Safety and efficacy of subcutaneous tocilizumab in adults with systemic sclerosis (faSScinate): a phase 2, randomised, controlled trial

Dinesh Khanna, Christopher P Denton, Angelika Jahreis, Jacob M van Laar, Tracy M Frech, Marina E Anderson, Murray Baron, Lorinda Chung, Gerhard Fierlbeck, Santhanan Lakshminarayanan, Yannick Allanore, Janet E Pope, Gabriela Riemekasten, Virginia Steen, Ulf Müller-Ladner, Robert Lafyatis, Giuseppina Stefano, Helen Spotwood, Haiyun Chen-Harris, Sebastian Dziadek, Alyssa Morimoto, Thierry Sornasse, Jeffrey Siegel, Daniel E Furst

Summary

Background Systemic sclerosis is a rare disabling autoimmune disease with few treatment options. The efficacy and safety of tocilizumab, an interleukin 6 receptor-α inhibitor, was assessed in the faSScinate phase 2 trial in patients with systemic sclerosis.

Methods We did this double-blind, placebo-controlled study at 35 hospitals in Canada, France, Germany, the UK, and the USA. We enrolled adults with progressive systemic sclerosis of 5 or fewer years’ duration from first non-Raynaud’s sign or symptom. Patients were randomly assigned (1:1) to weekly subcutaneous tocilizumab 162 mg or placebo. The primary endpoint was the difference in mean change from baseline in modified Rodnan skin score at 24 weeks. This study is registered with ClinicalTrials.gov, number NCT01532869.

Findings We enrolled 87 patients: 43 assigned to tocilizumab and 44 assigned to placebo. The least squares mean change in modified Rodnan skin score at 24 weeks was –3.92 in the tocilizumab group and –1.22 in the placebo group (difference –2.70, 95% CI –5.85 to 0.45; p=0.0915). The least squares mean change at 48 weeks was –6.33 in the tocilizumab group and –2.77 in the placebo group (treatment difference –3.55, 95% CI –7.23 to 0.12; p=0.0579). In one of several exploratory analyses, fewer patients in the tocilizumab group than in the placebo group had a decline in percent predicted forced vital capacity at 48 weeks (p=0.0373). However, we detected no significant difference in disability, fatigue, itching, or patient or clinician global disease severity. 42 (98%) of 43 patients in the tocilizumab group versus 40 (91%) of 44 in the placebo group had adverse events. 14 (33%) versus 15 (34%) had serious adverse events. Serious infections were more common in the tocilizumab group (seven [16%] of 43 patients) than in the placebo group (two [5%] of 44). One patient died in relation to tocilizumab treatment.

Interpretation Tocilizumab was not associated with a significant reduction in skin thickening. However, the difference was greater in the tocilizumab group than in the placebo group and we found some evidence of less decline in forced vital capacity. The efficacy and safety of tocilizumab should be investigated in a phase 3 trial before definitive conclusions can be made about its risks and benefits.

Funding F Hoffmann-La Roche, Genentech.

Introduction Systemic sclerosis is a rare connective tissue disorder characterised by fibrosis, inflammation, and microvascular injury with heterogeneous clinical presentations. Pulmonary, cardiac, gastrointestinal, and renal complications contribute to patient morbidity and decreased survival.1 Increasing evidence supports important roles for interleukin 6 in the pathogenesis of systemic sclerosis,1 including B-cell differentiation towards immunoglobulin-secreting plasma cells, T-cell differentiation towards Th17 cell types, and transformation of fibroblasts to myofibroblasts leading to synthesis of extracellular matrix.2 Dermal fibroblasts from patients with systemic sclerosis constitutively express more interleukin 6 than those from healthy controls,3 and serum and skin concentrations of interleukin 6 are elevated in patients with early systemic sclerosis.4 In patients with systemic sclerosis with or without interstitial lung disease, increased interleukin 6 has been associated with higher mortality, more severe skin involvement, and increased incidence of progressive pulmonary decline.5 Although the exact cellular mechanisms of the effects of interleukin 6 on fibrosis are unknown, myeloid cells are implicated in systemic sclerosis skin pathogenesis.6 mRNA expression of a cluster of macrophage genes, including CD14 in the skin, correlates strongly with modified Rodnan skin score, and CD14 expression is prognostic for progressive skin disease.7 M2-macrophages appear to have an important role in mediating inflammation and promoting fibrosis through the release of profibrotic factors.8,9 In the bleomycin mouse model, blockade of the interleukin 6 pathway reduced skin fibrosis, a smooth-muscle actin protein expression,10 hydroxyproline concentrations of interleukin 6 are elevated in patients...
High serum concentrations of CCL18 are associated with systemic sclerosis pathobiology. Serum concentrations of interleukin 6 is increased in patients with early systemic sclerosis can predict the extent of progression of skin disease. Furthermore, the extent of fibrosis was attenuated in interleukin 6 knockout mice. Two treatment-refractory patients with systemic sclerosis had improved skin thickening after treatment with tocilizumab for 6 months.

Methods
Study design and participants
We did this randomised, double-blind, placebo-controlled study in 35 hospitals across Canada, France, Germany, the UK, and the USA. Investigators from each centre enrolled patients aged 18 years or older who met the 1980 American College of Rheumatology criteria for systemic sclerosis with or without face involvement. Participants had to have had an increase of at least 3 on the modified Rodnan skin score at screening compared with the last visit within the previous 1–6 months or new-onset systemic sclerosis diagnosed within 1 year before screening. Involvement of one new body area with an increase of modified Rodnan skin score of at least 2 or two new body areas with increase of at least 1, documentation of worsening of skin thickening in the previous 6 months, or at least one tendon friction rub plus at least one laboratory criterion (C-reactive protein ≥10·0 mg/L, erythrocyte sedimentation rate ≥28 mm/h, or platelets ≥330 000/μL). All patients provided written informed consent. Patients or caregivers could administer subcutaneous injections of the investigational drug. Eligible patients had clinically uninvolved skin in at least one body area for study drug injections.

Randomisation and masking
Patients were randomly assigned (1:1) using an interactive voice and web response system to receive weekly subcutaneous treatment with tocilizumab 162 mg or placebo for 48 weeks followed by open-label weekly tocilizumab for 48 weeks. Randomisation numbers were generated by the sponsor, and randomisation was stratified by joint involvement at baseline (<four vs ≥four joints on the 28 tender joint count). Investigators, patients, and sponsor personnel were masked to treatment assignment. To prevent unmasking, separate assessors evaluated efficacy and safety. The efficacy assessor did not have access to safety data during the
double-blind phase of the trial, but the safety assessor had access to both efficacy and safety data. Although some sponsor personnel were unmasked after the primary analysis at 24 weeks, personnel interacting with sites and site staff remained masked to treatment assignment until the database lock at 48 weeks.

Procedures
Escape treatment with methotrexate, hydroxychloroquine, or mycophenolate mofetil was permitted after 24 weeks for patients with 20% or more worsening of modified Rodman skin score from baseline, worsening complications associated with systemic sclerosis such as arthritis, and idiopathic lung disease, as determined by the treating investigator. Safety monitoring for adverse events and serious adverse events and laboratory monitoring were done at least every 8 weeks.

Outcomes
The primary endpoint was the difference in mean change in modified Rodnan skin score from baseline to 24 weeks. Secondary endpoints were patient-reported and physician-reported outcomes to 24 weeks and 48 weeks (Health Assessment Questionnaire–Disability Index [HAQ–DI] score, patient global visual analogue scale [VAS, 0–100], physician global VAS, Functional Assessment of Chronic Illness Therapy [FACIT]-Fatigue score, and pruritus 5-D Itch score), change in modified Rodnan skin score from baseline to 48 weeks, proportion of patients with change in modified Rodnan skin score from baseline at 48 weeks greater than or equal to change from baseline to week 24, and change from baseline in VAS score at 24 weeks and 48 weeks (intestinal, breathing, Raynaud’s disease, finger ulcers, overall disease from the EQ-5D, change from baseline at 24 and 48 weeks in 28 tender joint count in patients with joint involvement (at least four tender joints at baseline), change from baseline at 24 and 48 weeks in EQ-5D-3L, change from baseline at 24 and 48 weeks in digital ulcer count, and change from baseline at 24 and 48 weeks in tender friction rub. Proportions of patients with 20%, 40%, and 60% improvement in modified Rodman skin score at 24 and 48 weeks and correlation analysis between improvement in modified Rodnan skin score and change in percent predicted FVC at 48 weeks were post hoc outcomes. Exploratory biomarker analyses included gene expression analysis of skin biopsy specimens collected at baseline and at 24 weeks and COMP, POSTN, autoptxin/EENP2, and CCL18 serum concentrations measured with immunoassays (appendix pp 4, 5). Additionally, protein expression in the skin was analysed with immunohistochemistry, and leucocyte immunophenotypes were analysed with flow cytometry. Pharmacokinetic and pharmacodynamic (tocilizumab, sIL-6R and IL-6) analyses were conducted.

Statistical analysis
A sample size of 36 patients per group (86 patients total, allowing for 15% dropout) was determined to provide 80% power to detect a difference in means for the change in modified Rodnan skin score of 4·7 from baseline to 24 weeks, based on an estimated common SD of 6·99 using a two-group t test with a 5% two-sided significance level.

The primary endpoint was analysed with a mixed-model repeated-measures approach. There was no imputation of missing data before these analyses, and all patients were stratified by joint involvement at baseline. Patients discontinuing treatment before the end of the double-blind period underwent their last evaluation at the time of discontinuation. Except for exploratory analysis of pulmonary function, which included all available data, all mixed-model repeated-measures analyses of secondary endpoints at week 48 used data that were censored after initiation of escape treatment. The mixed-model repeated-measures approach assumes that data are missing at random; therefore, sensitivity analyses were done for the primary endpoint to account for data that might not have been missing because of chance (appendix p 11). There was no preplanned statistical analysis of exploratory endpoints; when clinically important exploratory endpoints showed a treatment effect, further unplanned analyses were conducted.

Patients with missing modified Rodnan skin score data at 24 weeks or 48 weeks were considered non-responders for analysis of 20%, 40%, and 60% improvement from baseline. Similarly, patients without data were considered non-responders for the minimal clinically important difference for HAQ–DI, modified Rodnan skin score, and maintenance of modified Rodnan skin score response. The van Elteren test (stratified by joint involvement) was used to compare the treatment effect on the cumulative distribution of change from baseline in percent predicted FVC and percent predicted diffusing capacity for CO. No adjustment was made for multiplicity in statistical testing for any of the analyses. For serious adverse event rates rates per 100 patient-years, multiple occurrences of events in a patient were counted, and confidence intervals were based on the Poisson distribution. The efficacy analyses were done for the modified intention-to-treat population, which included all randomly assigned patients who
The primary endpoint was not met. Least squares mean change in modified Rodnan skin score from baseline to 24 weeks was −3.92 in the tocilizumab group and −1.22 in the placebo group (treatment difference −2.70, 95% CI −5.85 to 0.45; p=0.0915; table 2, figure 2). The observed mean modified Rodnan skin score at 24 weeks was 21.84 (SD 9.89) for tocilizumab (n=37) and 23.21 (SD 9.26) for placebo (n=38).

The least squares mean change from baseline to 48 weeks was −6.33 in the tocilizumab group and −2.77 in the placebo group (treatment difference −3.55, 95% CI −7.23 to 0.12; p=0.0977; table 2, figure 2). The observed mean modified Rodnan skin score at 48 weeks was 19.56 (10.08) in the tocilizumab group (n=32) and 22.27 (8.0) in the placebo group.

For the safety population. Data are mean (SD) or n (%) unless stated otherwise. CRP=C-reactive protein. DLCO (Hb corr)=diffusing capacity for CO corrected for haemoglobin. FVC=forced vital capacity. HAQ-DI=Health Assessment Questionnaire—Disability Index. mRSS=modified Rodnan skin score. TJC28=tender joint count using 28 joints. VAS=visual analogue scale. *Calculated from time of first non-Raynaud’s sign or symptom. †Previous or concomitant. ‡Treatment that ended before first study treatment administration excluded from the summary. Medication class “steroids” excluding the following routes of administration: nasal, nasogastric, topical, and respiratory (inhalation).

Table 1: Baseline characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Placebo group (n=44)</th>
<th>Tocilizumab group (n=43)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>48 (12.9)</td>
<td>51 (11.7)</td>
</tr>
<tr>
<td>Women</td>
<td>35 (80%)</td>
<td>32 (74%)</td>
</tr>
<tr>
<td>White</td>
<td>40 (91%)</td>
<td>38 (88%)</td>
</tr>
<tr>
<td>Duration of systemic sclerosis (months)*</td>
<td>19.5 (17.0)</td>
<td>17.6 (13.9)</td>
</tr>
<tr>
<td>Anti-RNA polymerase antibody positive</td>
<td>17 (39%)</td>
<td>13 (30%)</td>
</tr>
<tr>
<td>Previous biological drugs</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Immunosuppressive drugs†</td>
<td>19 (43%)</td>
<td>22 (51%)</td>
</tr>
<tr>
<td>Concomitant systemic corticosteroid use†</td>
<td>17 (39%)</td>
<td>21 (49%)</td>
</tr>
<tr>
<td>Anti-topoisomerase antibody positive</td>
<td>20 (45%)</td>
<td>18 (42%)</td>
</tr>
<tr>
<td>Patients with tendon friction n&gt;1</td>
<td>22 (50%)</td>
<td>20 (47%)</td>
</tr>
<tr>
<td>TJC28 ≥4</td>
<td>21 (49%)§</td>
<td>20 (47%)§</td>
</tr>
<tr>
<td>TJC28</td>
<td>7 (8.5)§</td>
<td>7 (8.9)§</td>
</tr>
<tr>
<td>Total mRSS</td>
<td>26 (5.9)</td>
<td>26 (7.2)</td>
</tr>
<tr>
<td>Overall HAQ-DI score</td>
<td>1 (0.7)</td>
<td>1 (0.6)§</td>
</tr>
<tr>
<td>Clinician global VAS</td>
<td>61 (15.2)</td>
<td>64 (15.3)</td>
</tr>
<tr>
<td>CRP (mg/L)</td>
<td>10 (13.5)</td>
<td>10 (13.5)</td>
</tr>
<tr>
<td>Platelet count (10³ per L)</td>
<td>308 (88.9)</td>
<td>308 (82.4)</td>
</tr>
<tr>
<td>Percent predicted FVC</td>
<td>82% (13%)</td>
<td>80% (14%)</td>
</tr>
<tr>
<td>Percent predicted DLCO</td>
<td>74% (21%)</td>
<td>73% (19%)§</td>
</tr>
<tr>
<td>Patients with ≥1 digital ulcer†</td>
<td>10 (23%)</td>
<td>11 (26%)§</td>
</tr>
</tbody>
</table>

Table 2: Treatment of patients who developed skin ulcer

For the safety population. Data are mean (SD) or n (%) unless stated otherwise. CRP=C-reactive protein. DLCO (Hb corr)=diffusing capacity for CO corrected for haemoglobin. FVC=forced vital capacity. HAQ-DI=Health Assessment Questionnaire—Disability Index. mRSS=modified Rodnan skin score. TJC28=tender joint count using 28 joints. VAS=visual analogue scale. *Calculated from time of first non-Raynaud’s sign or symptom. †Previous or concomitant. ‡Treatment that ended before first study treatment administration excluded from the summary. Medication class “steroids” excluding the following routes of administration: nasal, nasogastric, topical, and respiratory (inhalation).

Table 3: Treatment of patients who developed skin ulcer

For the safety population. Data are mean (SD) or n (%) unless stated otherwise. CRP=C-reactive protein. DLCO (Hb corr)=diffusing capacity for CO corrected for haemoglobin. FVC=forced vital capacity. HAQ-DI=Health Assessment Questionnaire—Disability Index. mRSS=modified Rodnan skin score. TJC28=tender joint count using 28 joints. VAS=visual analogue scale. *Calculated from time of first non-Raynaud’s sign or symptom. †Previous or concomitant. ‡Treatment that ended before first study treatment administration excluded from the summary. Medication class “steroids” excluding the following routes of administration: nasal, nasogastric, topical, and respiratory (inhalation).

Table 4: Treatment of patients who developed skin ulcer

For the safety population. Data are mean (SD) or n (%) unless stated otherwise. CRP=C-reactive protein. DLCO (Hb corr)=diffusing capacity for CO corrected for haemoglobin. FVC=forced vital capacity. HAQ-DI=Health Assessment Questionnaire—Disability Index. mRSS=modified Rodnan skin score. TJC28=tender joint count using 28 joints. VAS=visual analogue scale. *Calculated from time of first non-Raynaud’s sign or symptom. †Previous or concomitant. ‡Treatment that ended before first study treatment administration excluded from the summary. Medication class “steroids” excluding the following routes of administration: nasal, nasogastric, topical, and respiratory (inhalation).

Table 5: Treatment of patients who developed skin ulcer

For the safety population. Data are mean (SD) or n (%) unless stated otherwise. CRP=C-reactive protein. DLCO (Hb corr)=diffusing capacity for CO corrected for haemoglobin. FVC=forced vital capacity. HAQ-DI=Health Assessment Questionnaire—Disability Index. mRSS=modified Rodnan skin score. TJC28=tender joint count using 28 joints. VAS=visual analogue scale. *Calculated from time of first non-Raynaud’s sign or symptom. †Previous or concomitant. ‡Treatment that ended before first study treatment administration excluded from the summary. Medication class “steroids” excluding the following routes of administration: nasal, nasogastric, topical, and respiratory (inhalation).
(SD 8·05) in the placebo group (n=33). Of patients whose modified Rodnan skin score had improved by 24 weeks, more patients in the tocilizumab group (15 [68%] of 22) than in the placebo group (eight [44%] of 18) maintained or had further improvement in modified Rodnan skin score by 48 weeks. At 24 weeks, more patients in the tocilizumab group had modified Rodnan skin score improvement of at least 20% and of at least 4·7 points, but the same number of patients in each group had at least 40% or 60% improvement. At 48 weeks, more patients in the tocilizumab group than in the placebo group (15 [68%] of 22) had modified Rodnan skin score improvement of at least 20%, 40%, 60%, and of at least 4·7 points (table 3). Eight patients in the placebo group and four in the tocilizumab group had an increase in modified Rodnan skin score of more than 5 and at least 25% from baseline at any time to 24 weeks.

For clinician and patient global VAS and FACIT-fatigue scores, the difference was not significant at 24 weeks or 48 weeks, but scores were better in the tocilizumab group than in the placebo group (table 2). For the HAQ–DI, we found no significant difference between treatment groups (table 2). At 48 weeks, 12 (28%) of 43 patients in the tocilizumab group versus 3 (7%) of 44 in the placebo group achieved an improvement of 0·22 or more in HAQ–DI (p=0·0111), with identical results for improvement defined as 0·14 or more. At 24 weeks, 11 (26%) of 43 patients in the tocilizumab group versus 7 (16%) of 44 in the placebo group achieved an improvement of 0·22 or more in HAQ–DI, with identical results for improvement defined as 0·14 or more. Overall, with the exception of breathing VAS score, all Scleroderma HAQ–DI VAS scores at 48 weeks had no significant difference between groups, although the scores favoured tocilizumab (appendix p 12). At 24 weeks, little treatment benefit was observed.

20 (47%) of 43 patients in the tocilizumab group and 21 (49%) of 43 in the placebo group had joint involvement. Among these patients, mean tender joint counts changed from baseline by –4·3 (SD 7·3; median –2·5, IQR –10·0 to –1·0; n=36) at 4 weeks and –5·1 (SD 7·3; median –4·5, IQR –12·0 to 0·0; n=10) at 48 weeks with tocilizumab compared with –2·1 (SD 6·3; median –2·0, IQR –4·0 to –1·0; n=17) at 24 weeks and –2·9 (SD 7·1; median –2·5, IQR –5·5 to 2·5; n=12) at 48 weeks in the placebo group.

We showed a significantly smaller decrease in FVC for tocilizumab than for placebo from baseline to 24 weeks (tocilizumab –34 mL vs placebo –71 mL; least square mean difference 136 mL, 95% CI 9 to 264; p=0·0368) but from baseline to 48 weeks the difference was not significant (tocilizumab –117 mL vs placebo –237 mL; 120 mL, 95% CI –23 to 262; p=0·0990). Fewer patients in the tocilizumab group than in the placebo group had worsening of percent predicted FVC at 24 weeks (p=0·009; figure 3A) or at 48 weeks (p=0·037; figure 3B). At 24 weeks, one (3%) of 30 patients in the tocilizumab group versus seven (19%) of 36 in the placebo group had more than 10% (absolute) decreases in percent predicted FVC from baseline. At 48 weeks, the proportion was three (10%) of 30 versus seven (23%) of 31. At 24 weeks the least squares mean change from baseline in percent predicted FVC was –0·7 (95% CI –3·2 to 1·8) in the tocilizumab group versus –4·5 (–7·0 to –1·9) in the placebo group. At 48 weeks, the change was –2·6 (–5·2 to –0·1) versus –6·3 (–8·9 to –3·8).

For the modified intention-to-treat population. A negative change from baseline shows improvement for all efficacy measures, except for the FACIT-Fatigue Scale, where positive change indicates improvement. We did a mixed-model repeated-measures analysis that included the fixed categorical effects of treatment, visit, stratification factor of joint involvement at baseline, and treatment-by-visit interaction, and the continuous covariates of baseline score and baseline score-by-visit interaction. FACIT=Functional Assessment of Chronic Illness Therapy. HAQ–DI=Health Assessment Questionnaire—Disability Index. VAS=visual analogue scale.

For the modified intention-to-treat population. A negative change from baseline shows improvement for all efficacy measures, except for the FACIT-Fatigue Scale, where positive change indicates improvement. We did a mixed-model repeated-measures analysis that included the fixed categorical effects of treatment, visit, stratification factor of joint involvement at baseline, and treatment-by-visit interaction, and the continuous covariates of baseline score and baseline score-by-visit interaction. FACIT=Functional Assessment of Chronic Illness Therapy. HAQ–DI=Health Assessment Questionnaire—Disability Index. VAS=visual analogue scale.

Table 2: Least square mean change from baseline for mRSS and in patient-reported and physician-reported outcomes

<table>
<thead>
<tr>
<th>Table 2: Least square mean change from baseline for mRSS and in patient-reported and physician-reported outcomes</th>
<th>Placebo group (n=44)</th>
<th>Tocilizumab group (n=43)</th>
<th>Difference in means (95% CI)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>mRSS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 24</td>
<td>–1·22 (n=43)</td>
<td>–3·92 (n=43)</td>
<td>–2·70 (–5·85 to 0·45)</td>
<td>0·0915</td>
</tr>
<tr>
<td>Week 48</td>
<td>–2·77 (n=43)</td>
<td>–6·33 (n=43)</td>
<td>–3·55 (–7·23 to 0·12)</td>
<td>0·0579</td>
</tr>
<tr>
<td>HAQ-DI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 24</td>
<td>0·117 (n=42)</td>
<td>0·137 (n=41)</td>
<td>0·020 (–0·186 to 0·225)</td>
<td>0·8503</td>
</tr>
<tr>
<td>Week 48</td>
<td>0·205 (n=41)</td>
<td>0·002 (n=41)</td>
<td>0·207 (–0·471 to 0·056)</td>
<td>0·1212</td>
</tr>
<tr>
<td>Clinician Global VAS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 24</td>
<td>–7·25 (n=41)</td>
<td>–8·24 (n=39)</td>
<td>–0·99 (–9·20 to 7·23)</td>
<td>0·8118</td>
</tr>
<tr>
<td>Week 48</td>
<td>–9·39 (n=41)</td>
<td>–18·41 (n=40)</td>
<td>–9·02 (–19·04 to 1·00)</td>
<td>0·0768</td>
</tr>
<tr>
<td>Patient Global VAS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 24</td>
<td>1·53 (n=42)</td>
<td>–2·33 (n=42)</td>
<td>–3·85 (–13·04 to 5·34)</td>
<td>0·4063</td>
</tr>
<tr>
<td>Week 48</td>
<td>2·70 (n=41)</td>
<td>–11·00 (n=42)</td>
<td>–8·30 (–19·31 to 2·71)</td>
<td>0·1371</td>
</tr>
<tr>
<td>FACIT (Fatigue) Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 24</td>
<td>1·26 (n=41)</td>
<td>2·68 (n=42)</td>
<td>1·43 (–2·92 to 5·82)</td>
<td>0·5197</td>
</tr>
<tr>
<td>Week 48</td>
<td>0·36 (n=40)</td>
<td>3·11 (n=42)</td>
<td>2·75 (–13·8 to 6·88)</td>
<td>0·1886</td>
</tr>
<tr>
<td>S-D Itch Scale</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 24</td>
<td>–1·73 (n=41)</td>
<td>–0·94 (n=41)</td>
<td>0·79 (–0·94 to 2·51)</td>
<td>0·3651</td>
</tr>
<tr>
<td>Week 48</td>
<td>–1·08 (n=40)</td>
<td>–2·19 (n=41)</td>
<td>–1·11 (–3·16 to 0·94)</td>
<td>0·2841</td>
</tr>
</tbody>
</table>

![Figure 2: Change (95% CI) from baseline in modified Rodnan skin score](image-url)

![Figure 3: Change (95% CI) from baseline in modified Rodnan skin score](image-url)
A change of 4.7 or more is deemed clinically important. Patients can be counted more than once. For example, a patient with a 43% improvement would be counted as having had decrease of both more than 20% and more than 40%.

Table 3: Changes from baseline in modified Rodnan skin score

<table>
<thead>
<tr>
<th>Decrease ≥20%</th>
<th>Placebo group (n=44)</th>
<th>Tocilizumab group (n=43)</th>
<th>p value</th>
<th>Placebo group (n=44)</th>
<th>Tocilizumab group (n=43)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decrease ≥40%</td>
<td>6 (14%)</td>
<td>6 (14%)</td>
<td>1.00</td>
<td>3 (7%)</td>
<td>9 (21%)</td>
<td>0.069</td>
</tr>
<tr>
<td>Decrease ≥60%</td>
<td>2 (5%)</td>
<td>2 (5%)</td>
<td>1.00</td>
<td>0 (0%)</td>
<td>5 (12%)</td>
<td>0.026</td>
</tr>
<tr>
<td>Decrease ≥4.7 units</td>
<td>10 (23%)</td>
<td>16 (37%)</td>
<td>0.16</td>
<td>11 (25%)</td>
<td>16 (37%)</td>
<td>0.025</td>
</tr>
</tbody>
</table>

Figure 3: Cumulative distribution of patients by change in percent predicted FVC

A change is shown as absolute change in percent predicted FVC. From baseline to 24 weeks (A), and from baseline to 48 weeks (B). FVC=forced vital capacity.

We explored gene expression using microarray technology on skin biopsy specimens collected from participants at baseline (34 patients in the tocilizumab group vs 39 in the placebo group) and at 24 weeks (28 vs 30), and from 20 age-matched and sex-matched healthy controls from Epistem (Manchester, UK; median age 43 years [22–62] years, six male and 14 female). First, we conducted gene set enrichment analysis35 (using GSEA: Linear Model Toolset for Gene Set Enrichment Analysis; R, version 1.30.0) for fibrosis, interferon-α, interleukin 6, TGF-β, M1-macrophage, and M2-macrophage gene sets (appendix pp 4, 13–36) of the microarray data obtained from all available samples. We looked for gene set enrichment among the genes overexpressed in patients with systemic sclerosis versus healthy controls, and among the genes downregulated by tocilizumab at week 24 versus baseline. At baseline, gene sets associated with fibrosis, interferon-α, interleukin 6, TGF-β, M1-macrophage, and M2-macrophage were significantly enriched in the group of genes overexpressed in patients with systemic sclerosis compared with those of healthy controls (p<0.005; appendix p 37). None of these gene sets was significantly enriched in the genes downregulated by tocilizumab when comparing week 24 samples to baseline (appendix p 38). We recorded non-significant enrichment in interferon α (p=0.07), interleukin 6 (p=0.095), and M2-macrophage (p=0.105) gene sets in the genes downregulated by tocilizumab when comparing week 24 with baseline (appendix p 38); we recorded no enrichment in the placebo group when comparing week 24 with baseline (appendix p 39).

Genome-wide analysis of gene expression in the biopsy samples was used to select 83 genes by unsupervised clustering of genes overexpressed in systemic sclerosis and downregulated by tocilizumab, and genes with expression correlated with modified Rodnan skin score (appendix p 5).Confirmatory analysis was done on these 83 genes by use of nCounter technology (NanoString Technologies; appendix p 5). Of the 83 genes representing the fibrosis, TGF-β, interleukin 6, interferon, and myeloid pathways, 62 transcripts were significantly overexpressed and two were significantly underexpressed in patients with systemic sclerosis compared with healthy controls (t test, Bonferroni correction for multiple comparisons). Clustering analysis of the average gene expression for the placebo and tocilizumab groups at both baseline and 24 weeks yielded nine clusters representing TGF-β, interleukin 6/STAT3, M1-macrophage, M2-macrophage, compared with those of age-matched and sex-matched healthy controls from Bioreclamation IVT (New York, NY, USA; median age 46 years [30–59], seven male and 18 female; p<0.0001 for all biomarkers; two-tailed t test assuming unequal variance). Serum concentration of CCL18 was significantly lower in the tocilizumab group than in the placebo group but we detected no effect of tocilizumab on the serum concentrations of COMP, POSTN, or ENPP2 (figure 4).
and interferon (appendix pp 8, 9). Analysis of the effect of treatment on change in gene expression at 24 weeks identified 16 genes specifically downregulated by tocilizumab compared with placebo (uncorrected p<0·05; appendix pp 8, 9). Although most of these genes (12 [75%] of 16) belonged to the M2-macrophage cluster, two belonged to the M1-macrophage cluster, suggesting that tocilizumab might have inhibitory actions on macrophages in general, and M2-macrophages in particular. Gene expression in the interleukin 6/STAT3 cluster was numerically reduced by tocilizumab at 24 weeks compared with placebo, with only two genes (CCL2 and SOCS3) significantly downregulated by tocilizumab, possibly reflecting the residual activity of other STAT3-activating factors (eg, EGF, interleukin 5, interleukin 6, HGF, LIF, and BMP2, which are not blocked by tocilizumab). None of the genes in the TGF-β-1 or TGF-β-2 clusters was significantly downregulated by tocilizumab compared with placebo, probably because the expression of these genes was reduced at 24 weeks in the skin of patients in both treatment groups, suggesting a tocilizumab-independent reduction in the fibrotic pathway in both groups. Gene expression in the interferon cluster was highly heterogeneous, limiting our ability to identify any pattern of expression in this gene cluster.

Finally, using the same dataset, we tested the differential effect of treatment with tocilizumab using a multianalyte, longitudinal, pharmacodynamic biomarker (2GSSc skin biomarker27). This biomarker, which predicts modified Rodnan skin score on the basis of the weighted values for THBS1 and MS4A4A mRNA expression, was validated and has been used in two clinical trials, indicating that it is a robust surrogate outcome measure for the extent of systemic sclerosis skin disease.28 At baseline, the predicted modified Rodnan skin scores were similar between treatment groups (placebo mean 16·4, 95% CI 15·1 to 17·8; tocilizumab mean 15·5, 95% CI 14·1 to 17·0). Comparison of the change in predicted modified Rodnan skin score at 24 weeks indicated that treatment with tocilizumab resulted in a significant decrease compared with treatment with placebo (placebo mean change –0·98, 95% CI –2·99 to 1·03; tocilizumab mean change –4·03, 95% CI –7·58 to –0·49; p=0·0488; appendix p 10). We did not detect any numerically significant effect of tocilizumab on protein expression in the skin (aSMA, PTX3, pSTAT3, interleukin 6, and CCL2) or on leucocytes.
### Serious adverse events

<table>
<thead>
<tr>
<th>Event Category</th>
<th>Placebo group (n=44)</th>
<th>Tocilizumab group (n=43)</th>
<th>Placebo group (n=44)</th>
<th>Tocilizumab group (n=43)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients with ≥1 adverse event</td>
<td>11 (25%)</td>
<td>9 (21%)</td>
<td>15 (34%)</td>
<td>14 (33%)</td>
</tr>
<tr>
<td>Deaths</td>
<td>1 (2%)</td>
<td>1 (2%)</td>
<td>1 (2%)</td>
<td>1 (2%)</td>
</tr>
<tr>
<td>Withdrawals because of an adverse event</td>
<td>5 (11%)</td>
<td>4 (9%)</td>
<td>5 (11%)</td>
<td>6 (14%)</td>
</tr>
</tbody>
</table>

### Serious adverse events:

- **Infections and infestations**
  - Patients with ≥1 infection: 18 (41%) vs 5 (11%)
  - Deaths: 1 (2%) vs 0

- **Cardiac disorders**
  - Patients with ≥1 cardiac event: 3 (7%) vs 0

- **Musculoskeletal and connective tissue disorders**
  - Patients with ≥1 musculoskeletal event: 2 (5%) vs 1 (2%)

- **Gastrointestinal disorders**
  - Patients with ≥1 gastrointestinal event: 2 (5%) vs 0

- **Skin and subcutaneous tissue disorders**
  - Patients with ≥1 skin event: 0 vs 0

- **Osteoarthritis**
  - Patients with ≥1 osteoarthritis: 0 vs 0

- **Haemolytic anaemia**
  - Patients with ≥1 haemolytic anaemia: 0 vs 0

- **Iron deficiency anaemia**
  - Patients with ≥1 iron deficiency anaemia: 0 vs 0

### Discussion

This study is the first, to our knowledge, phase 2 randomised controlled trial of tocilizumab for the treatment of systemic sclerosis. Week 24 was selected for the primary endpoint assessment because it was assumed that the rapid response of tocilizumab seen in clinical trials in patients with rheumatoid arthritis would translate to a rapid response in patients with systemic sclerosis. We detected no significant difference for the primary endpoint (modified Rodnan skin score at 24 weeks) or for secondary efficacy endpoints at 24 weeks.

Laboratory parameters of interest for tocilizumab (high alanine or aspartate aminotransferase concentration, decreased neutrophils, and decreased platelets) were mostly grade 1 or 2 in intensity and were not associated with clinically relevant sequelae such as hepatic events, serious infections, or serious bleeding events.
and 48 weeks. However, despite not meeting our prespecified definition of minimal clinically important change (≥4.7), the change in modified Rodnan skin score at 24 weeks and 48 weeks was within previously reported minimal clinically important differences for early, diffuse systemic sclerosis (mRSS improvement of 3·2–5·3 from baseline to week 24 or week 24 to week 48). HAQ–DI, clinician global VAS, and other patient-reported outcomes did not differ significantly between groups, but, with the exception of breathing VAS, they favoured tocilizumab over placebo by week 48. Fewer patients receiving tocilizumab than patients receiving placebo had absolute FVC declines of more than 10%.

According to previous definitions, we recorded a clinically meaningful decline in modified Rodnan skin score over 48 weeks in the tocilizumab group compared with the placebo group. Selection of patients with early progressive disease might have contributed to this effect. Eight patients in the placebo group and four patients in the tocilizumab group had progressive skin disease based on an increase in modified Rodnan skin score of more than 5 and at least 25% worsening at any time between baseline and 24 weeks. Skin thickness is a surrogate for internal organ involvement and mortality in systemic sclerosis, and patients with attenuated skin thickness have improved survival and physical function. Conversely, patients with high modified Rodnan skin score have greater internal organ involvement and are at greater risk for death.

Systemic sclerosis has the highest case fatality among rheumatic diseases, with a cumulative survival of 75% at 5 years from diagnosis. Cardiopulmonary involvement is the leading cause of death, and FVC is the primary outcome measure in most trials of systemic sclerosis idiopathic lung disease and idiopathic pulmonary fibrosis. High concentrations of C-reactive protein have been associated with progressive idiopathic lung disease, and observational cohort studies suggest that high C-reactive protein concentration at baseline predicts long-term FVC decline. Although our study was not designed to enrol patients with progressive systemic sclerosis idiopathic lung disease, enrolling those with early progressive skin disease and high acute-phase reactants probably enriched for patients at high risk for idiopathic lung disease. Our data suggest that tocilizumab might have a disease-modifying effect by slowing the decline in lung function of patients with systemic sclerosis.

Treatment with tocilizumab resulted in the downregulation of skin myeloid-associated genes, including genes associated with M2-macrophage. Perivascular macrophages (resident or recruited) contribute to vascular inflammation, and M2-macrophages might have a role in systemic sclerosis skin pathology through the release of inflammatory and profibrotic factors. It is therefore tempting to speculate that the improvement in skin disease in patients treated with tocilizumab might be attributed to the inactivation or depletion of skin macrophages in general and M2-macrophages in particular. Furthermore, treatment with tocilizumab resulted in rapid, sustained reductions in serum concentrations of CCL18, a chemokine associated with M2-macrophage, strengthening the hypothesis that tocilizumab might act, at least in part, by modulating the activity of M2-macrophages. Given that serum CCL18 is a prognostic biomarker that identifies patients at higher risk for the progression of scleroderma lung disease, the effect of tocilizumab on circulating CCL18 could be related to the positive clinical effect of tocilizumab on lung function. These hypotheses will be further explored in the ongoing phase 3 study of tocilizumab in patients with systemic sclerosis.

Tocilizumab tended to inhibit skin expression of genes in the interleukin 6/STAT3 cluster, but only two genes from this cluster (CCL2 and SOCS3) were significantly downregulated by tocilizumab at 24 weeks compared with placebo. This relatively modest effect of tocilizumab on the interleukin 6 pathway could be the result of residual activity of the multiple additional factors (eg, EGF, interleukin 5, HGF, LIF, and BMP2) that activate the STAT3 pathway but are not inhibited by tocilizumab. We did not record specific effects of tocilizumab on genes belonging to the TGF-β clusters or the interferon cluster, suggesting that the clinical effect of tocilizumab might not be mediated through these pathways. However, heterogeneity in the expression of genes associated with these pathways might have prevented the detection of subtle effects of tocilizumab on these genes. Tocilizumab did not affect the serum concentrations of ENPP2, COMP, or POSTN, three biomarkers related to fibrosis, suggesting
that the clinical effect of tocilizumab on skin fibrosis might not be mediated by direct inhibition of the fibrosis pathway by tocilizumab. Finally, treatment with tocilizumab resulted in a significant decrease in the two-gene systemic sclerosis skin biopsy-predicted modified Rodnan skin score. The magnitude of change in predicted modified Rodnan skin score was similar to that recorded clinically.

Overall, after 48 weeks of treatment, safety was consistent with the natural history of systemic sclerosis and the safety profile for tocilizumab. Reflecting the high morbidity and mortality rates of systemic sclerosis, 1 we recorded more serious adverse events, serious infections, and deaths than in clinical trials of tocilizumab for rheumatoid arthritis.

Three patients in the tocilizumab group and one in the placebo group died, and one death was judged by the investigator to be related to tocilizumab. It will be important to assess mortality and causes of death carefully during the phase 3 study of tocilizumab for systemic sclerosis that has just begun. Furthermore, serious infections were more common in the tocilizumab group. These included bronchitis, lung infection, pneumonia, and sepsis, which are similar to the types of events reported during treatment of rheumatoid arthritis with tocilizumab. Small bone osteomyelitis and infected digital ulcers, however, were rare in patients treated with tocilizumab for rheumatoid arthritis and thus could require particular vigilance in patients with systemic sclerosis.

Although the gastrointestinal tract is the most commonly affected internal organ in patients with systemic sclerosis, 1 no patients in our study had gastrointestinal perforations. Most injection site reactions, increases in alanine and aspartate aminotransferase concentrations, and decreases in neutrophil and platelet counts were grade 1 or 2 and involved no clinically relevant sequelae.

No disease-modifying treatment is approved for systemic sclerosis, and its management is based on organ involvement. 3,36 Methotrexate affected skin thickening without establishing positive effects in patients with other organ manifestations, 3,35 whereas oral cyclophosphamide modestly improved lung function and skin thickness in studies of patients with systemic sclerosis idiopathic lung disease. 13 Data from a clinical trial 13 of haemopoietic stem-cell transplantation suggest a disease-modifying effect in patients with systemic sclerosis; thus, haemopoietic stem-cell transplantation can be considered an option for some patients with rapidly progressing and severe systemic sclerosis. Data from the present trial suggest that tocilizumab might have a broader effect on systemic sclerosis skin and lung disease than methotrexate and a more favourable risk–benefit profile than cyclophosphamide, which is associated with clinically significant risk of adverse reactions. 13

The high discontinuation rate should be taken into account when interpreting data from this study. FVC was an exploratory endpoint, and high-resolution CT, which was not done to substantiate the pulmonary function data, is planned for a phase 3 study. Cohort enrichment by retaining patients with a higher probability of skin progression and high acute-phase reactants might have contributed to the treatment responses we recorded. Although our results did not show a significant treatment difference in modified Rodnan skin score at week 24, by week 48 there seemed to be a treatment benefit across most endpoints, which was suggestive of efficacy for tocilizumab in systemic sclerosis. There were no previous positive studies in systemic sclerosis that could be used to inform the design, and the observed effect size for the primary endpoint was smaller than that used in the power calculation for the study (2.7 vs 4.7 modified Rodnan skin score units). The effect size from this phase 2 study and the improvement at week 48 have informed the design of a phase 3 trial (NCT02453256) with the primary endpoint at 48 weeks.

This study shows that skin sclerosis did not differ significantly between groups, but did decrease more in the tocilizumab group than in the placebo group, and lung function showed a clinically relevant improvement. The safety profile was consistent with complications of systemic sclerosis and with the safety profile of tocilizumab. The propensity of patients with systemic sclerosis to develop digital ulcers could increase susceptibility to small bone osteomyelitis and infections of digital ulcers during tocilizumab treatment. Overall, the data warrant further investigation of tocilizumab in a phase 3 trial.

Contributors
DK designed the study, recruited patients, advised on data analysis, interpreted the data, and wrote and revised the report. CPD designed the study, recruited participants, collected, analysed, and interpreted data, and wrote the report. AJ designed the study, oversaw the study, interpreted data, and wrote the report. JMvL and JS designed the study and collected data. TMF and SL collected data. MEA designed the study. MB collected and interpreted data. LC collected and interpreted data and wrote the report. GF collected, analysed, and interpreted data and recruited patients. YA and RL collected, analysed, and interpreted data. JEP did a literature search, designed the study, collected, analysed and interpreted data, and wrote the report. GR designed the study, and collected and interpreted data. VS designed the study and collected data. UM-L analysed and interpreted data. GS designed the study and collected, analysed, and interpreted data. HS designed the study and interpreted data. HC-H collected and analysed data and wrote the report. SD designed the study and collected data. AM and TS collected, analysed, and interpreted data and wrote the report. DEF designed the study, collected and analysed data, and wrote the report. All authors revised the report and approved the final draft for publication.

Declaration of interests
DK has received grants from Bristol-Myers Squibb, Genentech/Roche, National Institute of Allergy and Infectious Diseases, National Institute of Arthritis and Musculoskeletal and Skin Diseases, Patient-Centered Outcomes Research Institute, and the Scleroderma Foundation; and consultancy fees from Actelion, Bayer, Cytori, EMD Serono (Merck), Genkyotex, Gilead, GlaxoSmithKline, Genentech/Roche, Sanofi-Aventis, and Seattle Genetics. CPD has received grants from CSL Behring and GlaxoSmithKline (paid to his institution); consultancy fees from GlaxoSmithKline and Roche (paid to his institution); consultancy fees from Merck-Serono; and speaker fees from Actelion and Bayer. AJ is an employee of and owns stock options in Genentech and has been issued a patent for subcutaneously administered tocilizumab (US 8580264 B2). JMvL has received honoraria from Merck Sharp & Dohme, Pfizer, Roche, and Eli Lilly. MEA has received advisory board and related fees from

www.thelancet.com Published online May 5, 2016 http://dx.doi.org/10.1016/S0140-6736(16)00232-4
Acteon and honoraria from Acteon and Bristol-Myers Squibb and has served as principal investigator of clinical trials for Acteon and Roche. LC has served on an advisory board for Gilead and a data monitoring committee for Cytori. YA has received grants from Bristol-Myers Squibb, Roche/Genentech, Inventiva, Pfizer, Sanofi, and Servier; and personal fees from Acteon, Bayer, Roche/Genentech, Inventiva, Medac, Pfizer, Sanofi, Servier; and UCB. JEP has received research and consulting fees from Roche and consulting fees from Acteon, Bayer, Biogen, Celgene, and Genentech. GR has received honoraria for lectures and advisory boards from Acteon, Bayer, GlaxoSmithKline, Pfizer, and Roche. VS has received a grant and advisory board fees from Roche. UM-L has received grants and speaker or advisory board fees from Roche and Chugai. RL has received grants from Shire, Sanofi, Regeneron, Genetics and Immunology Branch of the National Institutes of Health, UCB, UC San Francisco, Genentech, Biogen, Bristol-Myers Squibb, Inception, Stromedix, Pfizer, Boston University, Bristol-Myers Squibb, and PRISM; and consultancy fees from Shire, Sanofi, Regeneron, Roche/Genentech, Biogen, Incyte, Novartis, Celgene, Bristol-Myers Squibb, Amira, Celldara, Cellcees, Dart Therapeutics, Idera, Inception, Immuneon, Immimmune, Precision Dermatology, Promedior, Zwitter, PRISM, UCB, Acteon, EMD Serono, Akros, Extera, Reneo, Scholar Rock, and Merck. HS is an employee of Roche Products and owns stock in Roche. HC-H is an employee of Genentech. SD is an employee of Roche Diagnostics/F Hoffmann-La Roche Products and owns stock in Roche. HC-H is an employee of Roche Diagnostics/F Hoffmann-La Roche. AM and JS are employees of Genentech. TS is an employee of Genentech, and holds stock, stock options, or bond holdings in Genentech. DEF has received grants from AbbVie, Acteon, Argen, Bristol-Myers Squibb, Gilead, GlaxoSmithKline, National Institutes of Health, Novartis, Pfizer, Roche/Genentech, UCB; consultancy fees from AbbVie, Acteon, Argen, Bristol-Myers Squibb, Celgene, Janssen, Gilead, GlaxoSmithKline, National Institutes of Health, Novartis, Pfizer, and Roche/Genentech, UCB; and speaker fees from AbbVie, Acteon, and UCB. The remaining authors declare no competing interests.

Acknowledgments

We thank the investigators in the faScinate trial (appendix p 1). We also thank Sara Duggan and Meryl Mandel, who provided writing services on behalf of F Hoffmann-La Roche, Sabrina Cheng for her involvement in the trial, and Robyn Rourick, Cécile Holweg, and Jason Hackney for their contributions to the biomarker analyses.

References


