Title

Twenty-two points to consider for clinical trials in systemic sclerosis, based on EULAR standards

Journal

Rheumatology, 54(1)

ISSN

1462-0324

Authors

Khanna, Dinesh
Furst, Daniel E
Allanore, Yannick
et al.

Publication Date

2015

DOI

10.1093/rheumatology/keu288

Peer reviewed

RHEUMATOLOGY

Rheumatology 2015;54:144–151
doi:10.1093/rheumatology/keu288
Advance Access publication 13 August 2014

Original article

Twenty-two points to consider for clinical trials in systemic sclerosis, based on EULAR standards

Dinesh Khanna[1], Daniel E. Furst[2], Yannick Allanore[3,4], Sangmee Bae[2], Vijay Bodukam[5], Philip J. Clements[2], Maurizio Cutolo[6], Laszlo Czirjak[7], Christopher P. Denton[8], Oliver Distler[9], Ulrich A. Walker[10], Marco Matucci-Cerinic[11,12], Ulf Müller-Ladner[13], James R. Seibold[14], Manjit Singh[15] and Alan Tyndall[10]

Abstract

Objective. SSc is clinically and aetiopathogenically heterogeneous. Consensus standards for more uniform trial design and selection of outcome measures are needed. The objective of this study was to develop evidence-based points to consider (PTCs) for future clinical trials in SSc.

Methods. Thirteen international SSc experts experienced in SSc clinical trial design were invited to participate. One researcher with experience in systematic literature review and three trainees were also included. A systematic review using PubMed and the Cochrane Central Register of Controlled Trials was conducted and PTCs when designing clinical trials in SSc were developed. As part of that development we conducted an Internet-based Delphi exercise regarding the main points to be made in the consensus statement. Consensus was defined as achieving a median score of $\geqslant 7$ of 9.

Results. By consensus, the experts decided to develop PTCs for each individual organ system. The current document provides a unifying outline on PTCs regarding general trial design, inclusion/exclusion criteria and analysis. Consensus was achieved regarding all the main points of the PTCs.

Conclusion. Using European League Against Rheumatism suggestions for PTCs, a general outline for PTCs for controlled clinical trials in SSc was developed. Specific outlines for individual organ systems are to be published separately. This general outline should lead to more uniform and higher-quality trials and clearly delineate areas where further research is needed.

Key words: systemic sclerosis, clinical trials, points to consider.

CLINICAL SCIENCE

Introduction

SSc is clinically and aetiopathogenically heterogeneous [1]. Among the many different immune-mediated rheumatic diseases, SSc stands out as a severely incapacitating and life-threatening disease for which therapeutic options are few and insufficient.

Recent years have seen important refinements in the development and validation of candidate outcome measures

[1]Division of Rheumatology, University of Michigan Scleroderma Program, Ann Arbor, MI, [2]Division of Rheumatology, University of California at Los Angeles, Los Angeles, CA, USA, [3]Department of Rheumatology A, Paris Descartes University, [4]Cochin Institut, INSERM U1016, Cochin Hospital, Paris, France, [5]Department of Medicine, Crozer Chester Medical Center, Upland, PA, USA, [6]Research Laboratory and Academic Unit of Clinical Rheumatology, Department of Internal Medicine, University of Genoa, Genoa, Italy, [7]Department of Immunology and Rheumatology, University of Pécs, Pécs, Hungary, [8]Centre for Rheumatology, Royal Free Hospital, London, UK, [9]Department of Rheumatology, University Hospital Zurich, Zurich, [10]Deparment of Rheumatology, Basel University, Basel, Switzerland, [11]Department of Biomedicine, Division of Rheumatology, Azienda Ospedaliero Universitaria Careggi, [12]Department of Medicine, Denothe Centre, University of Florence, Florence, Italy, [13]Department of Rheumatology and Clinical Immunology, Justus-Liebrig University Giessen, Kerckhoff Clinic, Bad Beuheinn, Germany, [14]Scleroderma Research Consultants, Avon, CT and [15]Department of Internal Medicine, Rochester General Health System, Rochester, NY, USA.

Submitted 29 July 2013; revised version accepted 28 May 2014.

Correspondence to: Daniel E. Furst, Division of Rheumatology, University of Michigan Scleroderma Program, 1000 Veteran Avenue, Room 32-59, Los Angeles, CA 90095-1670, USA.
E-mail: defurst@mednet.ucla.edu

Dinesh Khanna and Daniel E. Furst contributed equally to this study.

[2, 3] and increased sophistication in trial methodology in SSc [4]. This is paralleled by an increased understanding of the pathogenesis of SSc [5, 6] and thus the possibility to develop more targeted therapies [7, 8]. Controlled trials may target constitutive elements of the disease process (e.g. vasculopathy, fibrosis, immune activation) or might focus on more narrow clinical outcomes [e.g. digital ulcers, interstitial lung disease (ILD)]. This complexity hinders comparisons between trials and contributes to delay in evaluating the most appropriate therapeutics for this disease. Under the auspices of the European League Against Rheumatism (EULAR), we undertook the task of developing points to consider (PTCs) for conducting clinical trials in SSc [9] using a combination of research-based evidence and expert consensus.

Methods

Expert committee members

The steering committee consisted of 13 experts in the field of SSc. In addition, three trainees and one methodologist were recruited for the systematic reviews (discussed below). We also sought input from experts in cardiopulmonary and lung involvement in SSc.

Current and past trials have approached SSc within one of the following constructs: (i) overall survival, (ii) specific organ-based complications and (iii) measures of composite response including several related disease features. An agent with putative antifibrotic effects might reasonably target clinical features thought to represent tissue fibrosis (e.g. skin and parenchymal lung involvement), but effects might not be measurable in patients with very mild expression of these disease features. An agent with putative antivascular effects might reasonably target clinical features thought to represent vascular complications (e.g. digital ulceration or pulmonary vascular syndromes), but might permit the study of broader populations. These examples of heterogeneity of disease expression provide part of the rationale for consideration of organ-specific PTCs.

In SSc, clinical trials for each organ system have their own particular requirements [2, 10]. In addition, and because there are significantly different issues to consider among the various organs potentially affected in SSc, the experts decided to approach SSc based on individual organ systems against a background of a general, unifying approach. The present statement represents the discussion of the general unifying approach.

The 11 specific organ systems or aspects to be considered will be published elsewhere and will include cardiac, renal, digital ulcers, gastrointestinal, health-related quality of life and functional disability, joints, muscle, pulmonary fibrotic, pulmonary vascular, RP and skin involvement.

Structured search strategy

We wished to have an appropriate literature background and evidence base for our considerations. Thus We did a systematic literature review using PubMed and the Cochrane Central Register of Controlled Trials. We examined the literature between 1995 (the time of the publication of the previous guidelines) and January 2011. The only exceptions were the organ systems of muscles and joints, where we examined the literature from 1966 to 2011 because there were very few articles that met our search strategy in these areas since 1995. Further, the bibliographies of all articles unearthed by our search strategy were reviewed for additional articles.

All articles that were clinical trials in SSc and were in English were included, with the following exclusions: animal studies; not concerned with humans; not pertaining to SSc; not a case–control study, case series, cohort study, database or registry; not pertaining to instruments or diagnostic tests; editorial, review article, letter or opinion; pertaining to infants or children; genetic studies (i.e. polymorphisms, genetic associations with internal organs, etc.). We obtained 4901 titles and ultimately extracted 903 titles as our evidence-based literature. These articles were the basis for both the present, general PTC discussions and for the organ-specific PTCs.

The PTCs were drafted, considered and revised by all authors. The principle points themselves were subjected to a Delphi exercise done on the Internet. Although the first round achieved consensus (median scores of $\geqslant 7$ of 9, where 9 was totally appropriate) on all statements, a second round was undertaken for all questions for which at least one participant gave a score of 1–3 or at least two participants gave a score of 4–6.

Many of the PTCs are based on evidence from general clinical science in SSc. When specific references are available, they are cited. The statements are labelled according to the quality of the evidence supporting them, using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) System [11].

Quality of evidence

The following definitions were used:

High quality: further research is very unlikely to change confidence in the statement.
Moderate quality: further research is likely to have an important impact on confidence in the statement and may change the statement.
Low quality: further research is very likely to have an important impact on confidence in the statement and is likely to change the statement.
Very low quality: any statement is very uncertain.

Results

Table 1 outlines the questions and median scores after the second, and final, Delphi round.

General design

There was clear consensus that clinical trials in SSc should adhere to the general approach supported by the US Food and Drug Administration (FDA; http://www.fda.gov) and the European Medicines Agency (http://www.

Table 1 Results of the Delphi exercise on points to consider for clinical trials in SSc

Result of the Delphi exercise on points to consider for clinical trials in SSc	Median
All trials should be well controlled, meaning, in general, that they should be randomized, blinded and controlled, particularly when undertaking phase 2 or phase 3 trials	9
Non-randomized trials should only be considered in the very earliest phases of therapeutic development, although, even here, randomization may be the most appropriate approach	8
While trials could be open label, blinding, either single or double, is most appropriate	9
In general, placebo-controlled trials, allowing appropriate background therapy, are strongly favoured	8
It was agreed that when effective therapies for a given organ system are available, positively controlled trials can be considered	8
Trial duration is often a critical consideration and should be tailored for the specific medication and organ system manifestation	9
Biologic response trials may be particularly short, although clinical correlations are highly desirable	8
Ethical considerations need to be adhered to	9
The CONSORT guidelines are an appropriate outline for reporting clinical trials	9
Biosampling should be done whenever possible	8
Obvious considerations in designing clinical trials include gender, age and disease subsets	8
Most trials of therapeutic interventions in SSc are initially done in adults	9
Testing of medications in children may be required and important, so trials in the young should be considered at some point	8
Trials in scleroderma should generally be done in uniform disease subsets (e.g. diffuse, limited or diffuse/limited)	8
Generally patients with well-defined overlapping diseases should be excluded from SSc clinical trials	8
Disease duration needs to be considered	9
Well-defined SSc should be one of the inclusion criteria	9
Environmental exposure to substances that have been associated with scleroderma-like disease (e.g. vinyl chloride, trichloroethylene, silica dust) should be excluded so that a uniform group of patients is tested	8
Concomitant medications need to be carefully considered when defining a trial in SSc	9
Concomitant diseases need to be considered when defining a trial in SSc	9
The primary outcome measure should be a validated measure	8
If the contemplated primary outcome has not been validated, it might be prudent to develop or test such validation as a preliminary to a phase 3 trial	8
Other organ system manifestations that might confound the primary outcome or might result in dropout before study completion should be considered	9
When analysing a clinical trial, prespecified analyses are important	9
The patient population needs to be described in sufficient detail so that one is able to understand the type of patient for whom the intervention would be applicable	9
Depending on the phase of the study, power analysis to define the number of patients needed to have confidence in the results is appropriate	9
Power analyses are not always necessary if the result is aimed at understanding future study design or getting a sense of the safety of a treatment	8
Statistical tests should consider the distribution of results (e.g. parametric versus non-parametric distributions), characteristics of the outcomes (continuous, ordinal, dichotomous) and how to deal with the inevitable missing results (a strategy to account for missing data)	9
How to summarize and examine adverse events should be considered	9

ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002874.pdf) for good clinical practice. These guidelines indicate that

(i) All trials should be well controlled, meaning, in general, that they should be randomized, blinded and controlled, particularly when undertaking phase 2 or phase 3 trials (high quality).

There are other considerations, however, and there are some nuances. Non-randomized trials could be considered in the very earliest phases of therapeutic development, although, even here, some experts felt that randomization is the most appropriate approach (low quality).

While trials could be open label, a clear consensus for blinding, either single or double, was found, as the biases

in open trials have, in the past, led to incorrect conclusions (high quality). Examples are D-penicillamine, colchicine and dimethyl sulphoxide (DMSO) [12–14]. Some experts felt that single blinding and open studies might be appropriate under specific circumstances where double or single blinding might be impossible (e.g. when a therapeutic intervention could not be blinded, such as acupuncture and stem cell transplantation). Even here, attempts at blinding should be made, such as using a blinded observer while a non-blinded investigator cares for the patient or overviews laboratory tests that might unblind the study.

(ii) In general, placebo-controlled trials, allowing appropriate background therapy, were favoured (moderate quality). Because SSc can be such a severe and progressive disease, there was some sentiment for positively controlled trials. It was agreed that when effective therapies for a given organ system are available, positively controlled trials should be considered (high quality). An example of such an instance is the use of CYC for ILD; in this subset of patients, a positively controlled trial has, in fact, been undertaken and showed short-term efficacy (18 months) [15]. However, 2 year follow-up data of CYC questioned its long-term benefits and issues of statistical power and statistical non-inferiority may make positively controlled trials difficult to conduct in a disease in which the prevalence and incidence is low.

Trial duration

(iii) Trial duration is often a critical consideration (moderate quality). Selecting a trial duration that is too short may yield an inappropriate negative result despite the temptation to keep trials short for both time and cost considerations. On the other hand, choosing a trial duration that is too long (particularly if placebo controlled) may be ethically questionable. Some trial durations can be particularly short (e.g. 6 weeks) if all that is being examined is a biological response, although such a response without clinical correlation may be of very limited use. Trial duration can sometimes be derived from an examination of the literature. For example, a 12 week trial might be appropriate for RP and is supported by the literature [16]. A 12- to 16-week trial would be more appropriate for testing pulmonary hypertension, while 16–24 weeks may be appropriate for testing haemodynamic changes in pulmonary hypertension prevention and healing of digital ulcers, again supported by the literature [17, 18]. Longer trials of 6–24 months may be necessary to prove an effect on remodelling and fibrotic outcomes such as in ILD or on skin fibrosis [19–22].

Ethical considerations

(iv) Ethical aspects need to be considered, which is a given (high quality). In addition, the Consolidated Standards of Reporting Trials (CONSORT) guidelines were accepted as an appropriate outline for reporting clinical trials (moderate quality) [23].

Biosampling

Our understanding of the pathogenesis of SSc is incomplete and better treatments are clearly needed. In that context it was agreed that

(v) Biosampling should be done whenever possible, as such sampling allows exploring new pathways or new treatments and validating biomarkers (low quality). While it was understood that there may be barriers to biosampling, such as the difficulty with storage or transportation and human subject protection issues, it was also clear that results in other connective tissue diseases could not necessarily be transposed to SSc. Thus biosampling according to good laboratory and clinical practice with respect to collection, storage and distribution should be considered in any trial and is encouraged. Such guidelines have been proposed and include serum, plasma, cellular and biological samples such as skin or lung [24].

Inclusion/exclusion criteria

(vi) A major issue in clinical trials is choosing appropriate and uniform patient groups. Uniformity in patient groups improves the likelihood of a clear outcome (high quality). At the same time, uniformity may decrease the generalizability of the results.

(vii) Well-defined SSc should be one of the inclusion criteria (high quality). However, this may require some thought. There are, for example, several definitions of SSc, including the preliminary 1980 ACR criteria and new criteria attempting to define the disease at an earlier stage [25–27]. These may result in very different populations of patients and may make it difficult to compare patient groups across different trials. At the present time, most trials use the 1980 ACR criteria. If one wishes to consider more than one set of diagnostic criteria, it might be best to define one set of criteria as the primary one while analysing the trial in an exploratory manner in terms of other criteria. In this way, cross-trial comparisons can still be made.

Scleroderma can be classified as a systemic disease, with subsets of diffuse and limited cutaneous disease, overlap disease and non-systemic localized disease. Many SSc trials are done in SSc patients with diffuse disease because outcome measures have only been validated in the diffuse cutaneous disease subset. On the other hand, when considering various visceral involvements, both limited cutaneous and diffuse cutaneous disease might be appropriate. The latter is true, for example, when testing ILD or pulmonary arterial hypertension. Likewise, many patients have an overlap with other diseases such as RA, polymyositis or SLE. While including overlap patients will increase recruitment, the inclusion of patients with multiple diseases will likely confound any potential therapeutic benefit and might also result in adverse effects that are not seen in patients with pure SSc.

(viii) Other considerations include gender, age and disease duration (high quality). Gender considerations must include the fact that SSc is more frequent in

women, so both men and women will clearly need to be recruited. Pregnancy and fertility considerations are necessary for both men and women when seeking to do a trial in SSc. Breastfeeding subjects should generally be excluded as medications can be transported in breast milk in unknown amounts, thus exposing infants without appropriate knowledge of the medication's safety in infants. When such knowledge has been gained, limited studies in breastfeeding mothers may be considered.

(ix) Most trials of therapeutic interventions in SSc are initially done in adults, as the therapies may adversely affect growth, development and fertility, thus contraindicating their use in children (moderate quality). On the other hand, testing of medications in children may be required and is important, so trials in the young should be considered at some point. If the therapeutic intervention does not have an adverse potential in children (e.g. a physical therapy intervention), trials in children should be encouraged.

(x) Disease duration needs to be considered in terms of the definition of disease duration. Some believe that duration should be defined based on the first symptom or sign, such as RP; others feel that, because RP may occur many years before the next sign or symptom, duration should be from the first typical sign or symptom other than RP [15]. This has major implications because the duration of disease is thought by most to help define the likelihood of response to specific therapies. Early disease (3–5 years) changes rapidly and thus therapeutic change can be discerned in a relatively short period of time (e.g. 1 year), while late disease may change slowly and may require much longer trials. In general, most trials define disease duration from the first non-RP sign or symptom.

(xi) In general, environmental exposure to substances that have been associated with scleroderma-like disease (e.g. vinyl chloride, trichloroethylene, silica dust) should be excluded. These should be excluded by history so that a uniform group of patients is tested (moderate quality). While one might argue that targeting scleroderma-like disease might be desirable, the lack of understanding of how these external environmental stimulations result in disease make it unlikely that such a choice would be prudent.

(xii) Concomitant medications need to be carefully considered when defining a trial in SSc (high quality). It would be ethically inappropriate to insist on allowing no concomitant medication and it would also make recruitment impossible. On the other hand, some background medications may confound and/or obscure a therapeutic response. For example, excluding corticosteroids beyond a certain dose or requiring a stable background dose of corticosteroids would seem prudent. It was agreed that prednisone or its equivalent at $\leqslant$10 mg/day in a stable regimen might be acceptable in trials of the skin, joints or lungs, but higher doses might obscure results or increase the risk of scleroderma renal crisis. Excluding calcium channel blockers (CCBs), angiotensin-converting enzyme (ACE) inhibitors or selective serotonin reuptake inhibitors (SSRIs) in trials of RP would be necessary, as these all decrease vascular reactivity. The use of background immunosuppressants in a trial of immunosuppressants could obviously confound results, although a trial allowing background immunosuppressants might be considered if a particular therapeutic intervention might add to or enhance the effect of the background immunosuppression [19–22].

(xiii) Concomitant diseases need to be considered (high quality). Some diseases may interfere with assessment of the intervention and therefore should be excluded. For example, uncontrolled hypertension should be excluded in a protocol oriented towards cardiac or renal involvement. Malignancy, liver disease or diabetes should be excluded if one is considering the need for significant follow-up, if a drug is metabolized by the liver or if outcome measures may be interfered with by the presence of a polyneuropathy (e.g. in diabetics). On the other hand, allowing multiple concomitant diseases will allow better generalization, as the patients are more likely to be those found in the general population. Further, allowing multiple concomitant diseases will make recruitment easier. In general, however, allowing unstable concomitant illnesses will interfere with the ability to measure outcomes and thus will increase the probability of a false result.

(xiv) Baseline disease severity might also be considered (moderate quality). End-stage patients with severely damaged organs are very unlikely to be able to improve sufficiently to be measurable in a short (1 year) study. Some authorities suggest the use of quantitative or even qualitative nailfold capillaroscopy as such a measure of severity [28]. The Medsger severity scale has also been used to establish severity [29]. One issue here is the difficulty of separating damage (irreversible to a large extent) from activity (frequently reversible), but there is no agreement on how this can be done at the present time.

Outcomes

(xv) The primary outcome measures should be validated measures (high quality).
This is exemplified by adherence to standards such as the OMERACT filters of truth, discrimination and feasibility [30]. Although it was recognized that there is a relative paucity of validated outcome measures in SSc, it was agreed that carefully defined and validated primary outcome measures should be used, and there are, in fact, a number of validated measures available in SSc [31–35]. For example, in a therapeutic trial aimed at the skin, the validated modified Rodnan skin score should be used and carefully defined upper and lower limits of the skin score should be stated in the inclusion criteria [36, 37]. This is needed to avoid floor or ceiling effects (see the PTCs on skin for further discussion).

(xvi) If a therapeutic intervention might affect other organ systems in addition to the primary one, those organ systems also need to be carefully defined in the inclusion criteria (moderate quality). For example, in a trial where

the skin is the primary measure, the acceptable pulmonary function tests [e.g. forced vital capacity (FVC), a validated measurement in SSC)] should be defined if the therapeutic intervention might affect the lungs as well [37, 38].

If the contemplated primary outcome has not been validated, it might be necessary to develop and/or validate such a test as a preliminary to a phase 3 trial. This development could occur during a phase 1–2B trial while simultaneously using another valid and appropriate outcome [e.g. validating high-resolution CT (HRCT) of the lungs while using the FVC during phase 1–2B].

Other organ system manifestations that might confound the primary outcome or might result in dropout before study completion should be considered. For example, a trial of the lungs should exclude patients with significant cardiac involvement or myositis, as those illnesses may confound results or, if severe, the patient might die before completing the trial.

While an overall estimation of involvement by SSc, such as the DAS in RA, may be appropriate, such a combined score might best be used in a therapeutic trial of an intervention that has very widespread effects. An example of such a therapeutic approach would be stem cell transplantation in SSc [38]. Thus far the EULAR Scleroderma Trials and Research group (EUSTAR) activity measure is the closest to validation as a combined measure and several others are being considered [39].

Analysis

(xvii) When analysing a clinical trial, prespecified analyses are important (high quality). Without prespecification of analyses, a trial may simply become a fishing expedition where multiple analyses are done and only the one best fitting the desired outcome is published. Obviously such an approach is not appropriate or credible.

(xviii) The patient population needs to be described in sufficient detail to define the type of patient for whom the intervention would be applicable (high quality). For example, age, gender distribution, disease duration, organ involvement, concomitant diseases and concomitant medications all might be considered.

(xix) Depending on the phase of the study, power analysis to define the number of patients needed to have confidence in the results would be appropriate (moderate quality). Thus predefining the probability of a false positive and false negative result is helpful in understanding the meaning of the result. Power analyses should be done for all phase 3 and most phase 2 trials, but are not always necessary if the result is aimed at understanding future study design or getting a sense of the safety of a treatment. Likewise, power analyses are not needed if one is simply seeking a biological response without clinical correlates or a pharmacokinetic result.

(xx) If comparisons between groups are desired, statistical tests should consider the distribution of results (e.g. parametric versus non-parametric distributions), characteristics of the outcomes (continuous, ordinal, dichotomous) and how to deal with the inevitable missing results (a strategy to account for missing data) (high quality) [40]. Sometimes the distribution of results may not be known, but considering these possibilities was felt by all to be likely to enhance the credibility of the outcomes derived from the study. Appropriate tests could include analysis of variance, analysis of covariance, linear or logistic regression, generalized estimating equations, survival analyses etc.

(xxi) Considerations of how to summarize and examine adverse events should be undertaken (high quality). Thus, for example, one might wish to enumerate the adverse events or consider percentage occurrence or rates, when appropriate.

(xxii) As serious adverse events (e.g. death or hospitalization) are always of concern, they should be carefully described.

Conclusion

Using a literature review, Delphi exercises and a consensus-driven approach, a general set of PTCs when doing clinical trials for SSc is described. We hope that the PTCs presented here will help to clarify these issues and give some guidance for clinical trial design.

Rheumatology key messages

- Placebo-controlled trials using appropriate background therapy are needed in SSc.
- Well-defined SSc and uniform patient selection will improve outcomes.
- Validated SSc outcomes and predetermined analyses are required.

Funding: The work described in this article was funded by EULAR.

Disclosure statement: D.E.F. has received grants/research support from AbbVie, Actelion, Amgen, BMS, Gilead, GSK, NIH, Novartis, Pfizer, Roche/Genentech and UCB and has received honoraria as a consultant for AbbVie, Actelion, Amgen, BMS, Janssen, Gilead, GSK, NIH, Novartis, Pfizer, Roche/Genentech and UCB and for participation on speaker's bureaus (CME only) for AbbVie, Actelion and UCB. J.R.S. has received speaking and/or consulting fees from Actelion, Aires, Bayer, Boehringer-Ingelheim, Celgene, DART, Eiccose, Eiger, EMD Serono, Fibrogen, Gilead, InterMune, Novartis, Pfizer, Sanofi-Aventis and United Therapeutics. U.M.-L. has received speaker's honoraria from Actelion, Pfizer and GlaxoSmithKline and his scientific work is funded in part by the EULAR, EUSTAR and European Union Seventh Framework Programme for Research (FP7). Y.A. received honoraria and/or research grants from Roche, Actelion, Pfizer, Sanofi and UCB. C.P.D. has consulted for and received honoraria from Actelion, Genetech-Roche, Pfizer, GlaxoSmithKline,

Merck-Serono and Sanofi-Aventis. D.K. has received grants from the National Institutes of Health (NIH), Scleroderma Foundation and Pulmonary Hypertension Association and is a consultant for Actelion, Bayer, Biogen Idec, BMS, DIGNA, Genentech/Roche, InterMune, Merck and Sanofi-Aventis/Genzyme. O.D. has had consultancy relationships and/or received research funding from Actelion, Pfizer, Ergonex, BMS, Sanofi-Aventis, United BioSource, Roche/Genentech, Medac, Biovitrium, Boehringer Ingelheim Pharma, Novartis, 4D Science, Active Biotec, Sinoxa, Serodapharm, Bayer, GSK and EpiPharm in the area of potential treatments for scleroderma and its complications. O.D. has a licensed patent (miR-29) for the treatment of SSc. All other authors have declared no conflicts of interest.

References

1 Charles C, Clements PJ, Furst DE. Systemic sclerosis: hypothesis-driven treatment strategies. Lancet 2006;367: 1683–91.

2 Furst DE. Outcome measures in rheumatologic clinical trials and systemic sclerosis. Rheumatology 2008; 47(Suppl 5):v29–30.

3 Khanna D, Lovell DJ, Giannini E *et al*. Development of a provisional core set of response measures for clinical trials of systemic sclerosis. Ann Rheum Dis 2008;67:703–9.

4 White B, Bauer EA, Goldsmith LA *et al*. Guidelines for clinical trials in systemic sclerosis (scleroderma). I. Disease-modifying interventions. The American College of Rheumatology Committee on Design and outcomes in clinical trials in systemic sclerosis. Arthritis Rheum 1995; 38:351–60.

5 Abraham DJ, Varga J. Scleroderma: from cell and molecular mechanisms to disease models. Trends Immunol 2005;26:587–95.

6 Varga J, Abraham D. Systemic sclerosis: a prototypic multisystem fibrotic disorder. J Clin Invest 2007;117: 557–67.

7 Varga J. Antifibrotic therapy in scleroderma: extracellular or intracellular targeting of activated fibroblasts? Curr Rheumatol Rep 2004;6:164–70.

8 Baroni SS, Santillo M, Bevilacqua F *et al*. Stimulatory autoantibodies to the PDGF receptor in systemic sclerosis. N Engl J Med 2006;354:2667–76.

9 Dougados M, Betteridge N, Burmester GR *et al*. EULAR standardised operating procedures for the elaboration, evaluation, dissemination, and implementation of recommendations endorsed by the EULAR standing committees. Ann Rheum Dis 2004;63:1172–6.

10 Khanna D, Distler O, Avouac J *et al*. Measures of response in clinical trials of systemic sclerosis: the Combined Response Index for Systemic Sclerosis (CRISS) and Outcome Measures in Pulmonary Arterial Hypertension related to Systemic Sclerosis (EPOSS). J Rheumatol 2009; 36:2356–61.

11 Guyatt GH, Oxman AD, Vist GE *et al*. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336:924–6.

12 de Clerck LS, Dequeker J, Francx L *et al*. D-penicillamine therapy and interstitial lung disease in scleroderma. A long-term followup study. Arthritis Rheum 1987;30: 643–50.

13 Fuchs D, Fruchter L, Fishel B *et al*. Colchicine suppression of local inflammation due to calcinosis in dermatomyositis and progressive systemic sclerosis. Clin Rheumatol 1986; 5:527–30.

14 Scherbel AL, McCormack LJ, Layle JK. Further observations on the effect of dimethyl sulfoxide in patients with generalized scleroderma (progressive systemic sclerosis). Ann N Y Acad Sci 1967;141:613–29.

15 Tashkin DP, Elashoff R, Clements PJ *et al*. Effects of 1-year treatment with cyclophosphamide on outcomes at 2 years in scleroderma lung disease. Am J Respir Crit Care Med 2007;176:1026–34.

16 Henness S, Wigley FM. Current drug therapy for scleroderma and secondary Raynaud's phenomenon: evidence-based review. Curr Opin Rheumatol 2007;19:611–8.

17 Channick RN, Simonneau G, Sitbon O *et al*. Effects of the dual endothelin-receptor antagonist bosentan in patients with pulmonary hypertension: a randomised placebo-controlled study. Lancet 2001;358:1119–23.

18 Matucci-Cerinic M, Denton CP, Furst DE *et al*. Bosentan treatment of digital ulcers related to systemic sclerosis: results from the RAPIDS-2 randomised, double-blind, placebo-controlled trial. Ann Rheum Dis 2011;70:32–8.

19 Tashkin DP, Elashoff R, Clements PJ *et al*. Cyclophosphamide versus placebo in scleroderma lung disease. N Engl J Med 2006;354:2655–66.

20 Nannini C, West CP, Erwin PJ *et al*. Effects of CYC 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. Arthritis Res Ther 2008;10: R124.

21 Clements PJ, Furst DE, Wong WK *et al*. High-dose versus low-dose D-penicillamine in early diffuse systemic sclerosis: analysis of a two-year, double-blind, randomized, controlled clinical trial. Arthritis Rheum 1999;42:1194–203.

22 Seibold JR, Korn JH, Simms R *et al*. Recombinant human relaxin in the treatment of scleroderma. A randomized, double-blind, placebo-controlled trial. Ann Intern Med 2000;132:871–9.

23 Schulz KF, Altman DG, Moher D. CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340: c332.

24 Beyer C, Distler JHW, Allanore Y *et al*. EUSTAR biobanking: recommendations for the collection, storage and distribution of biospecimens in scleroderma research. Ann Rheum Dis 2011;70:1178–82.

25 Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Arthritis Rheum 1980;23:581–90.

26 Johnson SR, Fransen J, Khanna D *et al*. Validation of potential classification criteria for systemic sclerosis. Arthritis Care Res (Hoboken) 2012;64:358–67.

27 Avouac J, Fransen J, Walker UA *et al*. Preliminary criteria for the very early diagnosis of systemic sclerosis: results of a Delphi Consensus Study from EULAR Scleroderma Trials and Research Group. Ann Rheum Dis 2011;70: 476–81.

28 Cutolo M, Sulli A, Secchi ME *et al*. Nailfold capillaroscopy is useful for the diagnosis and follow-up of autoimmune rheumatic diseases A future tool for the analysis of microvascular heart involvement? Rheumatology 2006; 45(Suppl 4):iv43–6.

29 Medsger TA Jr, Silman AJ, Steen VD *et al*. A disease severity scale for systemic sclerosis: development and testing. J Rheumatol 1999;26:2159–67.

30 Boers M, Brooks P, Simon LS *et al*. OMERACT: an international initiative to improve outcome measurement in rheumatology. Clin Exp Rheumatol 2005;23(Suppl 39): S10–3.

31 Merkel PA, Clements PJ, Reveille JD *et al*. Current status of outcome measure development for clinical- trials in systemic sclerosis. Report from OMERACT 6. J Rheumatol 2003;30:1630–47.

32 Furst DE, Khanna D, Matucci-Cerenic M *et al*. Scleroderma—developing measures of response. J Rheumatol 2005;32:2477–80.

33 Khanna D, Hays RD, Maranian P *et al*. Reliability and validity of the University of California, Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract Instrument. Arthritis Rheum 2009;61:1257–63.

34 Khanna D, Furst DE, Maranian P *et al*. Minimally important differences of the UCLA Scleroderma Clinical Trial Consortium Gastrointestinal Tract Instrument. J Rheumatol 2011;38:1920–4.

35 Furst D, Khanna D, Matucci-Cerinic M *et al*. Systemic sclerosis—continuing progress in developing clinical measures of response. J Rheumatol 2007;34:1194–200.

36 Amjadi S, Maranian P, Furst DE *et al*. Course of the modified Rodnan skin thickness score in systemic sclerosis clinical trials: analysis of three large multicenter, double-blind, randomized controlled trials. Arthritis Rheum 2009;60:2490–8.

37 Khanna D, Furst DE, Hays RD *et al*. Minimally important difference in diffuse systemic sclerosis: results from the D-penicillamine study. Ann Rheum Dis 2006;65:1325–9.

38 Nash RA, McSweeney PA, Crofford LJ *et al*. High-dose immunosuppressive therapy and autologous hematopoietic cell transplantation for severe systemic sclerosis: long-term follow-up of the US multicenter pilot study. Blood 2007;110:1388–96.

39 Valentini G, Della Rossa A, Bombardieri S *et al*. European multicentre study to define disease activity criteria for systemic sclerosis. II. Identification of disease activity variables and development of preliminary activity indexes. Ann Rheum Dis 2001;60:592–8.

40 Wong WK, Boscardin WJ, Postlethwaite AE *et al*. Handling missing data issues in clinical trials for rheumatic diseases. Contemp Clin Trials 2011;32:1–9.