A randomised placebo-controlled phase 3 trial of tocilizumab in systemic sclerosis

Dinesh Khanna, MD,[1] Celia J F Lin, MD,[2] Daniel E Furst, MD,[3] Jonathan Goldin, MD,[3]

Grace Kim, PhD,[3] Masataka Kuwana, MD, PhD,[4] Yannick Allanore, PhD,[5] Marco Matucci-

Cerinic, MD,[6] Oliver Distler, MD,[7] Yoshihito Shima, MD,[8] Jacob M van Laar, MD,[9] Helen

Spotswood, PhD,[10] Bridget Wagner, MD,[2] Jeff Siegel, MD,[2] Angelika Jahreis, MD,[2*]

Christopher P Denton, FRCP[11*] for the focuSSced investigators[**]

*Drs. Jahreis and Denton contributed equally to this article.

**Investigators listed in the appendix.

[1]University of Michigan, Ann Arbor, MI, USA; [2]Genentech, South San Francisco, CA, USA;

[3]University of California, Los Angeles, Los Angeles, CA, USA; [4]Nippon Medical School,

Tokyo, Japan; [5]Paris Descartes University, Paris, France; [6]University of Florence, Florence,

Italy; [7]University of Zurich, Zurich, Switzerland; [8]Osaka University Graduate School of

Medicine, Osaka, Japan; [9]University Medical Center Utrecht, Utrecht, Netherlands; [10]Roche

Products Ltd, Welwyn Garden City, United Kingdom; [11]University College London, London,

UK

Corresponding author:
Professor Dinesh Khanna
Division of Rheumatology
Department of Internal Medicine
University of Michigan Scleroderma Program
300 North Ingalls Street
Ann Arbor, MI 48109, USA
khannad@med.umich.edu
734-936-5561

1

6

7

Research in Context

Evidence before this study

We searched PubMed for entries within the last 10 years using the terms (("systemic sclerosis" OR scleroderma) AND "interstitial lung disease" AND treatment) and reviewed the results to identify phase 3 trials for treatment of patients with systemic sclerosis–interstitial lung disease (SSc-ILD). Cyclophosphamide and mycophenolate mofetil demonstrated significant but modest effects on improving lung function in patients with SSc-ILD. A randomised controlled phase 2 trial of the interleukin-6 receptor inhibitor tocilizumab in systemic sclerosis (SSc), which preceded the current trial, showed no statistically significant effect of tocilizumab on skin thickness, but there was evidence of clinically relevant improvement in lung function with tocilizumab treatment. The multi-tyrosine kinase inhibitor nintedanib slowed the decline of lung function in patients with radiographically evident, established SSc-ILD. Nintedanib was recently approved in the United States to slow the rate of decline in pulmonary function in SSc-ILD. There is no approved disease-modifying therapy for the treatment of SSc, and treatment guidelines focus on the management of organ-specific manifestations.

Added value of this study

This is the first randomised controlled phase 3 trial of an interleukin-6 receptor antagonist in SSc. Although the primary skin fibrosis endpoint was not met, key secondary analysis of forced vital capacity and exploratory analysis of radiographically determined lung fibrosis suggest that tocilizumab treatment can confer clinically meaningful preservation of lung function and maintenance of pulmonary structure in patients with early diffuse SSc-ILD and elevated acute-phase reactants.

1

Implications of all the available evidence

The potential effect of tocilizumab in preserving lung function in patients with early SSc-ILD

in early SSc has important therapeutic implications.

ABSTRACT <<*300 of max 250 words*>>

Background We assessed skin fibrosis and systemic sclerosis–interstitial lung disease (SSc-ILD) in a phase 3 trial of tocilizumab, an anti–interleukin-6 receptor antibody, in systemic sclerosis.

Methods Participants were randomly assigned 1:1 to receive double-blind weekly tocilizumab 162 mg or placebo subcutaneously for 48 weeks (other immunomodulatory therapy was not permitted at baseline). The primary endpoint was the difference in change from baseline to week 48 in modified Rodnan skin score (mRSS). Percent predicted forced vital capacity (ppFVC) at week 48, time-to-treatment-failure, and patient-/physician-reported outcomes were secondary endpoints.

Findings Among 104 tocilizumab-treated and 106 placebo-treated participants, the least squares mean (LSM) change from baseline to week 48 in mRSS was –6·14 and –4·41, respectively (adjusted difference, –1·73 [95% CI –3·78 to 0·32]; p=0·10). The shift in distribution of change from baseline in ppFVC at week 48 favoured tocilizumab (van Elteren nominal p=0·002 vs placebo) with a difference in LSM change of 4·2 (95% CI 2·0–6·4; nominal p=0·0002). Time-to-treatment-failure favoured tocilizumab (hazard ratio 0·63 [95% CI 0·37 to 1·06]; nominal p=0·08). LSM (95% CI) differences between tocilizumab and placebo in change from baseline to week 48 in Health Assessment Questionnaire–Disability Index (–0·05 [-0·19 to 0·09]), patient-global (–2·4 [–8·6 to 3·70]), and physician-global (–2·5 [–8·7 to 3·8]) visual analogue scale assessments were not statistically significant. Infections were the most common adverse events (tocilizumab, 54/104 [51·9%]; placebo 53/106 [50·0%]). Serious adverse events were reported in 13 tocilizumab-treated participants and 18 placebo-treated participants; primarily infections (tocilizumab, 3 events; placebo, 8 events) and cardiac events (tocilizumab, 2 events; placebo, 7 events).

1

2   **Interpretation** The primary skin fibrosis endpoint was not met. Secondary and exploratory

3   results suggest that tocilizumab preserves lung function in patients with early SSc-ILD and

4   elevated acute-phase reactants. Safety was consistent with the tocilizumab safety profile.

5   **Funding** F. Hoffmann-La Roche Ltd.

6   **Trial registration:** ClinicalTrials.gov, NCT02453256

1 Introduction

2 Systemic sclerosis (SSc) is a rare, severe disease,[1] and up to 60% of patients diagnosed with

3 SSc die of it.[2,3] Pulmonary complications, such as interstitial lung disease (ILD), are the

4 primary causes of death,[1,4,5] and decline in forced vital capacity (FVC) is associated with

5 increased mortality in patients with SSc-ILD.[5] The multi–tyrosine kinase inhibitor nintedanib

6 slowed the decline of lung function in a study in patients with radiographically evident SSc-

7 ILD.[6] However, treatment of SSc-ILD is limited to managing organ-specific complications.[7]

8 Circulating levels of interleukin-6 (IL-6) are elevated in patients with SSc[8] and are associated

9 with the development of skin fibrosis and SSc-ILD.[9-11] Early studies suggested that inhibition

10 of IL-6 signalling via IL-6 receptor blockade with tocilizumab might reduce skin fibrosis in

11 patients with SSc.[12,13] FaSScinate, a phase 2 randomised controlled trial, investigated the

12 efficacy and safety of tocilizumab in SSc.[14] The primary endpoint was not met in faSScinate;

13 the least squares mean (LSM) change in modified Rodnan skin score (mRSS) from baseline

14 to week 24 was –3·92 with tocilizumab and –1·22 with placebo (difference –2·70; 95%

15 confidence interval [CI] –5·85 to 0·45; p=0·09).[14] Mechanistic support for an antifibrotic

16 effect of tocilizumab in faSScinate came from analysis of explant dermal fibroblasts that

17 highlighted reversal of the activated fibrotic phenotype after 24 weeks of treatment.[15] In a

18 prespecified exploratory analysis in faSScinate, significantly fewer participants treated with

19 tocilizumab than placebo experienced decline in lung function accompanied by reduction in

20 expression of profibrotic M-2 macrophage-associated genes, suggesting that tocilizumab

21 might be able to preserve lung function.[14] These data, together with the high need for

22 effective treatments for SSc patients with severe skin and lung manifestations, supported

23 investigation of tocilizumab in a phase 3 trial. Therefore, the phase 3 randomised controlled

24 trial focuSSced was conducted to assess the effect of tocilizumab and placebo treatment on

25 change in mRSS, with impact on lung function as a key secondary objective.

7

Participants and methods

Study design

focuSSced was a multicentre, randomised, double-blind, placebo-controlled phase 3 trial (ClinicalTrials.gov, NCT02453256) conducted at 75 sites in 20 countries across Europe, North America, Latin America, and Japan. Eligible participants were randomly assigned 1:1 to receive subcutaneous injections of tocilizumab 162 mg or placebo weekly for a 48-week, double-blind period, followed by a 48-week, open-label tocilizumab period. Interim futility analysis conducted at week 24, for which the sponsor remained blinded, was conducted by an independent data coordinating centre. Immunomodulatory therapy could be added to study medication from week 16 for participants who experienced a decline in percent predicted FVC (ppFVC) or from week 24 for those who experienced worsened skin thickening or other significant SSc complications. A dual-assessor approach was used to prevent potential unblinding owing to knowledge of laboratory results. The study was conducted in accordance with the principles of the Declaration of Helsinki and Good Clinical Practice, and approval was obtained from the investigators' independent ethics committees or institutional review boards. All participants provided written informed consent to participate in the study.

Randomisation and masking

Participants were randomly assigned using and interactive voice-based or web-based response system. Randomisation was centralised and stratified by serum IL-6 levels at screening (<10 or ≥10 pg/mL) because lower IL-6 levels were associated with a more favourable outcome in change from baseline in mRSS in an analysis of data from the phase 2

study of tocilizumab in SSc (faSScinate). Participants and study sponsor personnel were blinded to study treatment.

Participants

Adults with diffuse cutaneous SSc, classified according to 2013 American College of Rheumatology/European League Against Rheumatism criteria,[16] of ≤60 months' duration (from first non–Raynaud phenomenon manifestation) and mRSS 10–35 units at screening were eligible. Participants had to have elevated acute-phase reactant levels (≥1 of the following: CRP ≥6 mg/L, ESR ≥28 mm/h, or platelet count ≥330 × 10^9/L) and active disease defined as ≥1 of the following at screening: disease duration ≤18 months, mRSS increase ≥3 units, or involvement of one new body area and mRSS increase ≥2 units, or involvement of two new body areas (each within the previous 6 months), and ≥1 tendon friction rub. Additional eligibility criteria are shown in Supplementary Appendix 2.

Outcomes

The primary efficacy endpoint was the difference in change from baseline in mRSS at week 48.[17] Key secondary efficacy endpoints were difference in distribution of change from baseline to week 48 in ppFVC (assessed according to standardised methods and reviewed centrally by readers masked to treatment, analysed by van Elteren test as the preplanned FVC outcome), time to treatment failure (defined as time of death, time to decline in ppFVC >10%, relative increase in mRSS >20% and ≥5 mRSS points, or occurrence of a predefined and adjudicated SSc-related serious complication [Supplementary Table S1]), and Health Assessment Questionnaire–Disability Index (HAQ-DI). Other secondary endpoints were patient global assessment, and physician global assessment. Exploratory endpoints included

1 proportions of participants with ≥10% decline (worsening) in ppFVC and change from

2 baseline in high-resolution computed tomography (HRCT) of quantitative lung fibrosis–most

3 affected lobe (QLF-LM) at week 48 (an HRCT read was planned at baseline and week 48 for

4 all participants), the American College of Rheumatology provisional Composite Response

5 Index in Systemic Sclerosis (ACR-CRISS), and other patient-reported outcomes

6 (Supplementary Appendix 3). HRCT quantitative lung fibrosis–whole lung (QLF-WL) and

7 quantitative interstitial lung disease–whole lung (QILD-WL) were post hoc analyses.

8 Additional exploratory analyses were performed for the subset of participants who had ILD at

9 baseline on visual read of HRCT (hereafter referred to as SSc-ILD). ILD was identified

10 visually post hoc by a thoracic radiologist (J. G.) using a diagnostic algorithm for SSc as the

11 presence of ground-glass opacification and/or fibrosis with a basal predominance. Potential

12 causes other than SSc for the pattern of ground-glass opacification were excluded. Other post

13 hoc analyses are described in Supplementary Appendix 3. Safety was assessed as treatment-

14 emergent adverse events (AEs) according to MedDRA system organ classification and was

15 graded according to the National Cancer Institute Common Toxicity Criteria for Adverse

16 Events, version 4·0.

17

18 **Statistical analysis**

19 One hundred five participants per treatment group provided power in the range of >75% to

20 80% (depending on an estimated participant dropout rate of 20% to 15%) to detect a

21 between-group difference of 3·55 mRSS units (the treatment effect in the phase 2 faSScinate

22 trial) in change from baseline to week 48 (common standard deviation of 8·43 units using a

23 two-group *t* test and a 5% two-sided significance level). Efficacy was assessed in the

24 intention-to-treat population, which included all randomly assigned participants who received

≥1 dose of study treatment according to their originally assigned treatment group. Safety was

assessed in the safety population, which included all participants who received ≥1 dose of

study treatment and had ≥1 postdose safety assessment. A statistical testing hierarchy was

conducted based on significance for the primary endpoint at the 5% level (Supplementary

Figure S1). Analysis of the primary endpoint was performed using mixed model for repeated

measures. An unstructured covariance matrix was used to model within-participant errors,

and the Kenward-Roger approximation was used to estimate the denominator degrees of

freedom. The model included fixed categorical effects for treatment, visit, IL-6 stratification

criteria, IL-6 level at screening-by-visit interaction, and treatment-by-visit interaction, as well

as the continuous covariates of baseline mRSS and baseline mRSS-by-visit interaction; there

was no imputation for missing data. The study was not stratified by site because the large

number of sites with small participant numbers predicted at each site made this impractical;

site was therefore not included in the primary analysis. The LSM, differences in mean, 95%

CI, and p value were reported with the significance test based on a two-sided alpha of 0·05.

LSMs and differences in means for other endpoints, including FVC, used similar methods.

The Cochran-Mantel-Haenszel test adjusted for IL-6 stratification factors (<10 or ≥10 pg/mL)

was used for binary mRSS endpoints, and participants with missing mRSS assessments at

week 48 were considered nonresponders. Change from baseline in FVC and HRCT variables

was assessed primarily with the use of nonparametric analysis (van Elteren test; the 95% CI

for the between-arm difference in medians was derived using bootstrapping) assuming a non-

normal distribution stratified by screening IL-6 level. Time to treatment failure was

summarised descriptively by Kaplan–Meier curves, and treatment groups were compared

using a Cox proportional hazards model adjusted for IL-6 stratification. An analysis of log

cumulative hazard plotted against log survival time gave approximately parallel curves

accounting for treatment group and baseline IL-6 stratification, indicating that a proportional

hazards assumption was appropriate. For time to treatment failure, data were censored from

the time of discontinuation for participants who discontinued the study before week 48. No

analyses included censoring for participants who initiated immunomodulatory therapy or

discontinued study drug.

Role of the funding source

F. Hoffmann-La Roche Ltd. was involved in the design and conduct of the study; collection,

management, analysis, and interpretation of the data; writing and review of the manuscript;

and decision to submit the manuscript for publication. DK, CJFL, HS, and BW had access to

the raw data. The corresponding author had full access to all of the data and the final

responsibility to submit for publication.

Results

Participants

Among 343 participants screened, 212 were randomly assigned to receive weekly

subcutaneous placebo (n=107) or tocilizumab 162 mg (n=105). The first participant was

randomly assigned on November 20, 2015, and the last participant completed the week 48

assessment on January 15, 2018. Ninety-three participants (86·9%) in the placebo group and

95 participants (90·5%) in the tocilizumab group completed 48 weeks (Figure 1). The

intention-to-treat and safety populations comprised 106 participants in the placebo group and

104 participants in the tocilizumab group.

Baseline demographics and disease characteristics were similar between treatment groups for

the intention-to-treat population and for the subgroup of participants who had SSc-ILD at

1   baseline according to HRCT visual read (Table 1; Supplementary Tables S2 and S3). Most

2   participants were female (n=171/210; 81·4%), and the mean (SD) age was 48·2 (12·4) years.

3   Median disease duration was <2 years, and skin involvement was moderate to severe with a

4   mean (SD) baseline mRSS of 20·4 (7·0) in the placebo group and 20·3 (6·7) in the

5   tocilizumab group. Participants had normal to mild impairment in lung function at baseline

6   (mean [SD] ppFVC was 83·9 [15·0] in the placebo group and 80·3 [14·4] in the tocilizumab

7   group, and percent predicted diffusing capacity for carbon monoxide was 76.8 [18·6] and

8   74·4 [19·2], respectively), and 136/210 (64·8%) had evidence of SSc-ILD on HRCT.

9

10   By week 48, immunomodulating therapy was initiated by 22 participants (20·8%) in the

11   placebo group and nine participants (8·7%) in the tocilizumab group (Figure 1); most

12   participants started immunomodulating therapy after week 36.

13

14   **Efficacy**

15   The primary endpoint of change from baseline in mRSS at week 48 for tocilizumab versus

16   placebo was not met, though participants treated with tocilizumab had a numerically greater

17   reduction in skin sclerosis after 48 weeks; LSM change from baseline to week 48 in mRSS

18   was –4·41 in the placebo group and –6·14 in the tocilizumab group (adjusted difference in

19   LSM, –1·73 [95% CI –3·78 to 0·32]; p=0·10) (Table 2; Figure 2). Because the primary

20   analysis did not meet statistical significance at the 5% level, none of the secondary endpoints

21   were considered to have achieved statistical significance according to the hierarchy

22   (Supplementary Figure S1), and all p values for secondary, exploratory, and post hoc

23   analyses were nominal.

At week 48, the LSM change from baseline in ppFVC was –4·6 in the placebo group and –

0·4 in the tocilizumab group (difference, 4·2 [95% CI 2·0 to 6·4]; nominal p=0·0002) (Figure

3A; Table 3), and the absolute LSM change was –190 mL and –24 mL, respectively

(difference, 167 mL (95% CI 83 to 250); nominal p=0·0001). The difference in change from

baseline in ppFVC was confirmed in sensitivity analyses (Supplementary Table S4). Based

on prespecified exploratory analysis, the proportion of participants who experienced absolute

decline in FVC ≥10% was 16·5% (15/91) in the placebo group and 5·4% (5/93) in the

tocilizumab group. There was a clinically meaningful[18] shift in the distribution of change

from baseline in ppFVC at week 48 (key secondary endpoint) favouring tocilizumab (van

Elteren nominal p=0·002 vs placebo) (Figure 3B). Preplanned exploratory analysis of QLF-

LM and post hoc analysis of QLF-WL and QILD-WL showed numeric improvements in lung

fibrosis in participants treated with tocilizumab, supporting FVC results and consistent in

participants with SSc-ILD at baseline (Table 3; Supplementary Figure S2). Among

participants with SSc-ILD on visual read at baseline, the LSM change from baseline to week

48 in ppFVC was –6·40 in the placebo group and 0·07 in the tocilizumab group (difference,

6·47 [95% CI 3·43 to 9·50]; nominal p<0·0001), and the absolute LSM change was –255 mL

and –14 mL, respectively (difference: 241 mL [95% CI 124 to 358]; nominal p<0·0001)

(Figure 3C; Table 3). Among the participants with SSc-ILD on visual read at baseline and

available week 48 FVC data, the proportion who experienced absolute decline in FVC ≥10%

to week 48 was 25·0% (14/56) in the placebo group and 8·5% (5/59) in the tocilizumab group

(Figure 3D).

Kaplan–Meier analysis of time to treatment failure (key secondary endpoint) favoured

tocilizumab over placebo; hazard ratio adjusted for baseline IL-6 stratification factors was

0·63 (95% CI 0·37 to 1·06); nominal p=0·08 (Figure 4); unadjusted hazard ratio was 0·58

(95% CI 0·34 to 0·98). At week 48, numerically fewer participants in the tocilizumab group than the placebo group experienced treatment failure in any of its components (Table 2).

There was no difference between placebo and tocilizumab for patient- or physician-reported outcomes of HAQ-DI, patient global assessment, or physician global assessment (secondary endpoints) or for Functional Assessment of Chronic Illness Therapy–Fatigue, Scleroderma HAQ, or Saint George's Respiratory Questionnaire (exploratory endpoints) at week 48 (Table 2).

Safety

Most participants experienced ≥1 AE during the study (Table 4; Supplementary Table S5). Infections and infestations were the most frequently reported AEs for both groups (53/106 [50·0%] placebo, 54/104 [51·9%] tocilizumab). Thirty serious AEs (SAEs) were reported in 18/106 participants (17·0%) in the placebo group compared with 14 SAEs in 13/104 participants (12·5%) in the tocilizumab group; this difference was primarily driven by more serious infections (placebo, 8 events in 7 participants; tocilizumab, 3 events in 2 participants) and serious cardiac events (placebo, 7 events in 6 participants; tocilizumab, 2 events in 2 participants) in the placebo group. All serious infections were grade 3 in severity except for a soft tissue infection reported in a placebo-treated participant, which was grade 4. Three placebo-treated participants developed pneumonia and one developed a respiratory tract infection; two of these participants had a history of ILD, two were former smokers, and two were receiving prednisone. None of the serious infections reported in the placebo arm started after initiation of escape therapy. Infected skin ulcers were reported in 12/106 participants (11·3%) in the placebo group and 15/104 participants (14·4%) in the tocilizumab group; one event in the placebo group was an SAE. No demyelinating AEs, SAEs or medically

significant hepatic AEs or bleeding events, gastrointestinal perforations, stroke, or

anaphylactic reactions were reported. Four participants died during the study: three in the

placebo group (chronic cardiac failure, myocarditis, myocardial infarction) and one in the

tocilizumab group (unknown cause); none of the events were considered related to study

treatment. Laboratory abnormalities are shown in Supplementary Table S6.

Discussion

The primary mRSS endpoint was not met in this phase 3 trial of tocilizumab in early, active

SSc. This suggests that there was no difference in change in skin thickness between

participants treated with tocilizumab and those treated with placebo after 48 weeks. However,

secondary FVC results suggest stabilization of lung function in participants who received

tocilizumab treatment; this replicates the effect observed in the phase 2 faSScinate trial[14] and

the exploratory and post hoc HRCT results confirm the antifibrotic effect of tocilizumab[15] in

radiologically evident lung fibrosis. In addition to previous studies of explant dermal

fibroblasts from the faSScinate clinical trial,[15] a direct effect of IL-6 is supported by recent

reports of reduced myofibroblast activity after inhibition of STAT3, a putative link between

IL-6 and TGFβ intracellular signalling.[19-21] focuSSced is the first placebo-controlled phase 3

clinical trial to assess structural and functional pulmonary changes in early, active SSc. The

effect of tocilizumab on disease progression was supported by numerical improvement in

time to treatment failure and the fact that 21% of participants in the placebo group received

immunomodulating rescue therapy compared with 9% in the tocilizumab group. Safety was

consistent with the safety profile of tocilizumab and complications of SSc, and no new safety

concerns emerged.

1   ILD is one of the leading causes of death in patients with early, diffuse SSc. Current clinical

2   practice, supported by treatment recommendations,[7,22] is to treat patients with SSc-ILD after

3   clinically significant disease, defined by symptoms and evidence of restrictive lung disease,

4   has developed. This practice is based on the fact that most patients with early SSc will not

5   develop clinically meaningful progressive disease and on a limited understanding of risk

6   factors for progression of ILD.

7   Putative predictive factors for progression include diffuse cutaneous phenotype, anti–Scl-70

8   (anti-topoisomerase I) positivity, elevated acute-phase reactants, and ethnicity.[23] The

9   focuSSced eligibility criteria enriched the study population for participants with these factors,

10   and indeed 65% had evidence of ILD on baseline HRCT. The clinically meaningful[18] shift in

11   distribution of change from baseline in ppFVC favouring tocilizumab over placebo and the

12   observation that fewer tocilizumab-treated than placebo-treated participants experienced a

13   decline ≥10% in ppFVC suggest that tocilizumab can preserve lung function. This trial

14   confirms the data from faSScinate; additionally, in the focuSSced trial, FVC benefit was

15   supported by quantitative HRCT analysis showing stabilisation of ILD; post hoc HRCT

16   results for the subset of participants with SSc-ILD also supported FVC results.

17   In the trials of tocilizumab in early SSc, the difference between tocilizumab and placebo was

18   120 mL in the phase 2 faSScinate trial[14] and 167 mL overall and 238 mL among participants

19   with baseline SSc-ILD in the phase 3 focuSSced trial. focuSSced participants were selected

20   for worsening skin disease and increased acute-phase reactants; therefore, most had early,

21   mild ILD, which might explain the lack of statistical differences we observed in patient-

22   reported outcomes because participants might not have had overt respiratory symptoms. The

23   mean change in QLF-LM was 1·4% with tocilizumab treatment in our study and –2·6% with

24   cyclophosphamide treatment in the scleroderma lung study (SLS-1),[24] possibly because

25   focuSSced participants had early, mild lung involvement whereas SLS-1 participants had to

1   have definitive evidence of lung disease based on HRCT or bronchoalveolar lavage and were

2   therefore likely to have had more severe ILD.

3   Our findings relate to recently reported results from the Safety and Efficacy of Nintedanib in

4   Systemic Sclerosis (SENSCIS) trial in which nintedanib slowed the progression of

5   established lung fibrosis in a large cohort of participants with SSc-ILD.[6] The SENSCIS trial

6   reported a difference of 46·4 mL in absolute FVC decline over 1 year in favour of nintedanib

7   (–54·6 mL) versus placebo (–101·0 mL), with a 1·2% difference in mean change in ppFVC.

8   Results from SENSCIS provide context for our findings but cannot be directly compared

9   because of substantial differences between the study designs and populations of the trials.

10   focuSSced recruited a population with earlier stage SSc, and all participants had diffuse,

11   inflammatory, progressive skin involvement and were therefore at high risk for ILD, whereas

12   SENSCIS recruited a population with established, clinically relevant ILD, including

13   participants in the limited and diffuse SSc subsets. Indeed, the SENSCIS cohort had mean

14   baseline ppFVC of approximately 72% and lung fibrosis on HRCT of 35% to 37%, whereas

15   the focuSSced cohort had mean baseline ppFVC of 82% and lung fibrosis on HRCT of 2% to

16   17%.

17   A higher rate of progression, assessed by absolute decline in FVC, was observed in the

18   focuSSced cohort, who had early, active diffuse SSc enriched for skin activity, which is

19   consistent with an association between progressive skin fibrosis and FVC decline.[25] The

20   difference in mean change in ppFVC between tocilizumab and placebo was 4·2% overall and

21   6·4% in the ILD subgroup in focuSSced. In the SLS-1 trial, there was a 2·5% difference in

22   ppFVC between cyclophosphamide and placebo,[26] and in SENSCIS, there was a 1·2%

23   difference in ppFVC between nintedanib and placebo. These differences might reflect the

24   earlier disease of focuSSced participants. The absolute FVC decline observed with placebo

25   treatment in focuSSced is comparable to the rate of decline observed in idiopathic pulmonary

1   fibrosis.[27,28] Our results might indicate that IL-6 is a more important driver of lung fibrosis

2   progression in early SSc, but this requires further investigation.

3   The focuSSced trial did not meet its primary endpoint: change in skin sclerosis measured by

4   mRSS. Skin thickness was chosen as the primary endpoint because it is universally present in

5   SSc and can have a profound impact on function and quality of life[29] and because mRSS is a

6   feasible, reliable, and valid outcome measure (including for sensitivity to change).[30] The

7   difference in change in mRSS results between the phase 2 faSScinate trial and the phase 3

8   focuSSced trial might reflect differences in inclusion criteria and a possible effect of

9   unknown genetic differences and molecular heterogeneity on mRSS. Several recently

10   completed placebo-controlled trials have highlighted the limitation of mRSS as a primary

11   outcome given the variable natural history of SSc and a tendency for improvement in placebo

12   treatment arms,[31,32] which was also observed in focuSSced. It is likely that other endpoints,

13   such as the composite ACR-CRISS responder index,[33] will emerge as more robust and

14   reliable for testing disease-modifying therapies in SSc. Our ACR-CRISS data highlight the

15   importance of global assessment in a multisystem heterogeneous disease such as systemic

16   sclerosis. At present, our findings, together with those from the SENSCIS and faSScinate

17   trials, support the use of FVC as a robust endpoint to demonstrate a treatment effect in SSc-

18   ILD and show that this is not associated with a significant difference in mRSS.

19

20   The design of the focuSSced trial offers strengths and limitations. This phase 3 trial was

21   powered to determine the difference in mRSS as the primary endpoint assuming a dropout

22   rate of 15% and a between-group difference of 3.55 mRSS units based on the effect observed

23   in the phase 2 faSScinate study. Clinical significance and statistical significance were

24   considered in powering the focuSSced study to avoid statistical significance associated with a

1   negligible treatment effect. A difference of 1.7 mRSS units between tocilizumab and placebo

2   (primary endpoint) was observed, possibly reflecting the heterogeneity of SSc and a strong

3   placebo effect. This difference was not statistically significant; therefore, any differences in

4   secondary endpoints, including FVC, could not be considered statistically significant despite

5   the strength of the evidence. Missing data from dropouts or missed assessments were

6   accounted for in the analyses according to the methods described for handling of missing

7   data. Sensitivity analyses for FVC showed that the FVC results were robust to different

8   assumptions regarding the missing data. Patients in focuSSced and faSScinate had early

9   disease and mild lung involvement, which limited comparison to other trials, but provided

10   novel results in this patient population that support further investigation. Examining FVC and

11   composite endpoints, such as ACR-CRISS as primary endpoints is an important consideration

12   for future trials. A limitation of the lung findings is that DLCO was measured using the

13   investigators' own equipment, which limits the comparability and reliability of the DLCO

14   data.

15   The results for FVC and other secondary outcome measures should be interpreted with

16   caution because the primary outcome was not statistically significant. However, the effect on

17   FVC supported by two randomised controlled trials shows a clinically meaningful impact on

18   preservation of lung function. In conclusion, the skin primary endpoint was negative in this

19   phase 3 trial, but results suggest that tocilizumab preserves lung function in early diffuse

20   cutaneous SSc.

21

Contributors

Dr Khanna had full access to all data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: DK, CPD, AJ, CJFL, HS, JS, DEF

Acquisition, analysis, or interpretation of data: All authors

First draft of manuscript: DK, CPD, AJ, CJFL, HS

Critical revision of the manuscript for important intellectual content: All authors

Statistical analysis: HS

Approval of final version for submission: All authors

Declaration of interests

D. Khanna has ownership interest in Eicos Sciences; has received grants from the National Institutes of Health (NIAID and NIAMS); has received consulting fees from Actelion, AstraZeneca, Bayer, Bristol-Myers Squibb, Boehringer Ingelheim, ChemomAb, Corbus, CSL Behring, Cytori, GlaxoSmithKline, Horizon, Pfizer, Regeneron, Roche/Genentech, Sanofi Aventis, and UCB Pharma.

C. J. F. Lin owns stock in and is an employee of Genentech (a member of the Roche group).

D. E. Furst has received research grants from Actelion, Amgen, BMS, Corbus, Galapagos, GSK, National Institutes of Health, Novartis, Pfizer, Roche/Genentech, and Sanofi and has received consulting fees from Actelion, Amgen, BMS, Corbus, Galapgos, Novartis, and Pfizer.

J. Goldin has nothing to disclose.

Data Sharing Statement

1   Qualified researchers may request access to data through the clinical study data request

2   platform (www.clinicalstudydatarequest.com). Further details on Roche's criteria for eligible

3   studies are available here (https://clinicalstudydatarequest.com/Study-Sponsors/Study-

4   Sponsors-Roche.aspx). For further details on Roche's Global Policy on the Sharing of

5   Clinical Information and how to request access to related clinical study documents, see here

6   (https://www.roche.com/research_and_development/who_we_are_how_we_work/clinical_tri

7   als/our_commitment_to_data_sharing.htm).

8

References

1   Denton CP, Khanna D. Systemic sclerosis. *Lancet* 2017; **390:** 1685–99.

2   Nihtyanova SI, Schreiber BE, Ong VH, et al. Prediction of pulmonary complications and long-term survival in systemic sclerosis. *Arthritis Rheumatol* 2014; **66:** 1625–35.

3   Elhai M, Meune C, Boubaya M, et al. Mapping and predicting mortality from systemic sclerosis. *Ann Rheum Dis* 2017; **76:** 1897–905.

4   Komocsi A, Vorobcsuk A, Faludi R, et al. The impact of cardiopulmonary manifestations on the mortality of SSc: a systematic review and meta-analysis of observational studies. *Rheumatology* 2012; **51:** 1027–36.

5   Goh NS, Hoyles RK, Denton CP, et al. Short-term pulmonary function trends are predictive of mortality in interstitial lung disease associated with systemic sclerosis. *Arthritis Rheumatol* 2017; **69:** 1670–8.

6   Distler O, Highland KB, Gahlemann M, et al. Nintedanib for systemic sclerosis-associated interstitial lung disease. *N Engl J Med* 2019; **380:** 2518–28.

7   Kowal-Bielecka O, Fransen J, Avouac J, et al. Update of EULAR recommendations for the treatment of systemic sclerosis. *Ann Rheum Dis* 2017; **76:** 1327–39.

8   Kadono T, Kikuchi K, Ihn H, Takehara K, Tamaki K. Increased production of interleukin 6 and interleukin 8 in scleroderma fibroblasts. *J Rheumatol* 1998; **25:** 296–301.

9   Sato S, Hasegawa M, Takehara K. Serum levels of interleukin-6 and interleukin-10 correlate with total skin thickness score in patients with systemic sclerosis. *J Dermatol Sci* 2001; **27:** 140–6.

10 Khan K, Xu S, Nihtyanova S, et al. Clinical and pathological significance of interleukin 6 overexpression in systemic sclerosis. *Ann Rheum Dis* 2012; **71:** 1235–42.

11 De Laurentis A, Sestini P, Pantelidis P, et al. Serum interleukin 6 is predictive of early functional decline and mortality in interstitial lung disease associated with systemic sclerosis. *J Rheumatol* 2013; **40:** 435–46.

12 Kitaba S, Murota H, Terao M, et al. Blockade of interleukin-6 receptor alleviates disease in mouse model of scleroderma. *Am J Pathol* 2012; **180:** 165–76.

13 Shima Y, Kuwahara Y, Murota H, et al. The skin of patients with systemic sclerosis softened during the treatment with anti-IL-6 receptor antibody tocilizumab. *Rheumatology* 2010; **49:** 2408–12.

14 Khanna D, Denton CP, Jahreis A, et al. Safety and efficacy of subcutaneous tocilizumab in adults with systemic sclerosis (faSScinate): a phase 2, randomised, controlled trial. *Lancet* 2016; **387:** 2630–40.

15 Denton CP, Ong VH, Xu S, et al. Therapeutic interleukin-6 blockade reverses transforming growth factor-beta pathway activation in dermal fibroblasts: insights from the faSScinate clinical trial in systemic sclerosis. *Ann Rheum Dis* 2018; **77:** 1362–71.

16 van den Hoogen F, Khanna D, Fransen J, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2013; **72:** 1747–55.

17 Furst DE, Clements PJ, Steen VD, et al. The modified Rodnan skin score is an accurate reflection of skin biopsy thickness in systemic sclerosis. *J Rheumatol* 1998; **25:** 84–8.

18 Kafaja S, Clements PJ, Wilhalme H, et al. Reliability and minimal clinically important differences of forced vital capacity: results from the Scleroderma Lung Studies (SLS-I and SLS-II). *Am J Respir Crit Care Med* 2018; 197**:** 644–52.

19 Papaioannou I, Xu S, Denton CP, Abraham DJ, Ponticos M. STAT3 controls COL1A2 enhancer activation cooperatively with JunB, regulates type I collagen synthesis posttranscriptionally, and is essential for lung myofibroblast differentiation. *Mol Biol Cell* 2018; **29**: 84-95.

20 Zehender A, Huang J, Gyorfi AH, et al. The tyrosine phosphatase SHP2 controls TGFβ-induced STAT3 signaling to regulate fibroblast activation and fibrosis. *Nat Commun* 2018; **9**: 3259.

21 Chakraborty D, Sumova B, Mallano T, et al. Activation of STAT3 integrates common profibrotic pathways to promote fibroblast activation and tissue fibrosis. *Nat Commun* 2017; **8**: 1130.

22 Denton CP, Hughes M, Gak N, et al. BSR and BHPR guideline for the treatment of systemic sclerosis. *Rheumatology* 2016; **55:** 1906–10.

23 Roofeh D, Jaafar S, Vummidi D, Khanna D. Management of systemic sclerosis-associated interstitial lung disease. *Curr Opin Rheumatol* 2019; **31:** 241–9.

24 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:** 2455–65.

25 Wu W, Jordan S, Graf N, et al. Progressive skin fibrosis is associated with a decline in lung function and worse survival in patients with diffuse cutaneous systemic sclerosis in

the European Scleroderma Trials and Research (EUSTAR) cohort. *Ann Rheum Dis* 2019; **78:** 648–56.

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

27 Richeldi L, du Bois RM, Raghu G, et al. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med* 2014; **370:** 2071–82.

28 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:** 2083–92.

29 Wiese AB, Berrocal VJ, Furst DE, et al. Correlates and responsiveness to change of measures of skin and musculoskeletal disease in early diffuse systemic sclerosis. *Arthritis Care Res* 2014; **66:** 1731–9.

30 Khanna D, Furst DE, Clements PJ, et al. Standardization of the modified Rodnan skin score for use in clinical trials of systemic sclerosis. *J Scleroderma Relat Disord* 2017; **2:** 11–8.

31 Distler O, Allanore Y, Denton CP, et al. Riociguat in patients with early difuse cutaneous systemic sclerosis: a randomized, double-blind, placebo-controlled phase IIb study (RISE-SSc). *Arthritis Rheumatol* 2018; **70**(suppl 10). Abstract 903.

32 Khanna D, Spino C, Johnson S, et al. Abatacept in early diffuse cutaneous systemic sclerosis—results of a phase 2 investigator-initiated, multicenter, double-blind randomized placebo-controlled trial. *Arthritis Rheumatol* 2019; doi: 10.1002/art.41055. [Epub ahead of print].

1 33 Khanna D, Berrocal VJ, Giannini EH, et al. The American College of Rheumatology

2   Provisional Composite Response Index for Clinical Trials in Early Diffuse Cutaneous

3   Systemic Sclerosis. *Arthritis Rheumatol* 2016; **68:** 299–311.

1   **FIGURE LEGENDS**

2   *Figure 1:* **Screening, randomisation, and follow-up**

3   The most frequent immunomodulating treatment received was mycophenolate mofetil (13

4   participants [12·3%] in the placebo group; five participants [4·4%] in the TCZ group)

5   followed by methotrexate (four participants [3·8%] in the placebo group; three participants

6   [2·9%] in the TCZ group). Only deaths reported as the reason for withdrawal are shown; two

7   additional participants in the placebo arm withdrew for other reasons (one not stated and one

8   because of an adverse event) before they died. QW, weekly; SC, subcutaneous; TCZ,

9   tocilizumab.

10

11 *Figure 2:* **Mean change from baseline in mRSS (ITT population)**

12 Mixed-model repeated measures analysis was implemented that included the fixed

13 categorical effects of treatment, visit, IL-6 stratification (<10 or ≥10 pg/mL at screening), IL-

14 6 level at screening-by-visit interaction, and treatment-by-visit interaction, as well as the

15 continuous covariates of baseline score and baseline score-by-visit interaction. BL, baseline;

16 IL-6, interleukin-6; ITT, intention-to-treat; LSM, least squares mean; mRSS, modified

17 Rodnan skin score; PBO, placebo; TCZ, tocilizumab.

18

19 *Figure 3:* **Cumulative distribution (A, B) and mean change from baseline (C, D) for**

20 **ppFVC at week 48**

21 Data are shown (A, C) for all participants and (B, D) for participants with SSc-ILD at

22 baseline (subset of participants who had ILD on visual read of HRCT by a thoracic

23 radiologist). A mixed model repeated measures analysis was implemented that included the

24 fixed categorical effects of treatment, visit, IL-6 stratification (<10 or ≥10 pg/mL at

25 screening), IL-6 level at screening-by-visit interaction, and treatment-by-visit interaction, as

well as the continuous covariates of baseline score and baseline score-by-visit interaction.

Change from baseline was assessed using nonparametric analysis assuming a normal

distribution. HRCT, high-resolution computed tomography; IL, interleukin; ILD, interstitial

lung disease; PBO, placebo; ppFVC, percent predicted forced vital capacity; SSc, systemic

sclerosis; TCZ, tocilizumab.

Figure 4: **Kaplan-Meier analysis of time-to-treatment failure (ITT population).**

Treatment groups were compared using a Cox proportional hazards model adjusted for

baseline IL-6 stratification factors (<10 pg/mL; ≥10 pg/mL). Data were censored from the

time of discontinuation for participants who discontinued before week 48 but not for

participants who initiated immunomodulatory therapy. HR, hazard ratio; IL-6, interleukin-6;

ITT, intention-to-treat; PBO, placebo; TCZ, tocilizumab.

1 *Table 1:* **Baseline demographics and disease characteristics (intention-to-treat**

2 **population**)

Characteristic	Placebo SC QW N=106	Tocilizumab 162 mg SC QW N=104
Female, n (%)	90 (84·9)	81 (77·9)
Age, years, mean (SD)	49·3 (12·6)	47·0 (12·2)
Former or current smoker, n (%)	40 (37·7)	32 (30·8)
Race, n (%)		
American Indian or Alaskan native	3 (2·8)	1 (1·0)
Asian	9 (8·5)	16 (15·4)
Black or African American	3 (2·8)	2 (1·9)
White	90 (84·9)	85 (81·7)
Other	1 (0·9)	0
Duration of SSc, months		
Mean (SD)	23·1 (17·0)	22·2 (16·0)
Median (IQR)	17·9 (9·4 to 33·2)	17·2 (9·0 to 34·9)
mRSS		
Mean (SD)	20·4 (7·0)	20·3 (6·7)
Median (IQR)	19·0 (15·0 to 26·0)	19·0 (15·0 to 24·5)
ppFVC[a]		
Mean (SD)	83·9 (15·0)	80·3 (14·4)
Median (IQR)	85·9 (72·4 to 95·9)	80·0 (69·3 to 90·2)
ppDLCO,[a] Hb corrected		

Mean (SD)	76·8 (18·6)	74·4 (19·2)
Median (IQR)	75·6 (65·7 to 85·8)	71·5 (59·1 to 89·3)
	n=105	n=104
Baseline SSc-ILD, n (%)[b]	68 (65)	68 (67)
	n=104	n=102
Baseline QLF-LM		
Mean [95% CI][c]	4·2 [2·4 to 6·0]	5·4 [3·0 to 7·8]
Median (IQR)	2·1 (1·0 to 4·4)	1·8 (0·7 to 4·9)
	n=84	n=73
Baseline QLF-WL		
Mean [95% CI][c]	1·8 [1·2 to 2·4]	2·7 [1·8 to 3·5]
Median (IQR)	1·1 (0.5 to 2·1)	1·2 (0·5 to 3·0)
	n=102	n=100
Baseline QILD-WL		
Mean [95% CI][c]	14·1 [12·0 to 16·1]	16·9 [14·1 to 19·6]
Median (IQR)	12·3 (7·5 to 20·2)	14·2 (7·0 to 24·4)
	n=102	n=100
HAQ-DI		
Mean (SD)	1·3 (0·7)	1·1 (0·8)
Median (IQR)	1·3 (0·9 to 1·8)	1·1 (0·4 to 1·8)
	n=104	n=104
IL-6 at screening, pg/mL		

<10, n (%)	77 (72·6)	77 (74·0)
≥10, n (%)	29 (27·4)	27 (26·0)
CRP, mg/mL		
Mean (SD)	7·0 (11·1)	8·9 (14·8)
Median (IQR)	3·8 (1·1 to 8·7)	4·0 (1·3 to 9·1)
ESR, mm/h		
Mean (SD)	34·7 (18·5)	34·8 (16·3)
Median (IQR)	33·0 (23·0 to 43·0)	33·5 (26·0 to 42·0)
	n=103	n=100
Platelet count, $\times 10^9$/L		
Mean (SD)	298·7 (96·0)	311·1 (88·2)
Median (IQR)	286·5 (231·0 to 358·0)	306·0 (243·0 to 361·0)
Antinuclear antibody positive, n (%)	90 (91·8)	91 (92·9)
	n=98	n=98
Anti–centromere antibody positive, n (%)	9 (9·0)	8 (8·0)
	n=100	n=100
Anti–RNA polymerase III antibody positive, n (%)	16 (16·0)	19 (19·0)
	n=100	n=100
Anti–topoisomerase I antibody positive, n (%)	49 (49·0)	52 (52·0)
	n=100	n=100

1   CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; HAQ-DI; Health Assessment

2   Questionnaire–Disability Index; HRCT, high-resolution computed tomography; ILD,

3   interstitial lung disease; IL-6, interleukin-6; IQR, interquartile range; mRSS, modified

Rodnan skin score; ppDLCO, percent predicted diffusing capacity for carbon monoxide; ppFVC, percent predicted forced vital capacity; QILD-WL, quantitative interstitial lung disease–whole lung; QLF-LM, quantitative lung fibrosis–most affected lobe; QLF-WL, quantitative lung fibrosis–whole lung; QW, every week; SC, subcutaneously; VAS, visual analog scale.

[a]FVC was measured using a centralized spirometry system provided to all sites and DLCO was measured using each site's equipment.

[b]Subset of participants who had ILD on visual read of HRCT by a thoracic radiologist.

[c]Participants who had available baseline and week 48 data.

The number of participants with evaluable data for each characteristic is 106 for placebo and 104 for tocilizumab unless shown otherwise.

1   *Table 2:* **Efficacy endpoints (intention-to-treat population)**

	Placebo SC QW N=106	**Tocilizumab 162 mg SC QW N=104**	**Difference between treatment groups[a]**
mRSS			
Primary endpoint, LSM change in mRSS from baseline to week 48 [95% CI]	–4·4 [–6·0 to -2·9]	–6·1 [–7·7 to -4·6]	–1·7 [–3·8 to 0·3] p=0·10[b]
LSM change in mRSS from baseline to week 24, LSM [95% CI]	–3·1 [–4·3 to –1.8]	–3·7 [–5·0 to –2·4]	–0·6 [–2·3 to 1·0] Nominal p=0.455
Improvement in mRSS from baseline ≥20%, n (%) [95% CI]	53 (50·0) [40·0 to 60·0]	75 (72·1) [63·0 to 81·2]	21·9 [9·2 to 34·6] Nominal p=0·0007[c]
Improvement in mRSS from baseline ≥40%, n (%) [95% CI]	40 (37·7) [28·0 to 47·4]	44 (42·3) [32·3 to 52·3]	4·3 [–8·7 to 17·3] Nominal p=0·51[c]
Improvement in mRSS from baseline ≥60%, n (%) [95% CI]	24 (22·6) [14·2 to 31·1]	18 (17·3) [9·6 to 25·1]	–5·4 [–16·2 to 5·4] Nominal p=0·33[c]
Treatment failure			

Treatment failure, n (%) Median TTF, weeks [95% CI]	37 (34·9) NE [48.7, NE]	23 (22·1) NE [NE]	HR [95% CI], 0·6 [0·4 to 1·1] Nominal p=0·08[d]
Components of treatment failure			
ppFVC >10% decrease, n (%) Median TTF, weeks [95% CI]	25 (23·6) NE [NE]	13 (12·5) NE [NE]	HR, 0·55 [0·3 to 1·1] Nominal p=0·08[d]
mRSS increase >20% and ≥5 points, n (%) Median TTF, weeks [95% CI]	16 (15·1) NE [NE]	10 (9·6) NE [NE]	HR, 0·64 [0·3 to 1·4] Nominal p=0·26[d]
SSc-related complication, n (%) Median TTF, weeks [95% CI]	7 (6·6) NE [NE]	5 (4·8) NE [NE]	HR, 0·79 [0·3 to 2·5] Nominal p=0·68[d]
Death, n (%) Median TTF, weeks [95% CI]	3 (2·8) NE [NE]	1 (1·0) NE [NE]	HR, 0·37 [0·0 to 3·6] Nominal p=0·39[d]
Treatment failure excluding decline in ppFVC, n (%) Median TTF, weeks [95% CI][e]	20 (18·9) NE [NE]	13 (12·5) NE [NE]	HR, 0·67 [0·3 to 1·4] Nominal p=0·26[d]

Treatment failure excluding increase in mRSS, n (%)	29 (27·4)	17 (16·3)	HR, 0·62 [0·3 to 1·1] Nominal p=0·12[d]
Median TTF, weeks [95% CI][e]	NE [NE]	NE [NE]	
Patient- and physician-reported outcomes, LSM change from baseline to week 48			
HAQ-DI [95% CI]	–0·06 [–0·16 to 0·05] n=102	–0·11 [–0·22 to –0·01] n=103	–0·05 [–0·19 to 0·09] Nominal p=0·45[b]
Patient global assessment VAS [95% CI]	–7·7 [–12·3 to –3·0] n=102	–10·1 [–14·8 to –5·4] n=102	–2·4 [–8·6 to 3·70] Nominal p=0·43[b]
Physician global assessment VAS [95% CI]	–20·0 [–24·8 to –15·22] n=96	–22·5 [–27·3 to –17·6] n=98	–2·5 [–8·7 to 3·8] Nominal p=0·44[b]
FACIT-Fatigue [95% CI]	2·6 n=102	5·1 n=103	2·40 [0·08 to 4·73] Nominal p=0·04[b]
SHAQ VAS [95% CI]	–0·3 [–0·5 to –0·1]	–0·3 [–0·5 to –0·2]	NA
SGRQ [95% CI]	–2·1 [–6·0 to 1·7]	–3·2 [–5·9 to –0·4]	NA

ACR-CRISS			
Median (IQR)	0·3 (0·0 to 1·0) n=82	0·9 (0·1 to 1·0) n=84	Nominal p=0·02[f]
Predicted probability of improvement from baseline ≥0.6, n (%) [95% CI][g]	39 (36·8) [27·1 to 46·4]	53 (51·0) [40·9 to 61·1]	13·9% [1·0 to 26·8] Nominal p=0.04[h]

1   ACR-CRISS, American College of Rheumatology–Combined Response Index in Systemic

2   Sclerosis; HR, hazard ratio; mRSS, modified Rodnan skin score; NA, not assessed; NE, not

3   estimable; QW, every week; SC, subcutaneously; SHAQ, Scleroderma Health Assessment

4   Questionnaire; SRGQ, St George's Respiratory Questionnaire; TTF, time to treatment failure;

5   VAS, visual analogue scale.

6   [a]All p values are nominal because the primary endpoint analysis was not significant.

7   [b]Based on difference in means using mixed-model repeated measures analysis including the

8   fixed categorical effects of treatment, visit, IL-6 at screening stratification, IL-6 at screening-

9   by-visit interaction, and treatment-by-visit interaction and the continuous covariates of

10   baseline score and baseline score-by-visit interaction.

11   [c]Wald with continuity correction for 95% CI. Weighted difference in proportions with 95%

12   CI using the Cochran-Mantel-Haenszel test adjusted for the stratification factor (IL-6 <10 or

13   ≥10 pg/mL at screening). Participants with missing week 48 assessment were considered

14   nonresponders for p value.

15   [d]Cox proportional hazards model adjusted for the stratification factor (IL-6 <10 or ≥10

16   pg/mL) at screening. Comparison of the two treatment groups with a non-parametric test of

17   survival time gave consistent results (Wilcoxon p=0·045).

1   [e]Post hoc analysis.

2   [f]Van Elteren test stratified by IL-6 level at screening (<10 or ≥10 pg/mL).

3   [g]Wald with continuity correction.

4   [h]Cochran-Mantel-Haenszel test stratified by IL-6 level at screening (<10 or ≥10 pg/mL).

5   All endpoints are shown at week 48 unless stated otherwise.

1 *Table 3:* **Lung function efficacy endpoints**

	Intention-to-treat population			Participants with SSc-ILD[a]		
	Placebo SC QW N=106	**Tocilizumab 162 mg SC QW N=104**	**Difference between treatment groups[b]**	**Placebo SC QW N=68**	**Tocilizumab 162 mg SC QW N=68**	**Difference between treatment groups[b]**
ppFVC change from baseline, median [95% CI]	n=91 –3·9 [–4·8 to –1·6]	n=93 –0·6 [–2·4 to 0·9]	3·3 [0·9 to 4·8] Nominal p=0·002	n=56 –4·0 [–5·3 to –1·7]	n=59 –0·6 [–3·2 to 2·0]	3·4 [0·4 to 5·6] Nominal p=0·002
ppFVC change from baseline, LSM [95% CI]	n=104 –4·6	n=104 –0·4	4·2 [2·0 to 6·4] Nominal p=0·0002[c]	n=66 –6·4	n=68 –0·1	6·5 [3·4 to 9·5] Nominal p<0·0001[c]

ppFVC ≥10% decline, n/N (%)	15/91 (16·5)	5/93 (5·4)	NA[d]	14/56 (25·0)	5/59 (8·5)	NA[d]
Improvement in ppFVC (increase ≥0%), n/N (%)	26/91 (28·6)	43/93 (46·2)	NA[d]	13/56 (23·2)	27/59 (45·8)	NA[d]
Absolute change from baseline in FVC, mL, LSM [95% CI]						
Week 24	n=104 −101	n=104 −13	88 [24 to 152] Nominal p=0.008[c]	n=66 −133	n=68 −15	118 [31 to 205] Nominal p=0.008[c]
Week 48	n=104 −190	n=104 −24	167 [83 to 250] Nominal p=0.0001[c]	n=66 −255	n=68 −14	241 [124 to 358] Nominal p<0.0001[c]
Observed ppDLCO	−2·1 [−4·4 to −0·4]	−2·4 [−4·1 to 1·0]	NA[d]	NA	NA	NA

Change from baseline, median [95% CI] Participants with ≥15% decline in ppDLCO, n/N (%)	8/82 (9·8)	7/79 (8·9)		NA	NA	NA
Change from baseline in observed HRCT QLF-LM[e]	n=66	n=60		n=36	n=35	
Median	0·3	0·0	–0·3	1·4	–0·2	–1·6
[95% CI]	[0·0 to 0·8]	[–0·3 to 0·2]	[–0·6 to 0·0] Nominal p=0·02[f]	[0·3 to 2·1]	[–2·2 to 0·2]	[–3·3 to –0·4] Nominal p=0·002[f]
Mean [95% CI]	0.9 [0.1 to 1.7]	-1.4 [-2.8 to 0.0]		1·9 [0·6 to 3·2]	-2.2 [-4·5 to 0·2]	

43

	n=81	n=84		n=48	n=54	
Change from baseline in observed HRCT QLF-WL[e,g]	n=81	n=84		n=48	n=54	
Median [95% CI]	0·1 [0·0 to 0·3]	0·0 [–0·2 to 0·1]	–0·1 (95% CI: –0·3 to –0·05) Nominal p=0·005[f]	0·4 [0·1 to 0·9]	–0·2 [–0·8 to 0·0]	–0·6 (–1·2 to –0·3) Nominal p=0·0008[f]
Mean [95% CI]	0·4 [0·0 to 0·7]	–0·4 [–0·9 to 0·1]		0·7 [0·3 to 1·2]	–0·6 [–1·4 to 0·2]	
Change from baseline in observed HRCT QILD-WL[e,g]	n=80	n=84		n=47	n=54	
Median [95% CI]	0·4 [–1·0 to 2·0]	–0·9 [–2·0 to –0·2]	–1·3 [–2·8 to –0·3]	1·6 [–1·1 to 2·8]	–1·7 [–2·3 to –0·7]	–3·3 [–4·3 to –0·7]

Mean [95% CI]	0·1 [−1·4 to 1·6]	−1·7 [−3·0 to −0·4]	Nominal p=0·04[f]	1·5 [−0·3 to 3·4]	−2·1 [−4·0 to −0·2]	Nominal p=0·008[f]

HRCT, high-resolution computed tomography; ILD, interstitial lung disease; NA, not assessed; ppDLCO, percent predicted diffusing capacity

for carbon monoxide; ppFVC, percent predicted forced vital capacity; QILD-WL, quantitative interstitial lung disease–whole lung; QLF-LM,

quantitative lung fibrosis–most affected lobe; QLF-WL, quantitative lung fibrosis–whole lung; QW, every week; SC, subcutaneously.

[a]Subset of participants who had ILD on visual read of HRCT by a thoracic radiologist.

[b]All p values are nominal because the result of the primary endpoint analysis was not significant.

[c]Based on difference in means using mixed-model repeated measures analysis including the fixed categorical effects of treatment, visit, IL-6 at

screening stratification, IL-6 at screening-by-visit interaction, and treatment-by-visit interaction and the continuous covariates of baseline score

and baseline score-by-visit interaction. Exploratory analysis of FVC with a mixed-model repeated measures analysis including a treatment

baseline IL-6 interaction gave a similar treatment effect in both IL-6 stratification subgroups: IL-6 <10 pg/mL, treatment difference 160 mL

(95% CI 60 to 260); IL-6 ≥10 pg/mL, treatment difference 180 mL (95% CI 10 to 340). Therefore, an interaction term with treatment was not

fitted for the mixed-model repeated measures analysis of FVC.

[d]Exploratory endpoint; no statistical comparison.

[e]Negative change indicates improvement.

1   [f]Based on van Elteren test of the medians adjusted for the stratification factor (IL-6 <10 or ≥10 pg/mL) at screening.

2   [g]Post hoc analysis.

3   FVC was measured using a centralised spirometry system provided to all sites, and DLCO was measured using each site's equipment. All

4   endpoints are shown at week 48 unless stated otherwise.

1 *Table 4:* **Safety (safety population)**

	Placebo SC QW N=106	**Tocilizumab 162 mg SC QW N=104**
Participants with ≥1 AE	82 (77·4)	89 (85·6)
Participants with ≥1 infectious AE	53 (50·0)	54 (51·9)
Participants with injection site reactions	3 (2·8)	8 (7·7)
Participants with ≥1 SAE	18 (17·0)	13 (12·5)
Participants with ≥1 infectious SAE	7 (6·6)	2 (1·9)
Participants with ≥1 noninfectious SAE	11 (10·4)	11 (10·6)
Withdrawal because of an AE	4 (3·8)	3 (2·9)
Deaths	3 (2·8)	1 (1·0)
Most frequent (≥5% of participants in either treatment arm) SAEs by SOC		
Infections and infestations, no. of events	8	3
Pneumonia	3 (2·8)	0
Infected skin ulcer	1 (0·9)	0
Osteomyelitis	0	1 (1·0)
Pelvic inflammatory disease	0	1 (1·0)
Chronic pyelonephritis	1 (0·9)	0
Respiratory tract infection	1 (0·9)	0
Sepsis	1 (0·9)	0
Soft tissue infection	1 (0·9)	0
Wound infection	0	1 (1·0)
Cardiac disorders, no. of events	7	2

Acute myocardial infarction	1 (0·9)	0
Angina pectoris	0	1 (1·0)
Atrial fibrillation	1 (0·9)	0
Cardiac failure	0	1 (1·0)
Chronic cardiac failure	1 (0·9)	0
Microvascular coronary artery disease	1 (0·9)	0
Myocardial infarction	1 (0·9)	0
Myocarditis	1 (0·9)	0

AE, adverse event; QW, every week; SAE, serious adverse event; SC, subcutaneously; SOC, system organ class.

Data are number of participants with event (%) unless stated otherwise.

Figure 2
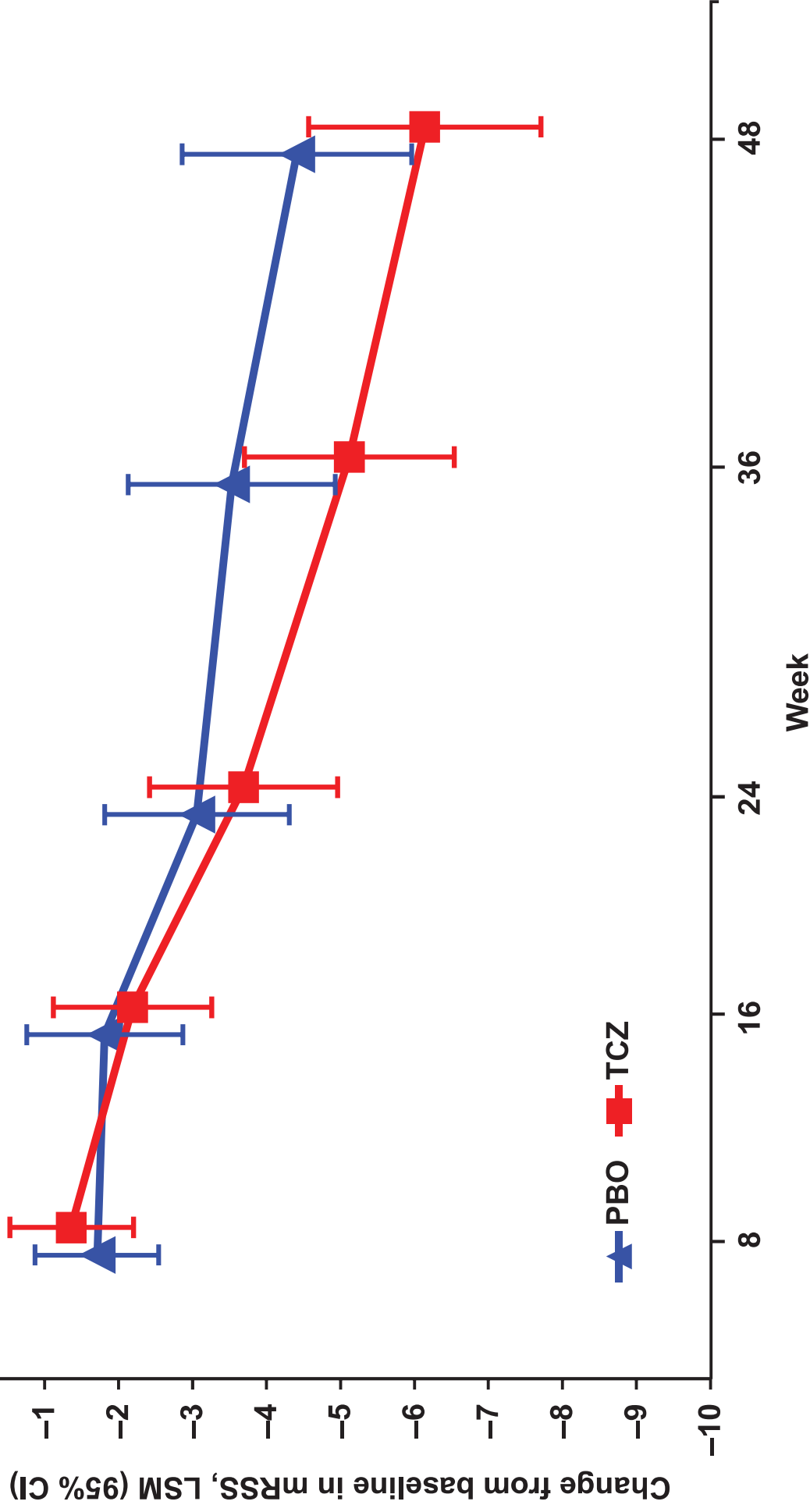
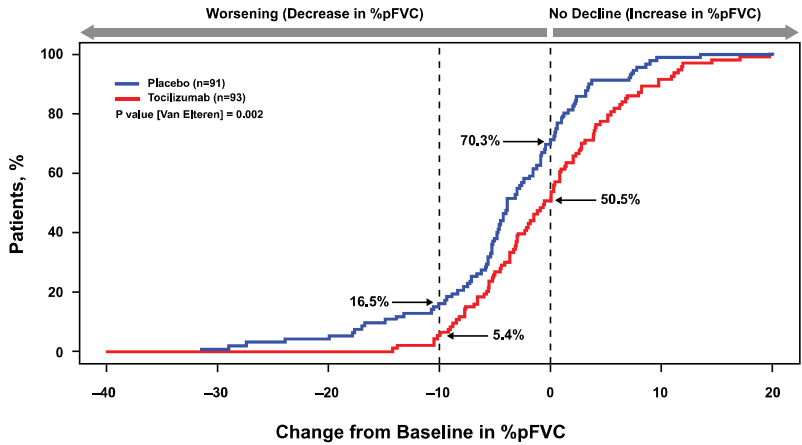|	PBO n=106	TCZ n=104	Difference (95% CI); p value
LSM change from BL at week 48 (primary outcome)	−4·4	−6·1	−1·7 (−3·8 to 0·3); p=0·10
LSM change from BL at week 24 (exploratory outcome)	−3·1	−3·7	−0·6 (−2·3 to 1·0); nominal p=0·45

Change from baseline in mRSS, LSM (95% CI)

Week

Figure 3A-D

A

Worsening (Decrease in %pFVC) ← → No Decline (Increase in %pFVC)

Placebo (n=91)
Tocilizumab (n=93)
P value [Van Elteren] = 0.002

70.3%
50.5%
16.5%
5.4%

Patients, %

Change from Baseline in %pFVC

B

Worsening (Decrease in %pFVC) ← → No Decline (Increase in %pFVC)

Placebo (n=56)
Tocilizumab (n=59)
P value [Van Elteren] = 0.0016

75.0%
50.8%
25.0%
8.5%

Patients, %

Change from Baseline in %pFVC

C

Change from Baseline in %pFVC, LSM (95% CI)

PBO (n=106)
TCZ (n=104)

Week

	PBO n=106	TCZ n=104	Difference (95% CI); P value
LSM change from baseline at week 48	−4.6	−0.4	4.2 (2.0-6.4); nominal P<0.001

D

Change from Baseline in %pFVC, mean (95% CI)

PBO (n=68)
TCZ (n=68)

Week

Figure 4

Figure 4