

Editorial

Interstitial Lung Disease with Connective Tissue Diseases

The interstitial lung diseases (ILD) are a heterogeneous group of pulmonary disorders which diffusely affect the lung parenchyma including the small airways and alveolar structures. While many are of unknown cause, others have known clinical associations including infection, a diversity of inhaled, oral or parenteral exposures (including drugs, tobacco smoke, occupational and environmental exposures), and a variety of systemic diseases. In particular, ILD frequently complicates connective tissue disease (CTD), having a major negative impact on quality of life and disease course; pulmonary involvement is the leading cause of death in systemic sclerosis and an important cause of morbidity and mortality in other connective tissue diseases including rheumatoid arthritis, Sjögren's syndrome, systemic lupus erythematosus and inflammatory myopathies.

In many ways, connective tissue disease-associated ILD (CTD-ILD) remains an enigma. Our understanding of underlying pathogenesis and ability to detect these conditions early remain quite limited. For the clinician and clinical researcher, there exists no consensus on appropriate measurement of disease activity or what constitutes a significant response to therapeutic intervention. The lack of this information inhibits effective drug development and hampers the regulatory evaluation of candidate therapies. For example, in recent years, using pathologic insights gleaned from the biologic investigation of autoimmunity, a diverse range of data-driven "rational" therapies has emerged, from broad immunosuppressants (cyclophosphamide, mycophenolate mofetil) and antifibrotics (tyrosine kinase inhibitors, pirfenidone) to more specifically targeted interventions (rituximab, TGF- β trapping strategies). While substantial progress might have been expected from trials evaluating their efficacy, lack of validated and accepted criteria for assessment of disease activity or response to treatment has hampered our ability to interpret their results. Without cohesive domains and measures, there is no way to safeguard the investments made by research and development.

This issue of *Current Rheumatology Reviews* addresses this important area. There are emerging data relevant to ILD in the setting of each specific CTD yet there are also strong commonalities between them. Many questions remain regarding shared pathogenic mechanisms but also importantly about approaches which should be taken by the bedside clinician and the clinical researcher. Is there a core group of measures that can define activity, response and outcome? Is quantification of the extent of lung disease by high resolution CT a generalizable measure or only a tool for focused research? Pulmonary physiologic testing, e.g. forced vital capacity, is accessible, reproducible, cost effective and sensitive to change but is it confounded by concurrent chest wall weakness or pleural disease? What about exercise capacity, e.g. six minute walk testing - a robust outcome measure in trials of pulmonary hypertension? How does this apply to the individual with active lower extremity joint disease?

Importantly, also, each of the CTDs has its own unique set of clinical features which may affect not only the type and expression of associated lung disease but also the exact nature of pulmonary involvement in individual patients. This relationship is particularly relevant to the study of specific CTD-ILD in randomized controlled trials (RCTs). For example, what is the utility of exercise testing in a patient with rheumatoid arthritis who has a high disease activity score and/or lower extremity joint involvement? How do these disease features affect the performance and interpretation of pulmonary function tests? To what extent does muscle involvement in patients with dermatomyositis, or sicca syndrome in patients with Sjögren's syndrome, adversely affect the performance of pulmonary function testing. To date, there is no agreement on the best approach to evaluation of lung involvement in the clinic or in the context of clinical trials for any single CTD-associated ILDs.

How best to approach the problem of lung disease assessment in CTD-ILD for study in an RCT? To address this, a review of the available relevant literature can serve as a useful starting point, as presented in this volume of CRR. Information from the literature, interpreted by clinician scientist's expert in these diseases, forms a useful basis for evaluating the best evidence for developing outcome criteria for RCTs in CTD-ILD. A formal process to develop consensus can then provide an initial approach for assessment of lung disease in these CTDs, which necessarily will be followed by formal testing in RCTs and longitudinal observational studies (LOS).

In developing an initial consensus, the structure of the OMERACT (Outcome Measures in Rheumatology Clinical Trials) filter deserves consideration [1]. Outcome measures considered to be validated by OMERACT satisfy the following criteria:

1. **Truth:** the measure is truthful, measuring what it is intended to measure. The result is unbiased and relevant. This criterion captures the issues of face, content, construct and criterion validity.
2. **Discrimination:** the measure discriminates accurately between sub-groups or scenarios of interest. This applies equally to evaluation at a single point in time (for classification or prognostic purposes), serial evaluation (in the quantification of change), and in the definition of sub-groups of interest. This criterion captures the issues of reliability as well as sensitivity to change.
3. **Feasibility:** the measure is practicable in the RCTs or LOS, despite constraints of time, money, and interpretability. This criterion addresses whether a measure is "user friendly", often, in reality, the major determinant of its successful application.

The need for this approach is apparent in scleroderma-associated ILD. Treatment studies of ILD in systemic sclerosis have demonstrated treatment effects, but have also laid bare the complexities of clinical trial design. It was widely believed that suppression of inflammation (“alveolitis”) might result in worthwhile regression of disease. Thus, cyclophosphamide was studied in subjects thought to have active inflammation, as defined by increased neutrophilia on bronchoalveolar lavage and/or “ground glass” appearance on high resolution chest tomography (HRCT). It is, perhaps, remarkable that a clear treatment effect was observed, despite the fact that neither the use of CT and bronchoalveolar lavage (BAL) for this purpose nor the conceptual basis of the study stands up to close scrutiny. Ground-glass on HRCT often denotes fine intralobular fibrosis, rather than the predominance of inflammation; in some reports, ground-glass on HRCT is resistant to anti-inflammatory treatment in the majority of cases. Similarly, neutrophilia seen by BAL may be indicative of interstitial inflammation in some cases but is also associated with extensive fibrotic disease. In large recent series, a BAL neutrophilia did not identify scleroderma patients with more progressive lung disease. More importantly, the treatment effect of cyclophosphamide in an RCT performed in the United States did not consist of regression of inflammation, but the prevention of disease progression in patients with greater impairment of forced vital capacity (FVC) and more extensive coarse fibrotic disease on HRCT [2, 3]. Thus, a concept of “alveolitis”-predominant inflammation drove trial design but outcome was determined by measures of “damage”. When viewed against the OMERACT filter, measures of activity lacked truth and discrimination.

There are no Food and Drug Administration (FDA) approved therapies available for the treatment of any of the ILDs. Diagnosis and treatment principles have been outlined in the interstitial lung disease guideline of the British Thoracic Society in collaboration with the Thoracic Society of Australia and New Zealand and the Irish Thoracic Society [4]. With increasing numbers of investigators evaluating promising therapies, it is important to describe current thinking regarding the design of RCTs in patients with ILD. The following protocol design elements should be considered when constructing a clinical development program in patients with CTD-ILD: 1) patient population, 2) concomitant medications, 3) dose selection, and 4) outcome measures/endpoints.

With regard to the patient population, the primary diagnostic criterion is the presence of CTD-ILD, as judged by HRCT, and it is imperative that HRCT interpretation be subject to quality control, with discussion of marginal cases by consensus panels. Fortunately, the detection of pulmonary fibrosis on HRCT is straightforward in most cases. The more difficult issue relates to excluding clinically insignificant disease, for which a treatment effect is unlikely to be demonstrable. This important design issue, which has been only partially addressed, requires the definition of severity thresholds that reliably designate lung involvement as clinically important. The need for standardized algorithms in future studies, to exclude patients with clinically insignificant disease on severity criteria, appears axiomatic. Furthermore, the severity of lung function impairment will, along with the choice of endpoints, determine the duration and number of subjects required. Finally, a priori statements of the intention to compare treatment effects in advanced and less advanced disease are essential; as such analyses lose credibility when performed post hoc.

Limiting the use of concomitant medications in a CTD-ILD study can be challenging from the clinical as well as trial design point of view. A wide variety of therapies have been used for ILD, and few are considered to be “standard of care”. Inclusion of multiple therapies may obscure the treatment effect, increase the risk of drug-drug interactions, and complicate analysis of the trial. As there are currently no FDA-approved therapies for an ILD indication, a study which prohibits the use of experimental drugs is more likely to provide the most interpretable information regarding the efficacy of the treatment under study. If concomitant medications must be allowed, they should be limited to those that are considered standard of care, and investigators should ensure that their use is balanced across treatment groups.

Clinical development programs should include full characterization of dosing regimens, including dose and dose schedule. Robust dose-finding may be precluded by the limited size of the population of patients with CTD-ILD, justifying a more mechanistic approach to dose selection. Nonetheless the use of more than one dose is particularly useful in phase III trials.

As there are currently no well-validated surrogate endpoints, RCTs in CTD-ILD should emphasize outcomes that are clinically meaningful to patients. Examples of such endpoints in ILD include mortality (survival), time to lung transplantation, disease progression, hospitalizations, exacerbations, and decline in lung function. The choice of primary endpoint(s) as well as baseline disease severity in pivotal trials will dictate the appropriate duration and number of RCTs required providing substantial evidence of efficacy. With this principle in mind, there are three classes of endpoint which, in our opinion, would currently be acceptable to provide substantial evidence of efficacy in CTD-ILD: 1) survival, 2) survival + disease progression, and 3) other clinically meaningful outcomes.

Disease progression with survival as a composite endpoint may be optimal. An example could be to define disease progression as a $\geq 15\%$ decline in the diffusion capacity for carbon monoxide (DLco) and $\geq 10\%$ decrease in forced vital capacity (FVC), while including death or lung transplantation as outcomes if they occur. It is important to use a by patient definition of success for the measure(s). Adequate separation of lung function decline (slope), along with the mortality outcome trending in support of the composite, would be necessary to establish convincing evidence of efficacy. This type of study design will likely require trials of longer duration.

Finally, other clinically meaningful efficacy variables such as hospitalizations, exacerbations, or preservation of lung function (FVC and DLco) are also plausible for clinical trials in ILD. However, these variables may require a longer duration than those described above, in order to establish separation of treatment groups; therefore, we encourage investigators to measure mortality as well, as a means of validating these variables.

From a safety standpoint, the challenge lies in accumulating a large number of ILD patients. As a result, the International Conference on Harmonization (ICH) guidelines for required drug exposure may not apply. In order to increase the size of the safety database, investigators may want to include CTD-ILD patients with other disease processes.

The standards of modern regulatory agencies serve to define the essence of the problem of assessment of treatment effect in patients with CTD-ILD. As the Chinese proverb states, "A journey of a thousand miles begins with a single step". The lessons of earlier trials lead to more robust designs. Hard but reliable endpoints will require sophistication in cohort enrichment strategies and large and longer collaborative trials. The characteristics of the study population likely to demonstrate change, inform bedside decision making in the individual patient. Ultimately, better understanding of the pathophysiology will guide the development of best treatments. It is time to develop robust and meaningful measures for evaluation of these diseases.

**Eric Matteson
(Guest Editor)**

Division of Rheumatology
Mayo Clinic College of Medicine
200 1st St. S.W.
Rochester, MN 55905, USA
Tel: 1-507-284-8450
Fax: 1-507-284-0564
E-mail: matteson.eric@mayo.edu

Kevin K. Brown

Department of Medicine
National Jewish Health
Denver, CO, USA

Athol U. Wells

Interstitial Lung Disease Unit
Royal Brompton Hospital and National Heart and Lung Institute
Imperial College, London, UK

Badrul A. Chowdhury

Division of Pulmonary and Allergy Products
Food and Drug Administration, Rockford
MD, USA

Banu A. Karimi-Shah, MD

Division of Pulmonary and Allergy Products
Food and Drug Administration, Rockford
MD, USA

Vibeke Strand

Division of Rheumatology
Stanford University, Palo Alto
CA, USA

James R. Seibold

Division of Rheumatology
University of Connecticut Health Center
Farmington, CT, USA

REFERENCES

- [1] Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT filter for outcome measures in rheumatology. *J Rheumatol* 1998; 25: 198-9.
- [2] Tashkin DP, Elashoff R, Clements PJ, *et al.* Cyclophosphamide *versus* placebo in scleroderma lung disease. *N Eng J Med* 2006; 354: 2655-66.
- [3] Nannini C, West CP, Erwin PJ, Matteson EL. Effects of cyclophosphamide on pulmonary function in patients with scleroderma and interstitial lung disease: A systematic review and meta-analysis of randomized controlled trials and observational prospective cohort studies. E-published. *Arthritis Res Ther* 2009; 11(2): 408.
- [4] Wells AU, Hirani N. On behalf of the BTS Interstitial Lung Disease Guideline Group, a subgroup of the British Thoracic Society Standards of Care Committee, in collaboration with the Thoracic Society of Australia and New Zealand and the Irish Thoracic Society. Interstitial lung disease guideline: The British Thoracic Society in collaboration with the Thoracic Society of Australia and New Zealand and the Irish Thoracic Society. *Thorax* 2008; 63(Suppl 5): v1-58.