient advisors can provide important information on feasibility issues related to patients that may impact cohort retention.

The consensus effort included patients at several junctures of design, data collection and as panelists in the final recommendations meeting [12]. Prior to the panel meeting, patient panelists and other patient advisors with either CTD-ILD or IPF attended a series of patient teleconference sessions (led by LAS)- unpublished data. The purpose of the pre-conference sessions was to prepare the patient panelists to participate in the panel debates with high level of confidence. To this end, education on the scientific relevance of the candidate domains and instruments was provided, and patient perceptions and experiences of the instruments were queried and explored in this forum.

Universally, there were patient concerns expressed regarding FVC. The most striking of which was the expression of dread mostly due to the incitation of cardiopulmonary debilitation with cough, dyspnea and tachycardia that was perceived as persisting for hours after the test. Some patients voiced fear that they were hurting themselves. Participants voiced being anxious because FVC testing was dependent on their effort and performance with the concern of spurious results on a 'bad day' that would effect physician treatment choices or prognostication. This was not true of HRCT. Patients were unanimously more comfortable with the prospect of serial HRCT despite radiation exposure.

Regarding PROMs, symptoms and disease impact on HRQoL were central to patients. Discussion of PROMs surrounded the relevance of the content and implementation. Discussion on content echoed the final panel results that ILD-specific may more accurately reflect the experiences of patients with restrictive lung diseases. In regard to mortality, 'survival' was the preferred term. Discussions supported capturing all-cause death as respiratory-related death was voiced as being of interest to participants' perceived efficacy of the drug; and capturing a drug-related harm signal was perceived as essential to participants. Lastly, in discussion of mortality as a primary endpoint, participants voiced overall concern about life-long allegiance to a trial drug with uncertain efficacy (or the comparative agent) as other novel drugs become available.

DISCUSSION

In both CTD-ILDs and IPF, new and novel therapeutics that target fibrosing disease are increasingly available for clinical trials. In this era of accelerated research and promising interventions, a minimum set of outcome measures has been proposed to encourage consistency across trials as well as the utilization of instruments bearing the highest yield of gauging efficacy. While for each IPF and CTD-ILD, there is no single instrument that confers consistent reliability and discrimination, absolute change in FVC may be the least flawed and most feasible of available objective measures. Extent of overall disease on HRCT is useful but current methods of assessment commonly used may be slow to describe discriminatory changes; the future of computer-aided scoring may yield finer measurements allowing for higher sensitivity to change.

Though mortality may ultimately be the most robust measure, powering a trial for mortality as a stand-alone primary endpoint requires trial duration over years, creating a cost prohibitive design and is virtually infeasible. However, all-cause mortality in combination with other measures over shorter trial duration may provide important efficacy data as well as, very importantly, a harm signal related to the therapeutic intervention. Changes in FVC, as a potential surrogate for mortality or conversely progression free survival is being explored but currently supported by inconsistent data.

Perhaps in diseases such as IPF, for which medicinal therapy only slows the decline in lung function (14,15), a disease-specific HRQoL, dyspnea measure and VAS-PG may be the most patient-centered approaches when considering the delicate paradigm of life duration in maintenance of QoL.

The preferred primary end-point would be a measure that matters to patients - such mortality, symptoms, or QoL. However, mortality as a marker of efficacy is not practicable and symptoms (dyspnea, cough) and QoL is not yet reliably measurable. The standard inclusion of patient-reported dyspnea and cough as a secondary outcome measure in future trials may yield useful correlative data and provide insight into both natural history and further instrument identification, utilization and development. However, these instruments should be developed with patients to accurately reflect patient priority content and language related to restrictive lung disease.

ACKNOWLEDGEMENTS

The editorial staff recommends the following link as useful for readers: <http://erj.ersjournals.com/content/26/5/948.long>

Biography

Lesley Ann Saketkoo

REFERENCES

1. Travis WD, Costabel U, Hansell DM, et al. ATS/ERS Committee on Idiopathic Interstitial Pneumonias. An official American Thoracic Society/European Respiratory Society statement: Update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. *Am J Respir Crit Care Med*. 2013 Sep 15; 188(6):733–748. [PubMed: 24032382]
2. Moore CF Jr, Perez RL. Comparison of Idiopathic Interstitial Pneumonias with Their Connective Tissue Disease-Associated Counterparts. *Curr Resp Med Rev*. 2015; 11(2) in press.
3. Jokerst C, Azok J, Cummings KW, Bhalla S. High-Resolution CT Imaging Findings of Collagen Vascular Disease-Associated Interstitial Lung Disease. *Curr Resp Med Rev*. 2015; 11(2) in press.
4. American Thoracic Society/European Respiratory Society Statement on Interstitial Pulmonary Autoimmune Fibrosis. American Thoracic Society/European Respiratory Society Task Force—"An International Working Group on Undifferentiated Forms of CTD-ILD". Aryeh Fischer, Chair. In press.
5. Solomon JJ, Ryu JH, Tazelaar HD, et al. Fibrosing interstitial pneumonia predicts survival in patients with rheumatoid arthritis-associated interstitial lung disease (RA-ILD). *Respir Med*. 2013 Aug; 107(8):1247–1252. [PubMed: 23791462]
6. Song JW, Lee HK, Lee CK. Clinical course and outcome of rheumatoid arthritis-related usual interstitial pneumonia. *Sarcoidosis Vasc Diffuse Lung Dis*. 2013 Aug 1; 30(2):103–112. [PubMed: 24071881]
7. 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]
8. 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]
9. Roth MD, Tseng CH, 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]
10. 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]
11. 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 (Hoboken)*. 2011; 63(11):1620–1628. [PubMed: 22034123]
12. 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]
13. Saketkoo LA, Mittoo SM, 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]
14. 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]
15. 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]
16. Borg G. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc*. 1982; 14:377–381. [PubMed: 7154893]
17. Medical Research Council. Committee on research on chronic bronchitis. Instructions for use of the questionnaire on respiratory symptoms. Devon, UK: W.J. Holdman; 1966.

18. Mahler DA, Ward J, Fierro-Carrion G, et al. Development of self-administered versions of modified baseline and transition dyspnea indexes in COPD. *COPD*. 2004; 1(2):165–172. [PubMed: 17136984]
19. Eakin, EG.; Resnikoff, PM.; Prewitt, LM.; Ries, AL.; Kaplan, RM. *Chest*. Vol. 113. San Diego: University of California; 1998. Validation of a new dyspnea measure: the UCSD Shortness of Breath Questionnaire; p. 619–624.
20. Yorke J, Moosavi SH, Shuldham C, Jones PW. Quantification of dyspnoea using descriptors: development and initial testing of the Dyspnoea-12. *Thorax*. 2010; 65(1):21–26. [PubMed: 19996336]
21. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med*. 2002; 166(1): 111–117. [PubMed: 12091180]
22. Wilson RC, Jones PW. A comparison of the visual analogue scale and modified Borg scale for the measurement of dyspnoea during exercise. *Clin Sci (Lond)*. 1989; 76(3):277–282. [PubMed: 2924519]
23. Wilson RC, Jones PW. Long-term reproducibility of Borg scale estimates of breathlessness during exercise. *Clin Sci (Lond)*. 1991; 80(4):309–312. [PubMed: 1851065]
24. 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]
25. Manali ED, Lyberopoulos P, Triantafyllidou 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]
26. Yorke J, Swigris J, Russell AM, 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]
27. Holland AE, Fiore JF Jr, Bell EC, et al. Dyspnoea and comorbidity contribute to anxiety and depression in interstitial lung disease. *Respirology*. 2014 Nov; 19(8):1215–1221. [PubMed: 25112470]
28. Mahler DA, Weinberg DH, Wells CK, Feinstein AR. The measurement of dyspnea Contents, interobserver agreement, and physiologic correlates of two new clinical indexes. *Chest*. 1984; 85(6):751–758. [PubMed: 6723384]
29. Khanna D, Tseng CH, Furst DE, et al. Minimally important differences in the Mahler's Transition Dyspnoea Index in a large randomized controlled trial--results from the Scleroderma Lung Study. *Rheumatology (Oxford)*. 2009; 48(12):1537–1540. [PubMed: 19776222]
30. 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]
31. 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]
32. Raj AA, Birring SS. Clinical assessment of chronic cough severity. *Pulm Pharmacol Ther*. 2007; 20(4):334–337. [PubMed: 17113330]
33. Brignall K, Jayaraman B, Birring SS. Quality of life and psychosocial aspects of cough. *Lung*. 2008; 186(Suppl 1):S55–S58. [PubMed: 17939003]
34. Raj, AA.; Pavord, ID.; Birring, SS. What is the minimal important difference of the Leicester Cough Questionnaire?. In: Chung, KF.; Widdicombe, J., editors. *Pharmacology and Therapeutics of Cough*. Berlin: Springer; 2009. p. 311–320.
35. Ryerson CJ, Abbritti M, Ley B, Elicker BM, Jones KD, Collard HR. Cough predicts prognosis in idiopathic pulmonary fibrosis. *Respirology*. 2011 Aug; 16(6):969–975. [PubMed: 21615619]
36. Theodore AC, Tseng CH, Li N, Elashoff RM, Tashkin DP. Correlation of cough with disease activity and treatment with cyclophosphamide in scleroderma interstitial lung disease: findings from the Scleroderma Lung Study. *Chest*. 2012 Sep; 142(3):614–621. [PubMed: 22156609]
37. Birring SS, Parker D, McKenna S, et al. Sputum eosinophilia in idiopathic pulmonary fibrosis. *Inflamm Res*. 2005 Feb; 54(2):51–56. [PubMed: 15750711]

38. Birring SS. Developing antitussives: the ideal clinical trial. *Pulm Pharmacol Ther.* 2009 Apr; 22(2):155–158. [PubMed: 19041729]
39. Birring SS, Parker D, Brightling CE, et al. Induced sputum inflammatory mediator concentrations in chronic cough. *Am J Respir Crit Care Med.* 2004 Jan 1; 169(1):15–19. [PubMed: 14512269]
40. Birring SS, Prudon B, Carr AJ, Singh SJ, Morgan MD, Pavord ID. Development of a symptom specific health status measure for patients with chronic cough: Leicester Cough Questionnaire (LCQ). *Thorax.* 2003 Apr; 58(4):339–343. [PubMed: 12668799]
41. French CT, Irwin RS, Fletcher KE, Adams TM. Evaluation of a cough-specific quality-of-life questionnaire. *Chest.* 2002 Apr; 121(4):1123–1131. [PubMed: 11948042]
42. Davenport PW. Clinical cough I: the urge-to-cough: a respiratory sensation. *Handb Exp Pharmacol.* 2009; 187:263–276. [PubMed: 18825345]
43. Birring SS, Matos S, Patel RB, Prudon B, Evans DH, Pavord ID. Cough frequency, cough sensitivity and health status in patients with chronic cough. *Respir Med.* 2006 Jun; 100(6):1105–1109. [PubMed: 16266801]
44. Birring SS, Fleming T, Matos S, Raj AA, Evans DH, Pavord ID. The Leicester Cough Monitor: preliminary validation of an automated cough detection system in chronic cough. *Eur Respir J.* 2008 May; 31(5):1013–1018. [PubMed: 18184683]
45. Barton A, Gaydecki P, Holt K, Smith JA. Data reduction for cough studies using distribution of audio frequency content. *Cough.* 2012; 8(1):12. [PubMed: 23231789]
46. Prudon B, Birring SS, Vara DD, Hall AP, Thompson JP, Pavord ID. Cough and glottic-stop reflex sensitivity in health and disease. *Chest.* 2005 Feb; 127(2):550–557. [PubMed: 15705995]
47. Huisman AN, Wu MZ, Uil SM, van den Berg JW. Reliability and validity of a Dutch version of the Leicester Cough Questionnaire. *Cough.* 2007; 3:3. [PubMed: 17313670]
48. Key AL, Holt K, Hamilton A, Smith JA, Earis JE. Objective cough frequency in Idiopathic Pulmonary Fibrosis. *Cough.* 2010; 6:4. [PubMed: 20565979]
49. Lechtzin N, Hilliard ME, Horton MR. Validation of the Cough Quality-of-Life Questionnaire in patients with idiopathic pulmonary fibrosis. *Chest.* 2013 Jun; 143(6):1745–1749. [PubMed: 23519393]
50. Fletcher KE, French CT, Irwin RS, Corapi KM, Norman GR. A prospective global measure, the Punum Ladder, provides more valid assessments of quality of life than a retrospective transition measure. *J Clin Epidemiol.* 2010 Oct; 63(10):1123–1131. [PubMed: 20303709]
51. 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 Sep 18; 157(6):398–406. [PubMed: 22986377]
52. Swigris JJ, Fairclough D. Patient-Reported Outcomes in Idiopathic Pulmonary Fibrosis Research. *Chest.* 2012; 142(2):291–297. [PubMed: 22871754]
53. Food and Drug Administration, Center for Drug Evaluation and Research. Guidance for Industry: Qualification Process for Drug Development Tools Silver. Spring, MD: US Department of Health and Human Services; 2010.
54. Aaronson N, Choucair A, Elliot T, Greenhalgh J, et al. User's guide to Implementing Patient Reported Outcome Measures Assessment in clinical Practice International society for Quality of Life Research. 2011 Nov.
55. Cochrane Handbook for Systematic Reviews of Interventions 5.1.0. 2011 Mar.
56. 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; 104:296–304. [PubMed: 19815403]
57. Ware JE, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): Conceptual framework and item selection. *Medical Care.* 1992; 30:473–483. [PubMed: 1593914]
58. Russell AM, Fraser U, Molyneaux P, et al. Quality of life measures in patients with idiopathic pulmonary fibrosis. *Eur Resp J.* 2013; 42(Suppl 57):P3368.
59. Swigris J, Gould MK, Wilson SR. Health-Related Quality of Life Among Patients With Idiopathic Pulmonary Fibrosis. *Chest.* 2005; 127:284–294. [PubMed: 15653996]

60. Patel AS, Brignall K, Gordon P, Steer S, Wells AU, Higginson I, Birring S. Assessment Of Health Related Quality Of Life (HRQOL) In Connective Tissue Disease Associated Interstitial Lung Diseases (CTD-ILD). *Am J Respir Crit Care Med*. 2011; 183:A3030.
61. Yorke J, Jones PW, Swigris JJ. Development and Validity Testing of an IPF-Specific Version of the SGRQ Thorax. 2010; 65 921e926.
62. Martinez J, Martinez T, Galhardo F, et al. Dyspnea scales as a measure of health-related quality of life in patients with idiopathic pulmonary fibrosis. *Med Sci Monit*. 2002; 8:CR405–CR410. [PubMed: 12070430]
63. Chang J, Curtis R, Patrick D, et al. Assessment of Health Related quality of life in patients with interstitial lung disease. *Chest*. 1999; 116:1175–1182. [PubMed: 10559073]
64. Swigris J, Wilson S, Green K, Spunger D, Brown K, Wambolt F. Development of the ATAQ-IPF: a tool to assess quality of life in IPF Health and Quality of Life Outcomes. 2010; 8:77. [PubMed: 20673370]
65. Yorke J, Spencer LG, Duck A, et al. Cross-Atlantic modification and validation of the A Tool to Assess Quality of Life in Idiopathic Pulmonary Fibrosis (ATAQ-IPF-cA) *BMJ Open Res*. 2014; 1:24.
66. 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; 67:804–810.
67. Steen VD, Medsger TA. The value of the Health Assessment Questionnaire and special patient-generated scales to demonstrate change in systemic sclerosis patients over time. *Arthritis Rheum*. 1997; 40:1984–1991. [PubMed: 9365087]
68. Akira M, Kozuka T, Inoue Y, Sakatani M. Long-term follow-up CT scan evaluation in patients with pulmonary sarcoidosis. *Chest*. 2005; 127:185–191. [PubMed: 15653982]
69. Best AC, Meng J, Lynch AM, et al. Idiopathic Pulmonary Fibrosis: Physiologic Tests, Quantitative CT Indexes, and CT Visual Scores as Predictors of Mortality. *Radiology*. 2008; 246:935–940. [PubMed: 18235106]
70. Edey AJ, Devaraj AA, Barker RP, Nicholson AG, Wells AU, Hansell DM. Fibrotic idiopathic interstitial pneumonias: HRCT findings that predict mortality. *Eur Radiol*. 2011; 21:1586–1593. [PubMed: 21380740]
71. Goh NS, Desai SR, Veeraraghavan S, et al. Interstitial lung disease in systemic sclerosis: a simple staging system. *Am J Respir Crit Care Med*. 2008; 177(11):1248–1254. [PubMed: 18369202]
72. Goldin JG. Quantitative CT of the lung. *Radiol Clin North Am*. 2002; 40(1):145–162. [PubMed: 11813816]
73. Kauczor H-U, Heussel CP, Herth FJ. . Longitudinal quantitative low-dose CT in COPD: ready for use? *Lancet Respir Med*. 2013; 1(2):95–96. [PubMed: 24429077]
74. Kazantzzi A, Costaridou L, Skiadopoulou S, et al. Automated 3D Interstitial Lung Disease Extent Quantification: Performance Evaluation and Correlation to PFTs. *J Digit Imaging*. 2014; 27(3): 380–391. [PubMed: 24448918]
75. Kim HJ, Brown MS, Elashoff R, et al. Quantitative texture-based assessment of one year changes in fibrotic reticular patterns on HRCT in scleroderma lung disease treated with oral cyclophosphamide. *Eur Radiol*. 2011; 21:2455–2465. [PubMed: 21927793]
76. Kim MY, Song JW, Do KH, Jang SJ, Colby TV, Kim DS. Idiopathic nonspecific interstitial pneumonia: changes in high-resolution computed tomography on long-term follow-up. *J Comput Assist Tomogr*. 2012; 36(2):170–174. [PubMed: 22446354]
77. Lee HY, Lee KS, Jeong YJ, et al. High-resolution CT findings in fibrotic idiopathic interstitial pneumonias with little honeycombing: serial changes and prognostic implications. *AJR Am J Roentgenol*. 2012; 199(5):982–989. [PubMed: 23096169]
78. Lynch DA, David GJ, Safrin S, et al. High-resolution computed tomography in idiopathic pulmonary fibrosis: diagnosis and prognosis. *Am J Respir Crit Care Med*. 2005; 172:488–493. [PubMed: 15894598]
79. Maldonado F, Moua T, Rajagopalan S, et al. Automated quantification of radiological patterns predicts survival in idiopathic pulmonary fibrosis. *Eur Respir J*. 2014; 43:204–212. [PubMed: 23563264]

80. Nishino M, Itoh H, Hatabu H. A practical approach to high-resolution CT of diffuse lung disease. *Eur J Radiol.* 2014; 83(1):6–19. [PubMed: 23410907]
81. Raghu G, Collard HR, Egan JJ, et al. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med.* 2011; 183:788–824. [PubMed: 21471066]
82. Raghunath S, Rajagopalan S, Karwoski RA, et al. Quantitative stratification of diffuse parenchymal lung diseases. *PLoS One.* 2014; 9(3):e93229. [PubMed: 24676019]
83. Ryu JH, Daniels CE, Hartman TE, Yi ES. Diagnosis of interstitial lung diseases. *Mayo Clinic Proceedings.* 2007; 82(8):976–986. [PubMed: 17673067]
84. Sumikawa H, Johkoh T, Colby TV, et al. Computed tomography findings in pathological usual interstitial pneumonia: relationship to survival. *Am J Respir Crit Care Med.* 2008; 177:433–439. [PubMed: 17975197]
85. Travis WD, Costabel U, Hansell DM, et al. An official American Thoracic Society/European Respiratory Society statement: Update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. *Am J Respir Crit Care Med.* 2013; 188:733–748. [PubMed: 24032382]
86. Vij R, Strek ME. Diagnosis and treatment of connective tissue disease-associated interstitial lung disease. *Chest.* 2013; 143(3):814–824. [PubMed: 23460159]
87. Walsh SL, Sverzellati N, Devaraj A, Keir GJ, Wells AU, Hansell DM. Connective tissue disease related fibrotic lung disease: high resolution computed tomographic and pulmonary function indices as prognostic determinants. *Thorax.* 2014; 69:216–222. [PubMed: 24127020]
88. Wells AU, Desai SR, Rubens MB, et al. Idiopathic pulmonary fibrosis: a composite physiologic index derived from disease extent observed by computed tomography. *Am J Respir Crit Care Med.* 2003; 167:962–969. [PubMed: 12663338]
89. Watadani T, Sakai F, Johkoh T, et al. Interobserver variability in the CT assessment of honeycombing in the lungs. *Radiology.* 2013; 266:936–944. [PubMed: 23220902]
90. Silva CI, Müller NL, Hansell DM, Lee KS, Nicholson AG, Wells AU. Nonspecific interstitial pneumonia and idiopathic pulmonary fibrosis: changes in pattern and distribution of disease over time. *Radiology.* 2008; 247:251–259. [PubMed: 18270375]
91. Du Bois RM, Weycker D, Albera C, et al. Ascertainment of individual risk of mortality for patients with idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med.* 2011; 184:459–466. [PubMed: 21616999]
92. Mura M, Porretta MA, Bargagli E, et al. Predicting survival in newly diagnosed idiopathic pulmonary fibrosis: a 3-year prospective study. *Eur Respir J.* 2012; 40:101–109. [PubMed: 22241745]
93. Ryerson CJ, Vittinghoff E, Ley B, et al. Predicting survival across chronic interstitial lung disease: The ILD-GAP model. *Chest.* 2013; 145:723–728. [PubMed: 24114524]
94. Mogulkoc N, Brutsche MH, Bishop PW, Greaves SM, Horrocks AW, Egan JJ. Pulmonary function in idiopathic pulmonary fibrosis and referral for lung transplantation. *Am J Respir Crit Care Med.* 2001; 164:103–108. [PubMed: 11435247]
95. Jeong YJ, Lee KS, Müller NL, et al. Usual interstitial pneumonia and non-specific interstitial pneumonia: serial thin-section CT findings correlated with pulmonary function. *Korean J Radiol.* 2005; 6:143–152. [PubMed: 16145289]
96. Iwasawa T, Ogura T, Sakai F, et al. CT analysis of the effect of pirfenidone in patients with idiopathic pulmonary fibrosis. *Eur J Radiol.* 2014; 83:32–38. [PubMed: 22465123]
97. Kim EA, Johkoh T, Lee KS, et al. Interstitial pneumonia in progressive systemic sclerosis: serial high-resolution CT findings with functional correlation. *J Comput Assist Tomogr.* 2001; 25:757–763. [PubMed: 11584237]
98. Winklehner A, Berger N, Maurer B, Distler O, Alkadhi H, Frauenfelder T. Screening for interstitial lung disease in systemic sclerosis: the diagnostic accuracy of HRCT image series with high increment and reduced number of slices. *Ann Rheum Dis.* 2012; 71:549–552. [PubMed: 22121134]
99. Wells AU, Behr J, Silver R. Outcome measures in the lung. *Rheumatology.* 2008; 47:v48–v50. [PubMed: 18784144]

100. American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement. American Thoracic Society (ATS), and the European Respiratory Society (ERS). *Am J Respir Crit Care Med*. 2000 Feb; 161(2 Pt 1):646–664. [PubMed: 10673212]
101. Bradford WZ, Cohen AH, Leff JA. Selection of clinically meaningful primary endpoints in phase 3 clinical trials in idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*. 2013; 187(11):1269–1270. [PubMed: 23725618]
102. King JRTE, Safrin S, Starko KM, Brown KK, Nobel PW, Raghu G, et al. Analyses of efficacy end points in a controlled trial of interferon- γ 1b for idiopathic pulmonary fibrosis. *Chest*. 2005; 127:171–177. [PubMed: 15653980]
103. duBois RM, Weycker D, Albera C, et al. Forced vital capacity in patients with idiopathic pulmonary fibrosis: test properties and minimal clinically important difference. *Am J Resp Crit Care Med*. 2011; 184:1382–1389. [PubMed: 21940789]
104. Steen VD, Conte C, Owens GR, Medsger TA Jr. Severe restrictive lung disease in systemic sclerosis. *Arthritis Rheum*. 1994; 37:1283–1289. [PubMed: 7945490]
105. Bouros D, Wells AU, Nicholson AG, et al. Histopathologic subsets of fibrosing alveolitis in patients with systemic sclerosis and their relationship to outcome. *Am J Resp Crit Care Med*. 2002; 165:1581–1586. [PubMed: 12070056]
106. Latsi PI, du Bois RM, Nicholson AG, et al. Fibrotic idiopathic interstitial pneumonia: The prognostic value of longitudinal function trends. *Am J Resp Crit Care Med*. 2003; 168:531–537. [PubMed: 12791580]
107. Steen V. Advancements in Diagnosis of Pulmonary Arterial Hypertension in Scleroderma. *Arth Rheum*. 2005; 52(12):3698–3700. [PubMed: 16320319]
108. Eaton T, Young P, Milne D, Wells AU. Six-minute walk, maximal exercise tests: Reproducibility in fibrotic interstitial pneumonia. *Am J Resp Crit Care Med*. 2005; 171:1150–1157. [PubMed: 15640367]
109. Flaherty KR, Andrel A, Murray S, et al. Idiopathic pulmonary fibrosis: Prognostic value of changes in physiology and six-minute-walk test. *Am J Resp Crit Care Med*. 2006; 174:803–809. [PubMed: 16825656]
110. Garin MC, Highland KB, Silver RM, Strange C. Limitations to the 6-minute walk test in interstitial lung disease and pulmonary hypertension in scleroderma. *J Rheumatol*. 2009; 35:330–336. [PubMed: 19208563]
111. Lama VM, Flaherty KR, Toews GB, et al. Prognostic value of desaturation during a 6-minute walk test in idiopathic interstitial pneumonia. *Am J Resp Crit Care Med*. 2003; 168:1084–1090. [PubMed: 12917227]
112. Swigris JJ, Zhou X, Wamboldt FS, et al. Exercise peripheral oxygen saturation (SpO₂) accurately reflects arterial oxygen saturation (SaO₂) and predicts mortality in systemic sclerosis. *Thorax*. 2009; 64:626–630. [PubMed: 19359269]
113. Raghu G, Collard HR, Anstrom KJ, et al. Idiopathic pulmonary fibrosis: clinically meaningful primary endpoints in phase 3 clinical trials. *Am J Resp Crit Care Med*. 2012; 185:1044–1048. [PubMed: 22505745]
114. Komocsi A, Vorobcsuk A, Faludi R, Pinter T, Lenkey Z, Kolto G, Czirjak L. The impact of cardiopulmonary manifestations on the mortality of SSc: a systematic review and meta-analysis of observational studies. *Rheumatology*. 2012; 51(6):1027–1036. [PubMed: 22223705]
115. Olson AL, Swigris JJ, Sprunger DB, et al. Rheumatoid arthritis-interstitial lung disease-associated mortality. *American journal of respiratory and critical care medicine*. 2011; 183(3):372–378. [PubMed: 20851924]
116. 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; 69:1809–1815. [PubMed: 20551155]
117. Noth I, Anstrom KJ, Calvert SB, et al. Idiopathic Pulmonary Fibrosis Clinical Research N: A placebo-controlled randomized trial of warfarin in idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*. 2012; 186(1):88–95. [PubMed: 22561965]

118. Raghu G, Anstrom KJ, King TE Jr, Lasky JA, Martinez FJ. Idiopathic Pulmonary Fibrosis Clinical Research Network. Prednisone, azathioprine, and N-acetylcysteine for pulmonary fibrosis. *The New England journal of medicine*. 2012; 366(21):1968–1977. [PubMed: 22607134]
119. Wells AU, Behr J, Costabel U, Cottin V, Poletti V, Richeldi L, European IPFCG. Hot of the breath: mortality as a primary end-point in IPF treatment trials: the best is the enemy of the good. *Thorax*. 2012; 67(11):938–940. [PubMed: 23047830]
120. Richeldi L, du Bois RM. Pirfenidone in idiopathic pulmonary fibrosis: the CAPACITY program. *Expert Rev Respir Med*. 2011 Aug; 5(4):473–481. [PubMed: 21859266]
121. Taniguchi H, Kondoh Y, Ebina M, et al. Pirfenidone Clinical Study Group in Japan. The clinical significance of 5% change in vital capacity in patients with idiopathic pulmonary fibrosis: extended analysis of the pirfenidone trial. *Respir Res*. 2011 Jul 15;12:93. [PubMed: 21756364]
122. de Lauretis A, Veeraraghavan S, Renzoni E. Review series: Aspects of interstitial lung disease: connective tissue disease-associated interstitial lung disease: how does it differ from IPF? How should the clinical approach differ? *Chron Respir Dis*. 2011; 8(1):53–82. [PubMed: 21339375]
123. Seibold JR, Denton CP, Furst DE, et al. Randomized, prospective, placebo-controlled trial of bosentan in interstitial lung disease secondary to systemic sclerosis. *Arthritis and rheumatism*. 2010; 62(7):2101–2108. [PubMed: 20506355]
124. Collard HR, Moore BB, Flaherty KR, et al. Idiopathic Pulmonary Fibrosis Clinical Research Network Investigators Acute exacerbations of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*. 2007 Oct 1; 176(7):636–643. [PubMed: 17585107]
125. King TE Jr, Tooze JA, Schwarz MI, Brown KR, Cherniack RM. Predicting survival in idiopathic pulmonary fibrosis: scoring system and survival model. *Am J Respir Crit Care Med*. 2001; 164:1171–1181. [PubMed: 11673205]
126. Wells AU, Desai SR, Rubens MB, et al. Idiopathic pulmonary fibrosis: a composite physiologic index derived from disease extent observed by computed tomography. *Am J Respir Crit Care Med*. 2003; 167:962–969. [PubMed: 12663338]
127. Watters LC, King TE, Schwarz MI, Waldron JA, Stanford RE, Cherniack RM. A clinical radiographic and physiologic scoring system for the longitudinal assessment of patients with idiopathic pulmonary fibrosis. *Am Rev Respir Dis*. 1986; 133:97–103. [PubMed: 3942381]
128. Baughman RP, Drent M, Culver DA, et al. Endpoints for clinical trials of sarcoidosis. *Sarcoidosis Vasc Diffuse Lung Dis*. 2012 Oct; 29(2):90–98. [PubMed: 23461070]