The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/174691

Please be advised that this information was generated on 2019-11-14 and may be subject to change.
EXTENDED REPORT

Treatment outcome in early diffuse cutaneous systemic sclerosis: the European Scleroderma Observational Study (ESOS)


Handling editor Tore K Kvien

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-210503).

For numbered affiliations see end of article.

Correspondence to
Professor Ariane Herrick, The University of Manchester; Salford Royal NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester M13 9PT, UK; ariane.herrick@manchester.ac.uk

Received 12 September 2016
Revised 24 November 2016
Accepted 25 November 2016
Published Online First 10 February 2017

CrossMark


ABSTRACT

Objectives The rarity of early diffuse cutaneous systemic sclerosis (dcSSc) makes randomised controlled trials very difficult. We aimed to use an observational approach to compare effectiveness of currently used treatment approaches.

Methods This was a prospective, observational cohort study of early dcSSc (within three years of onset of skin thickening). Clinicians selected one of four protocols for each patient: methotrexate, mycophenolate mofetil (MMF), cyclophosphamide or ‘no immunosuppressant’. Patients were assessed three-monthly for up to 24 months. The primary outcome was the change in modified Rodnan skin score (mRSS). Confounding by indication at baseline was accounted for using inverse probability of treatment (IPT) weights. As a secondary outcome, an IPT-weighted Cox model was used to test for differences in survival.

Results Of 326 patients recruited from 50 centres, 65 were prescribed methotrexate, 118 MMF, 87 cyclophosphamide and 56 no immunosuppressant. 276 (84.7%) patients completed 12 and 234 (71.7%) 24 months follow-up (or reached last visit date). There were statistically significant reductions in mRSS at 12 months in all groups: –4.0 (−5.2 to −2.7) units for methotrexate, −4.1 (−5.3 to −2.9) for MMF, −3.3 (−4.9 to −1.7) for cyclophosphamide and −2.2 (−4.0 to −0.3) for no immunosuppressant (p value for between-group differences=0.346). There were no statistically significant differences in survival between protocols before (p=0.389) or after weighting (p=0.440), but survival was poorest in the no immunosuppressant group (84.0%) at 24 months.

Conclusions These findings may support using immunosuppressants for early dcSSc but suggest that overall benefit is modest over 12 months and that better treatments are needed.

Trial registration number NCT02339441.

INTRODUCTION

The diffuse cutaneous subtype of systemic sclerosis (dcSSc) is rare (SSc incidence is around 10–20/ million/year,1 of whom approximately 25% will have diffuse disease) but carries high morbidity and mortality due to early internal organ involvement and rapidly progressive, painful skin thickening. Also, 5-year and 10-year survival rates, although improving, are in the order of 68% and 50%, respectively.2 3

At present, there is no drug known to favourably influence disease course. Randomised controlled trials (RCTs) have historically been confounded by disease rarity (only small numbers of patients are recruited, often over long periods) and strict entry criteria meaning that severe cases are often excluded.4 These strict criteria further restrict sample sizes and limit generalisability. Therefore, although RCTs represent a gold standard for assessing drug efficacy, results may not be applicable to
real-life clinical settings. Small trials run the risk of being underpowered, thus potentially yielding false-negative results. The past three decades have seen a number of promising treatments for early dcSSc failing to meet efficacy end points in RCTs: examples include methotrexate (multinational, 71 patients) and anti-transforming growth factor β1 antibody therapy (multinational, 45 patients).

A further difficulty in recruiting into RCTs of early dcSSc is that many clinicians have reservations about placebo therapy in a potentially life-threatening disease and favour immunosuppression, consistent with the European League Against Rheumatism (EULAR) recommendations, which advocate methotrexate for skin manifestations in early dcSSc, although this agent has been shown to be of only limited efficacy. Immunosuppressants are potentially hazardous, especially in patients prone to internal organ disease and infection.

Against this background, our aim was to compare, using an observational approach, the effectiveness of standard treatment approaches (mainly immunosuppressant treatments but including a ‘no immunosuppressant’ option to reflect that some patients or clinicians may choose this approach) in the early management of patients with dcSSc, capturing entry and outcome data in a systematic way. Modern statistical approaches allow robust interrogations of prospective observational studies, as an adjunct to, or even substitute for, RCTs in rare diseases, although the potential of these novel approaches has not yet been realised.

METHODS

Study design

The European Scleroderma Observational Study (ESOS) was a prospective, observational cohort study (ClinicalTrials.gov identifier: NCT02339441), in which standardised data were collected at study entry and at follow-up visits, and entered electronically by investigators at each centre into an electronic case record form. All data were checked by the project coordinator and any inconsistencies were discussed with the chief investigator and (if appropriate) the local principal investigator. The recommended treatment protocols, as decided by the Steering Committee to reflect international best clinical practice, were

1. Methotrexate (oral or subcutaneous with a target dose of 20–25 mg weekly).
2. MMF (500 mg twice daily for 2 weeks increasing to 1 g twice daily).
3. Cyclophosphamide.
4. No immunosuppressant treatment, to give the option of including patients in whom immunosuppression was not felt indicated or appropriate (or declined by the patient).

Patients were recruited at baseline, with subsequent visits scheduled three-monthly for 24 months (or between 12 and 24 months for those patients recruited after September 2013). To have 80% power to detect a difference between two treatment arms of five modified Rodnan skin score (mRSS) units at 12 months would require 63 patients per protocol. Allowing 20% loss to follow-up, and varying numbers recruited to the different protocols, recruitment target was 316 patients.

Patients

Patients were recruited between July 2010 and September 2014. Demographic characteristics including age, gender, smoking habit, ethnicity, antibody status (anti-topoisomerase-1 (anti-Scl70), anti-RNA III polymerase, anticientromere) and presence of visceral organ involvement were recorded for all patients. The algorithms to determine the presence of different types of organ involvement are summarised in online supplementary table S1.

Outcome measures

The primary outcome measure, assessed at each visit, was the change in mRSS over time. All mRSS assessments were performed by those experienced in skin scoring. The mRSS is assessed clinically at 17 body sites on a 0–3 scale (maximum score 51) and measures the extent of skin thickening. It is the most commonly used primary outcome measure in RCTs of dcSSc, reflecting disease severity and predicting mortality. All other outcomes/recorded variables were mainly part of routine clinical practice and are summarised in online supplementary table S2. Secondary end points included pulmonary function (forced vital capacity (FVC: % predicted) and carbon monoxide diffusing capacity (DLCO: % predicted)), quality of life (including the Health Assessment Questionnaire Disability Index (HAQ-DI) and Cochin Hand Function Scale, occurrence of side effects and survival.

Statistical analysis

In an observational study, patient characteristics differ between groups and any differences in outcomes might be driven by those characteristics rather than the treatments (confounding by indication). In each of the analyses (for the different outcome measures), all variables associated with the outcome were considered as confounders.

Differences between protocols at baseline

Kruskal-Wallis test was applied for continuous variables and Fisher’s test for categorical variables.

Influence of baseline characteristics on mRSS at baseline and over time

The association between baseline variables and mRSS was assessed by simple linear regressions, entering each characteristic separately as a predictor of mRSS. To examine how each variable affected the progression of mRSS, the regression equation was modified by adding a term for time and its interaction with the baseline predictor value.

Differences in the changes between groups for all outcomes

Inverse probability of treatment (IPT) weights equalise the distributions of confounders between the treatment groups, thus removing confounding by indication. Treatment probabilities
were computed using multinominal logistic regressions, with the baseline values of the selected confounders as predictors.\textsuperscript{22} Censoring weights rebalance the data such that the distributions of confounders remain unchanged throughout the study. For each observation, the probability of remaining uncensored given the baseline values of the confounders, the initial protocol and a cubic spline for time was calculated using a pooled logistic regression model.\textsuperscript{23} Multiplying both weights yielded the IPT and inverse probability of censoring (IPTC) weights. Weights $>20$ were truncated at that value.\textsuperscript{24}

Treatment effects were assessed using IPTC-weighted linear regression models, which include an intercept, a time term, indicator variables for treatment groups and interactions between time and treatments. The model followed an intention-to-treat approach. Differences in the interaction terms reflected differences in the evolution of outcome.

Cochin hand function data were log-transformed (after adding one to each value) to correct for a highly left-skewed distribution. CIs for the difference of logs were back-transformed, adding one to each value) to correct for a highly left-skewed distribution. CIs for the difference of logs were back-transformed, yielding a percentage difference between predicted baseline and 12-month levels.

Because of missing data at baseline for confounders, multiple imputation by chained equations was applied with STATA 12-month levels. CIs for the difference of logs were back-transformed, yielding a percentage difference between predicted baseline and 12-month levels.

Because of missing data at baseline for confounders, multiple imputation by chained equations was applied with STATA V13.1. Imputations were performed separately for each different outcome model. Moreover, each analysis was restricted to the subset of patients with available outcome data at baseline.

\textbf{Survival analysis}

Kaplan-Meier curves, adjusted using IPT weights, provide estimates of the cumulative probability of surviving in each of the protocols. An IPTC-weighted Cox regression, including indicator variables for the protocols, was used to test for differences in survival between protocols. Both overall and adverse event-free survival were examined.

\textbf{RESULTS}

In total, 326 patients from 50 centres (19 countries) were recruited into the study (figure 1): 160 from mainland Europe and the Middle East, 134 from the UK, 15 from Australia and 17 from North America (six centres from Australia and North America joined after the initial recruitment wave). Not being a native English speaker, the number of patients starting on each protocol differed: 65 (19.9%) methotrexate, 118 (36.2%) MMF, 87 (26.7%) cyclophosphamide and 56 (17.2%) no immunosuppressant treatment. Median (IQR) doses are shown in online supplementary tables S3.

\textbf{Baseline characteristics of patients}

The median mRSS (21, IQR 16–27) and its distribution did not differ across all four treatment groups ($p=0.306$ (table 1). There were significant differences between treatment groups in gender (patients in the cyclophosphamide group less likely to be female, $p=0.003$) and duration of skin thickening (the ‘no immunosuppressant’ group had the longest, $p=0.001$). Also, patients in the cyclophosphamide group were more likely to have had previous immunosuppression ($p=0.007$) or steroid treatment ($p=0.001$). At baseline, 94 (28.8%) patients were taking oral corticosteroids, with a median dose of 10 mg/day (range 2.5–60 mg/day).

\textbf{Organ involvement}

There were significant differences between groups for presence of pulmonary fibrosis, cardiac, renal and muscle involvement. Patients on cyclophosphamide were more likely to have pulmonary fibrosis ($p=0.306$ across groups) or cardiac involvement ($p=0.009$ across groups). Patients in the ‘no immunosuppressant’ group were more likely to have renal involvement ($p=0.039$), and the methotrexate group had more frequent muscle involvement ($p=0.002$).

\textbf{Functional ability}

Scores for the HAQ-DI, Functional Assessment of Chronic Illness Therapy (FACIT) fatigue and Short-Form 36 (SF36) physical and mental indexes did not differ significantly between groups. However, there were significant differences across groups in the cochin hand function scale (CHFS), which was poorest in the cyclophosphamide group ($p=0.025$).

\textbf{Concomitant medications}

As anticipated in a study of patients with early dcSSc, there was substantial use of concomitant medications (see online supplementary table S4).

\textbf{Withdrawals and deaths}

In total, 35 patients (10.7%) died and 42 (12.9%) withdrew from the study (including lost to follow-up). Of the 35 deceased patients, 31 cases were primarily attributed to SSC-related causes (26 most likely primarily cardiorespiratory, 2 renal crises, 2 gastrointestinal (one aspiration) and 1 peritonitis (on peritoneal dialysis following renal crisis)), 3 died of cancer (1 nasopharyngeal, 1 rectal, 1 colorectal) and in 1 case the cause was unknown.

\textbf{Influence of baseline variables on the initial skin score and on skin score trajectory}

Table 2 summarises the effect of different characteristics on the initial mRSS and its subsequent trajectory, as analysed with linear regression.

Using the associations described by table 2, the confounders identified for the skin score were age, duration of skin thickening, current or previous steroid use, anti-topoisomerase, anti-RNA polymerase III, pulmonary fibrosis, pulmonary hypertension, cardiac, renal and muscle involvement, as well as HAQ-DI, Cochin hand function and FACIT fatigue scores (see online supplementary table S5 for lists of confounders and online supplementary tables S6–S13 for each model’s confounder selection process).

\textbf{Changes in skin score over time in the different treatment groups}

The mean change in mRSS after 12 and 24 months was $-2.9$ and $-6.7$ units. Based on a weighted regression model, there were
Figure 1  Progression of patients through the study.


Clinical and epidemiological research

- 348 patients assessed for eligibility
  - 22 ineligible
    - 11 no proximal skin thickening
    - 2 previous immunosuppression
    - 2 skin thickening over 3 years
    - 1 recurrence of cancer
    - 1 revised diagnosis (not dCSc)
    - 5 ineligible for other reasons

- 326 eligible

**Baseline**
- 65 assigned methotrexate
  - 12 study exits
    - 4 deaths
    - 4 withdrawals
    - 4 protocol violations
  - 19 protocol changes
    - 6 adverse reactions
    - 9 progression
    - 4 other reasons

- 118 assigned mycophenolate mofetil
  - 13 study exits
    - 4 deaths
    - 8 withdrawals
    - 1 protocol violation
  - 15 protocol changes
    - 8 adverse reactions
    - 3 progression
    - 4 other reasons

- 87 assigned cyclophosphamide
  - 16 study exits
    - 10 deaths
    - 3 withdrawals
    - 1 protocol violation
  - 14 protocol changes
    - 6 adverse reactions
    - 4 progression
    - 4 other reasons

- 56 assigned no immunosuppressant
  - 9 study exits
    - 6 deaths
    - 2 withdrawals
    - 1 protocol violation
  - 8 protocol changes
    - 4 progression
    - 4 other reasons

**12 months**
- 53 treatment ongoing (incl. 5 missing)
- 105 treatment ongoing (incl. 17 missing)
- 71 treatment ongoing (incl. 6 missing)
- 47 treatment ongoing (incl. 6 missing)

**24 months**
- 42 complete study
- 69 complete study
- 48 complete study
- 30 complete study

---

- Withdrawals include patients who left the study for the following reasons: unable to travel, changed hospital, enrolled in a randomized clinical trial, no longer willing to participate or lost to follow-up with no given reason. Among the 8 patients who were lost to follow-up in the first year, 5 patients were due for their 24-month visit after the study deadline.

- Protocol violations brought about a premature exit from the study. There were 20 protocol violations mid-study: 9 who started rituximab (as their main treatment), 5 who received a stem cell transplantation, 4 who started azathioprine (not as part of a post-cyclophosphamide transfer), 1 who started tocilizumab and 1 who entered a randomized clinical trial.

- Protocol changes designate patients who changed immunosuppressants, started taking or stopped receiving one (in non-observance to their initial assigned protocol). Changes that were a part of a post-cyclophosphamide transfer were not considered protocol changes. If a patient deviated from her/his initial protocol more than once, only the initial change (along with its reason) was recorded in the flow-chat.

- There were 58 post-cyclophosphamide transfers: 40 patients went on to mycophenolate mofetil, 9 on to azathioprine, 5 on to methotrexate and 4 on to no immunosuppressant.

- In the no immunosuppressant cohort, 10 out of 56 patients deviated from their protocol at any point: 2 switched to methotrexate, 5 to mycophenolate mofetil and 3 to cyclophosphamide. 7 of those 10 cases were linked to pulmonary complications, one to associated inflammatory arthritis, one to skin thickening and one because of preference.
### Table 1  Baseline characteristics and differences between protocols

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Protocol 1 Methotrexate n=65</th>
<th>Protocol 2 Mycophenolate n=118</th>
<th>Protocol 3 Cyclophosphamide n=87</th>
<th>Protocol 4 No immunosuppressant n=56</th>
<th>P Value*</th>
<th>Total n=326</th>
<th>Missing at baseline, no. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>52.5 (41.5–61.1)</td>
<td>50 (40.7–61.1)</td>
<td>54.1 (45.7–60.5)</td>
<td>52.3 (42.3–60.3)</td>
<td>0.324</td>
<td>52.3 (43–60.8)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Female, no. (%)</td>
<td>50 (76.9%)</td>
<td>94 (79.7%)</td>
<td>49 (56.3%)</td>
<td>40 (71.4%)</td>
<td>0.003</td>
<td>233 (71.5%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Months since onset of skin thickening</td>
<td>10.3 (5.5–20.8)</td>
<td>12.6 (7.3–21.8)</td>
<td>10.2 (5.9–14.5)</td>
<td>16.5 (8.7–27)</td>
<td>0.001</td>
<td>11.9 (7–21)</td>
<td>18 (5.5%)</td>
</tr>
<tr>
<td>Years since onset of first non-Raynaud’s manifestation</td>
<td>1.3 (0.7–2.1)</td>
<td>1.4 (0.9–2)</td>
<td>1.0 (0.6–1.4)</td>
<td>1.7 (0.9–2.5)</td>
<td>0.001</td>
<td>1.2 (0.8–2)</td>
<td>6 (1.8%)</td>
</tr>
<tr>
<td>Years since onset of Raynaud’s phenomenon</td>
<td>1.3 (0.7–2.5)</td>
<td>1.9 (1.1–3.1)</td>
<td>1.4 (0.9–2.4)</td>
<td>2.2 (1.1–3.7)</td>
<td>0.014</td>
<td>1.7 (1–2.9)</td>
<td>22 (6.7%)</td>
</tr>
<tr>
<td>Previous immunosuppress use, no. (%)‡</td>
<td>3 (4.6%)</td>
<td>5 (4.2%)</td>
<td>15 (17.2%)</td>
<td>3 (5.4%)</td>
<td>0.007</td>
<td>26 (8%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Current or previous steroid use, no. (%)§</td>
<td>33 (50.8%)</td>
<td>39 (34.2%)</td>
<td>51 (58.6%)</td>
<td>18 (32.1%)</td>
<td>0.001</td>
<td>141 (43.8%)</td>
<td>4 (1.2%)</td>
</tr>
<tr>
<td>Current smoker (%)</td>
<td>11 (16.9%)</td>
<td>17 (15.3%)</td>
<td>18 (21.7%)</td>
<td>11 (20.8%)</td>
<td>0.646</td>
<td>57 (18.3%)</td>
<td>14 (4.3%)</td>
</tr>
<tr>
<td>History of cancer, no. (%)</td>
<td>5 (7.8%)</td>
<td>4 (3.4%)</td>
<td>4 (4.7%)</td>
<td>7 (12.5%)</td>
<td>0.121</td>
<td>20 (6.2%)</td>
<td>4 (1.2%)</td>
</tr>
<tr>
<td>Caucasian, no. (%)</td>
<td>54 (83.1%)</td>
<td>92 (78.0%)</td>
<td>77 (88.5%)</td>
<td>49 (87.5%)</td>
<td>0.201</td>
<td>272 (83.4%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>mRSS (0–51)</td>
<td>21 (17–24)</td>
<td>21 (16–27)</td>
<td>22 (17–29)</td>
<td>20 (15.5–26)</td>
<td>0.306</td>
<td>21 (16–27)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Haemoglobin (g/L)</td>
<td>12.7 (118–136)</td>
<td>126 (118–137)</td>
<td>130 (116–140)</td>
<td>130 (118–139)</td>
<td>0.721</td>
<td>128 (118–137)</td>
<td>12 (4%)</td>
</tr>
<tr>
<td>White blood count (×10^9/L)</td>
<td>7.7 (6.7–9.4)</td>
<td>7.9 (6.4–9.5)</td>
<td>8.8 (7.3–10.6)</td>
<td>8.0 (6.7–9.4)</td>
<td>0.029</td>
<td>8 (6.8–9.9)</td>
<td>14 (4.3%)</td>
</tr>
<tr>
<td>Platelets (×10^9/L)</td>
<td>300 (254–337)</td>
<td>298 (246–370)</td>
<td>309 (259–359)</td>
<td>281 (250–337)</td>
<td>0.459</td>
<td>298 (253–356)</td>
<td>15 (4.6%)</td>
</tr>
<tr>
<td>ESR (mm/hour)</td>
<td>17 (10–29)</td>
<td>18 (8–30)</td>
<td>21 (9–48)</td>
<td>20 (10–35)</td>
<td>0.341</td>
<td>18 (8–34)</td>
<td>77 (23.6%)</td>
</tr>
<tr>
<td>CRP (mg/L)</td>
<td>4.0 (2.2–7.5)</td>
<td>5.0 (1.1–11)</td>
<td>5.9 (3.2–20.0)</td>
<td>4.8 (3.0–11.7)</td>
<td>0.026</td>
<td>5 (2.1–11.8)</td>
<td>90 (27.6%)</td>
</tr>
<tr>
<td>Anti-topoisomerase (anti-ScI70), no. (%)</td>
<td>20 (31.3%)</td>
<td>49 (42.6%)</td>
<td>39 (45.3%)</td>
<td>18 (33.3%)</td>
<td>0.228</td>
<td>126 (39.5%)</td>
<td>7 (2.1%)</td>
</tr>
<tr>
<td>Anti-RNA polymerase III, no. (%)</td>
<td>9 (18.8%)</td>
<td>23 (23.2%)</td>
<td>9 (13.0%)</td>
<td>9 (19.6%)</td>
<td>0.433</td>
<td>50 (19.1%)</td>
<td>64 (19.6%)</td>
</tr>
<tr>
<td>Anticentromere, no. (%)</td>
<td>9 (14.1%)</td>
<td>6 (5.4%)</td>
<td>4 (4.7%)</td>
<td>3 (5.7%)</td>
<td>0.147</td>
<td>22 (7%)</td>
<td>12 (3.7%)</td>
</tr>
<tr>
<td>Organ involvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulmonary fibrosis, no. (%)</td>
<td>8 (12.3%)</td>
<td>13 (11%)</td>
<td>21 (24.1%)</td>
<td>5 (8.9%)</td>
<td>0.036</td>
<td>47 (14.4%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>FVC (% predicted)</td>
<td>93.7 (81–106)</td>
<td>90 (75–102)</td>
<td>82.5 (68.5–96.5)</td>
<td>90 (75–100)</td>
<td>0.026</td>
<td>89 (75–102)</td>
<td>19 (5.8%)</td>
</tr>
<tr>
<td>DLCO (% predicted)</td>
<td>73 (62–83)</td>
<td>64.5 (48–77)</td>
<td>57 (41–73)</td>
<td>64 (52–75)</td>
<td>&lt;0.0005</td>
<td>64 (50–78)</td>
<td>35 (10.7%)</td>
</tr>
<tr>
<td>Pulmonary hypertension, no. (%)</td>
<td>4 (6.3%)</td>
<td>7 (5.9%)</td>
<td>10 (11.5%)</td>
<td>5 (8.9%)</td>
<td>0.488</td>
<td>26 (8%)</td>
<td>1 (0.3%)</td>
</tr>
<tr>
<td>sPAP or RVSP (mm Hg)†</td>
<td>27 (23–32)</td>
<td>27 (20–33)</td>
<td>30 (23–37)</td>
<td>27 (23–35)</td>
<td>0.472</td>
<td>29 (21–34)</td>
<td>124 (38%)</td>
</tr>
<tr>
<td>Cardiac involvement, no. (%)</td>
<td>5 (7.7%)</td>
<td>8 (6.8%)</td>
<td>19 (21.8%)</td>
<td>7 (13.0%)</td>
<td>0.009</td>
<td>39 (12%)</td>
<td>2 (0.6%)</td>
</tr>
<tr>
<td>Renal involvement, no. (%)§</td>
<td>1 (1.5%)</td>
<td>13 (11%)</td>
<td>10 (11.5%)</td>
<td>8 (14.3%)</td>
<td>0.039</td>
<td>32 (9.8%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>eGFR (ml/min)</td>
<td>90 (63–90)</td>
<td>84.5 (60–90)</td>
<td>90 (60–90)</td>
<td>80 (60–90)</td>
<td>0.339</td>
<td>85 (60–90)</td>
<td>92 (28.2%)</td>
</tr>
<tr>
<td>Renal crisis</td>
<td>0 (0%)</td>
<td>8 (6.8%)</td>
<td>4 (4.6%)</td>
<td>4 (7.1%)</td>
<td>0.110</td>
<td>16 (4.9%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Plasma creatinine (μmol/L)</td>
<td>63.5 (55–72)</td>
<td>69 (55–83)</td>
<td>64.5 (53–86)</td>
<td>64 (56–77)</td>
<td>0.422</td>
<td>65 (55–79.5)</td>
<td>70 (21.5%)</td>
</tr>
</tbody>
</table>

Continued
Table 1 Continued

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Protocol 1 Methotrexate n=65</th>
<th>Protocol 2 Mycophenolate n=118</th>
<th>Protocol 3 Cyclophosphamide n=87</th>
<th>Protocol 4 No immunosuppressant n=56</th>
<th>P Value*</th>
<th>Total n=326</th>
<th>Missing at baseline, no. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any Gl involvement, no. (%)</td>
<td>25 (38.5%)</td>
<td>32 (27.1%)</td>
<td>30 (34.5%)</td>
<td>26 (46.4%)</td>
<td>0.078</td>
<td>113 (34.7%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Muscle involvement, no. (%)</td>
<td>12 (18.5%)</td>
<td>9 (7.6%)</td>
<td>10 (11.5%)</td>
<td>0 (0%)</td>
<td>0.002</td>
<td>31 (9.5%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Current digital ulcers</td>
<td>10 (15.4%)</td>
<td>17 (14.4%)</td>
<td>17 (19.5%)</td>
<td>11 (19.6%)</td>
<td>0.705</td>
<td>55 (16.9%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Patient questionnaire data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAQ-DI (0–3)</td>
<td>1.1 (0.4–1.8)</td>
<td>1 (0.5–1.6)</td>
<td>1 (0.4–1.9)</td>
<td>0.7 (0.1–1.5)</td>
<td>0.400</td>
<td>1 (0.4–1.8)</td>
<td>19 (5.8%)</td>
</tr>
<tr>
<td>FACIT fatigue score (0–52)</td>
<td>33 (21.5–42.5)</td>
<td>30 (20–37)</td>
<td>31 (17–40)</td>
<td>37 (21–44)</td>
<td>0.165</td>
<td>31 (20–41)</td>
<td>16 (4.9%)</td>
</tr>
<tr>
<td>SF36 physical score (0–100)</td>
<td>39.2 (30.9–45.1)</td>
<td>36.4 (30.8–43)</td>
<td>36 (27.2–44.5)</td>
<td>39.2 (32–48.4)</td>
<td>0.203</td>
<td>37.4 (29.9–45)</td>
<td>15 (4.6%)</td>
</tr>
<tr>
<td>SF36 mental score (0–100)¶</td>
<td>37.2 (33.3–44)</td>
<td>36.8 (32.2–41.9)</td>
<td>39.5 (35.7–44.7)</td>
<td>40.8 (35.9–44.1)</td>
<td>0.029</td>
<td>38.3 (34.3–44)</td>
<td>15 (4.6%)</td>
</tr>
<tr>
<td>Cochin hand function score (0–90)  **</td>
<td>11.5 (2–30.5)</td>
<td>10 (4–24)</td>
<td>16 (5–40)</td>
<td>6.5 (0–23)</td>
<td>0.025</td>
<td>11 (3–29)</td>
<td>96 (29.4%)</td>
</tr>
</tbody>
</table>

Median (IQR) unless otherwise indicated.
*P indicates significance of Kruskal-Wallis test (for continuous variables) or Fisher’s exact test (for categorical variables).
†Of the 26 patients who had previously received immunosuppressant therapy, in 2 patients this was for cancer.
‡Despite the significant p-value for the Kruskal-Wallis test, post hoc tests reject any between-group differences in the SF36 mental scores.
¶Renal involvement is defined as renal crisis and/or moderate-to-severe renal impairment.
** Cochin hand function scores were not performed in all centres because of translational issues.
CRP, C reactive protein; DLCO, carbon monoxide diffusing capacity; eGFR, estimated glomerular filtration rate; ESR, erythrocyte sedimentation rate; FACIT, Functional Assessment of Chronic Illness Therapy; FVC, forced vital capacity; GI, gastrointestinal; HAQ-DI, Health Assessment Questionnaire Disability Index; mRSS, modified Rodnan skin score (17 sites); RVSP, right ventricular systolic pressure; SF36, Short-Form 36; sPAP, systolic pulmonary artery pressure.

statistically significant reductions in mRSS in all four treatment groups at 12 months (−4.0 (−5.2 to −2.7) units for methotrexate, −4.1 (−5.3 to −2.9) for MMF, −3.3 (−4.9 to −1.7) for cyclophosphamide and −2.2 (−4.0 to −0.3) for the no immunosuppressant group), but the differences between treatments were not significant (p=0.346) (table 3 and figure 2).

Changes in secondary outcomes over time in the different treatment groups

Lung function
After adjusting for potential confounders, the change rates of FVC and DLCO were not significantly different in the four treatment groups (p=0.460 and 0.505) (table 3).

However, in a subset of patients with pulmonary fibrosis or suspected pulmonary fibrosis (cases confirmed on high-resolution CT (HRCT) irrespective of FVC or DLCO, or with one of the following if HRCT not performed: FVC or DLCO under 55% predicted or definite basal shadowing on X-ray), there was a significant difference in the change rate of FVC over time (p=0.035). Patients initially prescribed cyclophosphamide demonstrated 7.4% absolute increase in FVC (% predicted) compared with 2.0% decrease for methotrexate, 3.2% increase for MMF and 4.0% increase for the ‘no immunosuppressant’ group (table 3).

Functional ability and hand function
Changes over time for the HAQ-DI and CHFS did not differ between protocols (p=0.130 and 0.073), regardless of adjusting (table 3).

Development of internal organ involvement
This is described in online supplementary figure S2.

Comparison of survival between treatment protocols
Survival was lowest in the no immunosuppressant group at both 12 and 24 months but differences between protocols were not statistically significant either before (p=0.389) or after weighting (p=0.440). In the adjusted model, at 24 months, those in the no immunosuppressant group had a predicted survival rate of 84.0% compared with 94.1% for methotrexate, 88.8% for MMF and 90.1% for cyclophosphamide (figure 3). Patients with lung involvement (pulmonary fibrosis and/or hypertension) at baseline had significantly poorer survival than those without: at 24 months, their predicted survival rate was 74.6% versus 91.7% (p<0.0005) and similarly for cardiac involvement, 76.3% versus 90.7% (p<0.0005).

Adverse effects
Of the 75, 182 and 101 patients who were ever on methotrexate, MMF or cyclophosphamide, respectively, 29 (38.7%), 40 (22.0%) and 23 (22.8%) were reported to have had side effects, necessitating drug discontinuation in 9 (12.0%), 14 (7.7%) and 5 (4.5%) patients, respectively. A survival analysis on protocol exit due to adverse effects showed no differences in the tolerability of the three treatments (p=0.212) (see online supplementary figure S3).

**DISCUSSION**

Our main findings were, first, that there were no significant differences in outcome between the four treatment protocols (methotrexate, MMF, cyclophosphamide, no immunosuppression), although there may be a signal in favour of immunosuppression for early dcSSc. Although skin score improved in all treatment groups, this was least in the no immunosuppressant category, who also had the highest mortality. Second, ESOS
Table 2  Associations between baseline characteristics and skin score

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>(A) Association with baseline mRSS</th>
<th>(B) Effect on mRSS evolution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline predictor coefficient (95% CI)</td>
<td>p(1)</td>
</tr>
<tr>
<td>Age, per 10 years</td>
<td>0.9 (0.3 to 1.6)</td>
<td>0.004</td>
</tr>
<tr>
<td>Female</td>
<td>−1.8 (−3.7 to 0.2)</td>
<td>0.072</td>
</tr>
<tr>
<td>Months since onset of skin thickening</td>
<td>−0.1 (−0.2 to 0)</td>
<td>0.156</td>
</tr>
<tr>
<td>Previous immunosuppressant use</td>
<td>−2.1 (−5.3 to 1.2)</td>
<td>0.205</td>
</tr>
<tr>
<td>Current or previous steroid use</td>
<td>1.2 (−0.6 to 3.0)</td>
<td>0.191</td>
</tr>
<tr>
<td>mRSS (5–51), per 5 units</td>
<td></td>
<td>0.9</td>
</tr>
<tr>
<td>White blood count (×10⁹/L)</td>
<td>0.5 (0.2 to 0.8)</td>
<td>0.002</td>
</tr>
<tr>
<td>ESR (mm/hour)</td>
<td>0.1 (0 to 0.1)</td>
<td>0.003</td>
</tr>
<tr>
<td>CRP (mg/L)</td>
<td>0 (0 to 0.2)</td>
<td>0.003</td>
</tr>
<tr>
<td>Anti-topoisomerase (anti-Scl70)</td>
<td>−2.6 (−4.4 to −0.8)</td>
<td>0.005</td>
</tr>
<tr>
<td>Anti-RNA polymerase III</td>
<td>4.5 (2.1 to 6.9)</td>
<td>&lt;0.0005</td>
</tr>
<tr>
<td>Anticitromere</td>
<td>−0.4 (−3.9 to 3.1)</td>
<td>0.816</td>
</tr>
</tbody>
</table>

Organ involvement

| Pulmonary fibrosis | 2.9 (0.4 to 5.4) | 0.021 | −3.5 | 0.3 (−0.6 to 1.2) | 0.534 |
| DLCO (% predicted) | −0.1 (−0.1 to 0) | 0.013 | −5.5 | 0 (0 to 0) | 0.005 |
| Pulmonary hypertension | 2.5 (−0.7 to 5.8) | 0.128 | −3.3 | −2.1 (−3.3 to −0.8) | 0.001 |
| Cardiac involvement | 2.5 (−0.2 to 5.2) | 0.075 | −3.5 | 0 (−1.0 to 0.9) | 0.929 |
| Renal involvement | 2.2 (−0.7 to 5.2) | 0.140 | −3.3 | −1.6 (−2.6 to −0.5) | 0.004 |
| Any GI involvement, no. (%) | 2.2 (0.3 to 4.0) | 0.021 | −3.4 | −0.1 (−0.8 to 0.5) | 0.664 |
| Muscle involvement | 1.2 (−1.8 to 4.2) | 0.425 | −3.2 | −2 (−2.9 to −1.0) | <0.0005 |
| Current digital ulcers | 2.5 (0.1 to 4.8) | 0.038 | −3.3 | −0.9 (−1.8 to 0) | 0.047 |
| Patient questionnaire data | | | | |
| HAQ-DI (0–3) | 3.5 (2.5 to 4.6) | <0.0001 | −3.0 | −0.4 (−0.8 to 0) | 0.039 |
| FACIT fatigue score (0–52), per 10 units | −1.3 (−2.0 to −0.6) | <0.0005 | −3.1 | −0.1 (−0.3 to 0.2) | 0.484 |
| SF36 physical score (0–100) | −0.2 (−0.3 to −0.1) | <0.0001 | −3.3 | 0 (0 to 0) | 0.873 |
| SF36 mental score (0–100) | 0.1 (−0.1 to 0.2) | 0.297 | −1.8 | 0 (−0.1 to 0) | 0.081 |
| Cochin hand function score (0–90), per 10 units | 1.4 (0.9 to 1.9) | <0.0001 | −3.7 | 0 (−0.2 to 0.2) | 0.979 |

Example for interpretation of results: the presence of anti-RNA polymerase III is associated with (A) a higher mRSS by 4.5 units at baseline and (B) losing an extra 2.1 units per year compared with an average of −3.0 units per year for all patients.

*Renal involvement is defined as renal crisis and/or moderate-to-severe renal impairment.

p(1): Significance p value for characteristic coefficient in linear regression of baseline mRSS on baseline predictor.
p(2): Significance p value for interaction coefficient between time and baseline characteristic in a longitudinal regression model.

CRP, C reactive protein; DLCO, carbon monoxide diffusing capacity; ESR, erythrocyte sedimentation rate; FACIT, Functional Assessment of Chronic Illness Therapy; FVC, forced vital capacity; GI, gastrointestinal; HAQ-DI, Health Assessment Questionnaire Disability Index; mRSS, modified Rodnan skin score (17 sites); SF36, Short-Form 36.

confirms the relative effectiveness of cyclophosphamide in patients with pulmonary fibrosis.25 26

An important point when interpreting our findings (and therefore a note of caution) is that the ‘no immunosuppressant’ group was not a control group. Patients in this group had a longer disease duration than the other three groups and were more likely to have renal involvement.

Our findings lend support to two recently published studies (the Autologous Stem Cell Transplantation International Scleroderma trial (ASTIS) trial of autologous stem cell transplantation27 and the Scleroderma Lung Study (SLS) II (comparing MMF and cyclophosphamide),26 which suggest benefit, including in mRSS, from immunosuppression (as did SLS I25). In ASTIS, those patients randomised to cyclophosphamide had an 8.8 unit fall in mRSS (from 23.8) at 24 months (compared with 3.3 in ESOS over 12 months), but the cyclophosphamide protocol was more intense, and the patients had more severe disease (patients with the highest mRSS at baseline tend to improve much more4 as also demonstrated by our own findings (table 2)). mRSS fell by 19.9 units in those patients randomised to stem cell transplantation7,7 and therefore intensive immunosuppression). In SLS I,25 patients with dcSSc randomised to cyclophosphamide experienced a 5.3 unit fall in mRSS at 12 months (compared with 3.3 in ESOS), whereas mRSS fell by 1.7 on placebo (compared with 2.2 units in the ESOS ‘no immunosuppressant’ group). In SLS II,26 mRSS at 24 months fell 4.9 units on MMF (compared with 4.1 units in ESOS at 12 months) and by 5.4 after 12 months treatment with cyclophosphamide, although these values are not directly comparable because they relate to patients with limited cutaneous and dcSSc combined.

The methodological strength of ESOS, which built upon experience gained in a previous, smaller observational study,28 and its design: its standardised protocols emulated the conditions of a clinical trial, and although not randomised, patients were enrolled into four homogenous treatment arms with well-defined interventions and a systematic record of protocol changes and exits. Entry criteria were deliberately inclusive: RCTs often exclude patients with internal organ involvement and for whom immunosuppression is most likely to be
Table 3  Predicted yearly changes in outcomes and survival rates according to initial protocol, with and without adjusting (95% CI)

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Model specification</th>
<th>p Value</th>
<th>Methotrexate</th>
<th>Mycophenolate mofetil</th>
<th>Cyclophosphamide</th>
<th>No immunosuppressant</th>
</tr>
</thead>
<tbody>
<tr>
<td>mRSS (0–51)</td>
<td>No adjusting, n=326</td>
<td>0.252</td>
<td>−4.4 (−5.7 to −3.2)</td>
<td>−3.8 (−4.9 to −2.8)</td>
<td>−3.5 (−5.0 to −2.0)</td>
<td>−2.4 (−3.9 to −1)</td>
</tr>
<tr>
<td></td>
<td>Adjusting for confounding (weighted model), n=326</td>
<td>0.346</td>
<td>−4.0 (−5.2 to −2.7) (n=65)</td>
<td>−4.1 (−5.3 to −2.9) (n=118)</td>
<td>−3.3 (−4.9 to −1.7) (n=87)</td>
<td>−2.2 (−4.0 to −0.3) (n=56)</td>
</tr>
<tr>
<td>FVC (% predicted)</td>
<td>No adjusting, n=307</td>
<td>0.045</td>
<td>−1.7 (−4.4 to 1.0)</td>
<td>2.1 (−0.1 to 4.3)</td>
<td>4.0 (0.5 to 7.5)</td>
<td>2.7 (−0.8 to 6.2)</td>
</tr>
<tr>
<td></td>
<td>Adjusting for confounding (weighted model), n=307</td>
<td>0.460</td>
<td>−0.5 (−3.7 to 2.6) (n=59)</td>
<td>2.0 (−0.7 to 4.6) (n=11)</td>
<td>3.3 (−0.6 to 7.2) (n=84)</td>
<td>2.0 (−1.6 to 5.6) (n=53)</td>
</tr>
<tr>
<td></td>
<td>Subset with PF on HRCT, weighted, n=1291</td>
<td>0.035</td>
<td>−2.0 (−5.9 to 2.0) (n=19)</td>
<td>3.2 (−0.6 to 7.0) (n=31)</td>
<td>7.4 (2.2 to 12.7) (n=57)</td>
<td>4.0 (−1 to 9.0) (n=22)</td>
</tr>
<tr>
<td>DLCO (% predicted)</td>
<td>No adjusting, n=291</td>
<td>0.703</td>
<td>−1.2 (−3.5 to 1.0)</td>
<td>0.3 (−1.7 to 2.3)</td>
<td>0.6 (−2.0 to 3.3)</td>
<td>−0.3 (−4.2 to 3.7)</td>
</tr>
<tr>
<td></td>
<td>Adjusting for confounding (weighted model), n=291</td>
<td>0.505</td>
<td>−1.6 (−3.8 to 0.6) (n=51)</td>
<td>0.8 (−1.5 to 3.1) (n=110)</td>
<td>0.1 (−3.1 to 3.2) (n=79)</td>
<td>−0.1 (−3.7 to 3.4) (n=51)</td>
</tr>
<tr>
<td></td>
<td>Subset with PF on HRCT, weighted, n=116†</td>
<td>0.809</td>
<td>−0.8 (−5.4 to 3.8) (n=12)</td>
<td>1.9 (−1.8 to 5.6) (n=31)</td>
<td>1.8 (−1.9 to 5.4) (n=53)</td>
<td>1.6 (−2.6 to 5.7) (n=20)</td>
</tr>
<tr>
<td>HAQ-DI (0–3)</td>
<td>No adjusting, n=307</td>
<td>0.070</td>
<td>−0.1 (−0.2 to 0.1)</td>
<td>0 (−0.1 to 0.1)</td>
<td>−0.2 (−0.3 to −0.1)</td>
<td>0.1 (−0.1 to 0.2)</td>
</tr>
<tr>
<td></td>
<td>Adjusting for confounding (weighted model), n=307</td>
<td>0.130</td>
<td>−0.1 (−0.3 to 0) (n=59)</td>
<td>0 (−0.1 to 0.1) (n=113)</td>
<td>−0.1 (−0.3 to 0) (n=81)</td>
<td>0.1 (−0.1 to 0.2) (n=54)</td>
</tr>
<tr>
<td>Cochin hand function (0–90‡)</td>
<td>No adjusting, n=230</td>
<td>0.072</td>
<td>−1.1 (−3.5 to 2.2)</td>
<td>−0.3 (−1.6 to 1.2)</td>
<td>−3.1 (−5.1 to −0.5)</td>
<td>1.1 (−0.3 to 3.0)</td>
</tr>
<tr>
<td></td>
<td>Adjusting for confounding (weighted model), n=230</td>
<td>0.073</td>
<td>−1.4 (−3.5 to 1.4) (n=36)</td>
<td>−0.6 (−2.0 to 1.1) (n=103)</td>
<td>−2.4 (−4.7 to 0.6) (n=49)</td>
<td>1.7 (−0.1 to 4.0) (n=42)</td>
</tr>
<tr>
<td>Survival</td>
<td>No adjusting, n=326</td>
<td>0.389</td>
<td>(12 months) 93.5% (24 months) 93.5%</td>
<td>(12 months) 96.5% (24 months) 89.3%</td>
<td>(12 months) 88.1% (24 months) 85.4%)</td>
<td>(12 months) 88.9% (24 months) 85.1%</td>
</tr>
<tr>
<td></td>
<td>Adjusting for confounding (weighted model), n=326</td>
<td>0.440</td>
<td>(12 months) 94.1% (24 months) 94.1%</td>
<td>(12 months) 96.1% (24 months) 88.8%</td>
<td>(12 months) 91.7% (24 months) 90.1%</td>
<td>(12 months) 88.6% (24 months) 84.0%</td>
</tr>
</tbody>
</table>

Significance p: Fisher’s test for equality of change rates between protocols, for each outcome variable.

*Results are reported in terms of changes after 12 months. However, all study data (from baseline to the 24-month end point) were used in estimation. To obtain 24-month changes, multiply results above by 2.

†For the subanalysis involving the subset of patients with pulmonary fibrosis at baseline, patients with definite bibasal pulmonary fibrosis confirmed on HRCT were included, irrespective of FVC value. If no HRCT scan was performed at baseline, an FVC<55%, DLCO<55% predicted or definite bibasal shadowing on X-ray was also a basis for inclusion.

‡Changes expressed in units for the Cochin regression are an approximation derived from the 95% CI of percentage changes between baseline and 12 months (on a scale shifted by one unit), applied to the predicted baseline values for each group in the original scale.

FVC, forced vital capacity; DLCO, carbon monoxide diffusing capacity; HAQ-DI, Health Assessment Questionnaire Disability Index; HRCT, high-resolution CT; mRSS, modified Rodnan skin score (17 sites); PF, pulmonary fibrosis.
Figure 2  Modified Rodnan skin score (mRSS) during baseline and follow-up visits, by initial protocol. For each group of patients, according to their initial protocol, the distribution of the skin score is illustrated on the left-hand side by box and whisker plots (indicating the median and IQR) at baseline, 12 and 24 months. On the right-hand side, the distribution of individual 1-year changes in the skin score is described by histograms and a kernel density estimate. In addition, a vertical green line indicates the value of the average 1-year change in the skin score, irrespective of treatment choice. The bottom panel in the figure describes the estimated changes in mRSS (with 95% CI) according to initial protocol, based on the results from the adjusted model (described in Table 3).

beneficial. By recruiting 326 patients from 50 centres, ESOS represents a large cohort of patients with very early dcSSc (median duration of skin thickening 11.9 months): its data will serve as a benchmark when designing and interpreting future clinical trials. This is especially relevant with a number of novel treatment approaches currently being explored including biological agents. For example, in a recent RCT of tocilizumab, mRSS fell over 24 weeks by 3.9 units from 26 in the 43 tocilizumab-treated patients and by 1.2 units from 26 in the 44 placebo-treated patients, this latter fall comparable to the ESOS ‘no immunosuppressant’ response. In comparing between these studies, the higher baseline mRSS in the tocilizumab study should be borne in mind.

The main weakness of observational studies is that each patient’s outcome on her/his treatment arm cannot be completely disentangled from her/his initial characteristics. For instance, ESOS has verified that patients with lung and cardiac involvement tend to be prescribed cyclophosphamide. However, adjusting using IPT weights minimises the problem of confounding by indication.

In conclusion, observational studies offer a rich population-wide perspective assessing treatment effects in a real-world setting. ESOS achieved its aim of following a large international cohort of patients with early dcSSc over 2 years, each of whom was treated according to one of four protocols. The message for clinicians is that there is a weak signal to support using immunosuppressants for early dcSSc (and in particular cyclophosphamide for patients with pulmonary fibrosis). However, it is clear that there remains a pressing need for the development of more effective and targeted treatments.

Figure 3  Kaplan-Meier estimated survival curves by treatment group.
Clinical and epidemiological research

20 University of Liverpool, Aintree University Hospital, Liverpool, UK
21 Service de Médecine Intérim, Hôpital Bretonneau Tours Cedex, France
22 Centre for Rheumatic Diseases, Glasgow Royal Infirmary, Glasgow, UK
23 Sheffield Teaching Hospitals, Sheffield, UK
24 Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds and NIHR Leeds Musculoskeletal Biomedical Research Unit, Leeds Teaching Hospitals NHS Trust, UK
25 Stanford University, Stanford, California, USA.
26 University of Belgrade School of Medicine, Institute of Rheumatology, Belgrade, Serbia
27 Clinical and Academic Rheumatology, North Bristol NHS Trust, Bristol, UK
28 Nottingham University Hospitals NHS Trust, and Nottingham NHS Treatment Centre, Nottingham, UK
29 Peter Maddison Rheumatology Centre, Llandrindod, UK
30 Queens Hospital, Romford, UK
31 University of Copenhagen, Copenhagen Lupus and Vasculitis Clinic, Center for Rheumatology and Spine Diseases, Rigshospitalet, Copenhagen, Denmark
32 Norwich Medical School, University of East Anglia, Norwich, UK
33 Royal National Orthopaedic Hospital, Barnet, London, UK
34 Department of Rheumatology and Clinical Immunology, Justus-Liebig University Giessen, Bad Nauheim, Germany
35 Department of Rheumatology, University of Lübeck, Lübeck, Germany
36 Department of Rheumatology and Clinical Immunology, University Hospital Charité Berlin, Berlin, Germany
37 Department of Rheumatology, Royal Perth Hospital, Perth, Australia
38 Servicio de Reumatología, Hospital Universitario 12 de Octubre, Madrid, Spain
39 Internal Medicine Unit, Limoges University Hospital, France
40 Département de Médecine Interne et Immuinologie Clinique, Centre National de Référence Maladies Systémiques etAuto-immunes Rares, University de Lille, Inserm, U995, FHI Immune-Mediated Inflammatory Diseases and Targeted Therapies, Lille, France
41 Gateshead Hospitals Foundation Trust, Gateshead, UK
42 Department of Internal Medicine, Division of Rheumatology, Istanbul University, Istanbul, Turkey
43 File Rheumatic Diseases Unit, Whyteman's Brae Hospital, Kirkcaldy, UK
44 Department of Rheumatology and Clinical Immunology, UMC Utrecht, Utrecht, The Netherlands
45 James Cook University Hospital, Middlesbrough, UK
46 Rheumatology Unit, Royal Adelaide Hospital, and Discipline of Medicine, University of Adelaide, Adelaide, South Australia
47 Department of Rheumatology and Inflammation Research, The Sahlgrenska Academy at Gothenburg University, Gothenburg, Sweden
48 Monash Centre for Inflammatory Diseases, Monash University, Clayton, Melbourne, Australia
49 Department of Internal Medicine, Hôpital Edouard Herriot, Lyon, France
50 Department of Internal Medicine, Foundation Hospital Saint Joseph, Marseille, France
51 Cannock Chase Hospital, Cannock, UK
52 St Vincent’s University Hospital, Dublin, Ireland
53 Department of Internal Medicine, Centre Hospitalier Lyon Sud, Pierre Benite, France
54 Internal Medicine, Ambroise Paré Hospital, Boulogne Billancourt, France

Acknowledgements

The authors are grateful to Dr Holly Ennis for study set-up and to her and Dr Grahame Dinsdale for project coordination during the earlier phase of the study. Thanks also to members of the independent oversight board: Stephen Cole, Dinesh Khanna and Frank Wollheim.

Contributors

ALH, ML, RH, LM, AS, EB, LC, JHWD, OD, KE, WJG, RO, MV and CPD were members of the Steering Committee and designed the study. ALH, RH, ML, RH, LM, AS and CPD wrote the draft report, and all authors reviewed the report, provided comments and approved the final report.

Funding

ESOS was funded by a grant from the European League Against Rheumatism (EULAR) Orphan Disease Programme. Additional funding from Scleroderma and Raynaud’s UK allowed a 1-year extension of the study.

Competing interests

ALH has done consultancy work for Actelion, served on a Data Safety Monitoring Board for Apricus, received research funding and speaker’s fees from Actelion, and speaker’s fees from GSK. JHWD has consultancy relationships and/or has received research funding from Actelion, BMS, Celgene, Bayer Pharma, Boehringer Ingelheim, J&J Therapeutics, Sanofi-Aventis, Novartis, UCB, GSK, Array Biopharma, Active Biotech, Galapagos, Inventiva, Medac, Pfizer, Anakam and RuBiGen and is stock owner of 4D Science GmbH. OD has received consultancy fees from 4D Science, Actelion, Active Biotech, Bayer, Biogenidec, BMS, Boehringer Ingelheim, EpiPharm, Ergonex, espeRare Foundation, Genentech/Roche, GSK, Inventiva, Lilly, Medac, Medimmune, Pharmaceuticals, Pfizer, Serodapharm, and Sinoxa and received research grants from Actelion, Bayer, Boehringer Ingelheim, Ergonex, Pfizer and Sanofi, and has a patent mir-29 for the treatment of systemic sclerosis licensed. WG has received teaching fees from Pfizer. FH has received research funding from Actelion. MEA has undertaken advisory board work and received honoraria from Actelion, and received speaker’s fees from Bristol-Myers Squibb. LC has done advisory board work for Gilead and served Data Safety Monitoring Boards for Cytori and Reata. HG has done consultancy work and received honoraria from Actelion. UM-L is funded in part by EUSTAR/EULAR. JMW has received honoraria from Eli Lilly, Pfizer, Roche, MSD and BMS. AR receives funding from AstaRenequa. CPD has done consultancy for GSK, Actelion, Bayer, Invernessa and Merck-Serono, received research grant funding from GSK, Actelion, CSL Behring and Inventiva, received speaker’s fees from Bayer and given trial advice to Merck-Serono.

Patient consent

Obtained.

Ethics approval

The Ethics Committee of each centre approved the study.

Provenance and peer review

Not commissioned; externally peer reviewed.

Data sharing statement

At present, unpublished data from the study are not available for sharing. This position may change in 6–12 months time.

Open Access

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

REFERENCES


This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/
Assessment Questionnaire (HAQ), Systemic Sclerosis HAQ, and Medical Outcomes Study 36-Item Short Form Health Survey. *Arthritis Rheum* 2007;57:94–102.


