

ORIGINAL ARTICLE

## Cyclophosphamide versus Placebo in Scleroderma Lung Disease

Donald P. Tashkin, M.D., Robert Elashoff, Ph.D., Philip J. Clements, M.D., M.P.H., Jonathan Goldin, M.D., Ph.D., Michael D. Roth, M.D., Daniel E. Furst, M.D., Edgar Arriola, Pharm.D., Richard Silver, M.D., Charlie Strange, M.D., Marcy Bolster, M.D., James R. Seibold, M.D., David J. Riley, M.D., Vivien M. Hsu, M.D., John Varga, M.D., Dean E. Schraufnagel, M.D., Arthur Theodore, M.D., Robert Simms, M.D., Robert Wise, M.D., Fredrick Wigley, M.D., Barbara White, M.D., Virginia Steen, M.D., Charles Read, M.D., Maureen Mayes, M.D., Ed Parsley, D.O., Kamal Mubarak, M.D., M. Kari Connolly, M.D., Jeffrey Golden, M.D., Mitchell Olman, M.D., Barri Fessler, M.D., Naomi Rothfield, M.D., and Mark Metersky, M.D., for the Scleroderma Lung Study Research Group\*

### ABSTRACT

#### BACKGROUND

We conducted a double-blind, randomized, placebo-controlled trial to determine the effects of oral cyclophosphamide on lung function and health-related symptoms in patients with evidence of active alveolitis and scleroderma-related interstitial lung disease.

#### METHODS

At 13 clinical centers throughout the United States, we enrolled 158 patients with scleroderma, restrictive lung physiology, dyspnea, and evidence of inflammatory interstitial lung disease on examination of bronchoalveolar-lavage fluid, thoracic high-resolution computed tomography, or both. Patients received oral cyclophosphamide ( $\leq 2$  mg per kilogram of body weight per day) or matching placebo for one year and were followed for an additional year. Pulmonary function was assessed every three months during the first year, and the primary end point was the forced vital capacity (FVC, expressed as a percentage of the predicted value) at 12 months, after adjustment for the baseline FVC.

#### RESULTS

Of 158 patients, 145 completed at least six months of treatment and were included in the analysis. The mean absolute difference in adjusted 12-month FVC percent predicted between the cyclophosphamide and placebo groups was 2.53 percent (95 percent confidence interval, 0.28 to 4.79 percent), favoring cyclophosphamide ( $P < 0.03$ ). There were also treatment-related differences in physiological and symptom outcomes, and the difference in FVC was maintained at 24 months. There was a greater frequency of adverse events in the cyclophosphamide group, but the difference between the two groups in the number of serious adverse events was not significant.

#### CONCLUSIONS

One year of oral cyclophosphamide in patients with symptomatic scleroderma-related interstitial lung disease had a significant but modest beneficial effect on lung function, dyspnea, thickening of the skin, and the health-related quality of life. The effects on lung function were maintained through the 24 months of the study.

From the University of California at Los Angeles, Los Angeles (D.P.T., R.E., P.J.C., J.G., M.D.R., D.E.F., E.A.); Medical University of South Carolina, Charleston (R.S., C.S., M.B.); University of Medicine and Dentistry of New Jersey—Robert Wood Johnson Medical School, New Brunswick (J.R.S., D.J.R., V.M.H.); University of Illinois, Chicago (J.V., D.E.S.); Boston University, Boston (A.T., R.S.); Johns Hopkins School of Medicine, Baltimore (R.W., F.W., B.W.); Georgetown University, Washington, D.C. (V.S., C.R.); University of Texas at Houston Medical School, Houston (M. Mayes, E.P.); Wayne State University, Detroit (K.M.); University of California San Francisco, San Francisco (M.K.C., J.G.); University of Alabama, Birmingham (M.O., B.F.); and University of Connecticut Health Center, Farmington (N.R., M. Metersky). Address reprint requests to Dr. Tashkin at the Division of Pulmonary and Critical Care Medicine, David Geffen School of Medicine, 10833 Le Conte Ave., Los Angeles, CA 90095-1690, or at dtashkin@mednet.ucla.edu.

\*Members of the Scleroderma Lung Study Research Group are listed in the Appendix.

N Engl J Med 2006;354:2655-66.

Copyright © 2006 Massachusetts Medical Society.

**S**CLERODERMA (ALSO CALLED SYSTEMIC sclerosis) is an autoimmune connective-tissue disorder that is characterized by microvascular injury, excessive fibrosis of the skin, and distinctive visceral changes that can involve the lungs, heart, kidneys, and gastrointestinal tract. Forty percent of patients with scleroderma have ventilatory restriction, mainly as a result of interstitial lung disease. Together with pulmonary hypertension, ventilatory restriction has emerged as the leading cause of death related to scleroderma.<sup>1,2</sup> The mortality rate among patients with severe ventilatory restriction (as reflected by a forced vital capacity [FVC] that is less than 50 percent of the predicted value) from scleroderma-related interstitial lung disease is approximately 42 percent within 10 years after the onset of the disease (as determined by the patient's recall of the first symptom of scleroderma other than Raynaud's phenomenon),<sup>1</sup> an estimate that underscores the need for treatment of this disorder.

A number of agents have been evaluated as treatments for scleroderma-related interstitial lung disease, but none have been proven effective.<sup>3</sup> Only cyclophosphamide has shown promise in several retrospective studies, which suggested that it may slow the decrease in or even improve the FVC over time.<sup>2-11</sup> Although these results have encouraged many physicians to adopt cyclophosphamide as a standard of care for patients with this disease, the lack of a randomized, prospective study makes it difficult to draw solid conclusions regarding the efficacy, toxicity, and risk-benefit ratio of this drug.

The Scleroderma Lung Study, a 13-center, double-blind, randomized, placebo-controlled trial sponsored by the National Institutes of Health, was designed to evaluate the effectiveness and safety of oral cyclophosphamide for one year in patients with active, symptomatic scleroderma-related interstitial lung disease who were followed for a total of two years. Results from the study indicate that cyclophosphamide has disease-modifying effects on pulmonary function and such measures of response as dyspnea, the health-related quality of life, functional ability, and skin thickness.

## METHODS

### SUBJECTS

Between September 2000 and January 2004, we enrolled patients with limited or diffuse systemic

scleroderma<sup>12,13</sup> if they had evidence of active alveolitis on examination of bronchoalveolar-lavage (BAL) fluid (defined as neutrophilia of  $\geq 3$  percent, eosinophilia of  $\geq 2$  percent, or both)<sup>2,14</sup> or on thoracic high-resolution computed tomography (CT), any ground-glass opacity, the onset of the first symptom of scleroderma other than Raynaud's phenomenon within the previous seven years, an FVC between 45 and 85 percent of the predicted value,<sup>15</sup> and grade 2 exertional dyspnea according to the baseline instrument of the Mahler Dyspnea Index (as measured with the use of the magnitude-of-task component).<sup>16</sup> Primary exclusion criteria included a single-breath carbon monoxide diffusing capacity (DLCO) that was less than 30 percent of the predicted value,<sup>17</sup> a history of smoking within the preceding six months, other clinically significant pulmonary abnormalities, or clinically significant pulmonary hypertension requiring drug therapy. Patients taking prednisone at a dose of more than 10 mg per day, those who had previously been treated for more than four weeks with oral cyclophosphamide or had received two or more intravenous doses, and those who had recently received other potentially disease-modifying medications were also excluded. The inclusion and exclusion criteria are described in detail in Supplementary Appendix 1 (available with the full text of this article at [www.nejm.org](http://www.nejm.org)). All patients provided written informed consent, and the study was approved by the medical institutional review board at each clinical center.

### SCREENING AND RANDOMIZATION

Patients were screened at each center as outlined in Supplementary Appendix 2, and their eligibility was determined centrally at the data coordinating center at the University of California at Los Angeles. Patients who met all the inclusion criteria were randomly assigned with the use of a permuted-block design and a 1:1 allocation (in blocks of four to six patients per center) to receive either daily oral cyclophosphamide or matching placebo for 12 months, followed by another year of follow-up without study medication. Treatment of any type was allowed during the second year at the discretion of the primary treating physician.

### BASELINE MEASUREMENTS

Baseline measurements, including tests of pulmonary function, scores for skin thickness, measures of the quality of life and functional activity, and

indexes of cough and breathlessness, are presented in detail in Supplementary Appendix 3.<sup>18-28</sup> High-resolution CT scans were scored for the extent of pure ground-glass opacity, lung fibrosis, honeycombing, and emphysema, as described in Supplementary Appendix 4.

#### TREATMENT

Cyclophosphamide (Cytoxan, Bristol-Myers Squibb) and placebo were formulated into matching gelcaps at a dose of 25 mg, and treatment was initiated with a dose of 1 mg per kilogram of body weight per day (to the nearest 25 mg). The doses were increased monthly by one capsule up to 2 mg per kilogram. Assessments to monitor toxic effects included a complete blood count and urinalysis every two to four weeks and a chemistry panel every three months. To preserve blinding of investigators, an independent medication-control officer assessed adverse events and regulated all doses of study medication, in accordance with the study protocol. The study drug was discontinued temporarily or permanently if there was evidence of drug-related toxic effects. Reasons for withholding or changing the dose of the study drug and predefined rules for the reintroduction of the drug or adjustment of the dose are given in Supplementary Appendix 5. A data safety and monitoring board provided oversight of the study. An independent morbidity and mortality review committee reviewed all serious adverse events and rendered opinions regarding their relation to the study.

#### SERIAL MONITORING AND OUTCOME MEASURES

Measurements were made at baseline and at three-month intervals throughout the study, except those made with the baseline instrument of the Mahler Dyspnea Index, which was replaced by the Mahler transitional dyspnea index (with a scale from -9 to +9, with the plus sign indicating improvement and the minus sign indicating worsening) at subsequent visits.<sup>16</sup> Lung-volume measurements were repeated at six-month intervals.

#### TREATMENT FAILURE

Treatment failure was defined as an absolute decrease from baseline in the FVC of at least 15 percent of the predicted value occurring at least three months after treatment was initiated and lasting for at least one month. Patients whose response met this definition were withdrawn from the double-blind phase of the study but were encour-

aged to continue scheduled visits that included spirometry, the administration of questionnaires, and other study procedures. These patients were informed of their treatment assignment and offered open-label therapy with cyclophosphamide.

#### STATISTICAL ANALYSIS

We estimated that we would need to enroll 163 patients, given an expected mean ( $\pm$ SD) annual decline in the FVC of  $9\pm 16$  percent of the predicted value<sup>2,4-6</sup> and a dropout rate of 15 percent, to achieve a two-sided alpha level of 0.05 and a statistical power of 90 percent. A prespecified analysis of covariance was used to assess treatment effects, adjusted for baseline values and treatment group, with the FVC (expressed as a percentage of the predicted value) at 12 months identified as the primary outcome. A Huber covariance estimation and testing procedure was used, because the data were not normally distributed and included influential extreme values.<sup>29</sup> Patients for whom outcome measures at six months or more were not available or whose primary outcome data were technically not interpretable were excluded from analysis. For all remaining patients who withdrew prematurely, a generalized-estimating-equation regression model was fitted, and data missing at 12 months were imputed. At 24 months after randomization, the FVC was the prespecified primary end point; the analysis was based on a longitudinal model that included terms for treatment and time and an interaction between treatment and time with the baseline FVC included in the model.

Other planned analyses of the primary outcome included a nonparametric analysis with a modified Wilcoxon score test,<sup>30</sup> in which the FVC at 12 months and the time to death were considered a combined end point, and an extension of the covariance analysis in which the baseline score for fibrosis (the mean reading by two independent radiologists of the worst degree of fibrosis in any region of the baseline high-resolution CT, as described in Supplementary Appendix 4) was used as an additional covariate.

Prespecified secondary outcomes included values at month 12, adjusted for baseline values, for total lung capacity (expressed as a percentage of the predicted value), DLCO, the diffusing capacity adjusted for alveolar volume (DL:VA), the disability index of the Health Assessment Questionnaire (HAQ), and the Medical Outcomes Study 36-item

Short-Form General Health Survey (SF-36). No adjustment was made to the P values for these multiple variables. A post hoc analysis of scores for skin thickness was performed in a similar fashion. The three subscales and the summed focal score for the transitional dyspnea index (a secondary outcome) were analyzed with the use of a mixed-model method.<sup>16</sup>

The data were collected, held, and analyzed in a secure manner at the data coordinating center, with frequent review by the data and safety monitoring board. The statistical analyses were performed with SAS software (version 9.1, SAS), where appropriate. Bristol-Myers Squibb, the manufacturer of the oral cyclophosphamide used in the study, donated the drug but had no role in the design or conduct of the study or the analysis of the data.

## RESULTS

### ENROLLMENT AND BASELINE CHARACTERISTICS

The disposition of the 267 patients who underwent screening and of the 158 eligible patients who underwent randomization is shown in Figure 1. Baseline characteristics of those who underwent randomization are listed in Table 1 and summarized in Supplementary Appendix 6. There were no significant differences between the two groups except that the scores for the HAQ disability index were significantly lower (indicating greater health) in the placebo group than in the cyclophosphamide group.

The frequency, type, and severity of baseline findings on high-resolution CT are described in detail in Supplementary Appendix 7. Ground-glass opacity, a criterion for entry into the study, was present in 90.1 percent of the patients. Fibrosis was the most common finding on high-resolution CT (89.5 percent of the patients), with 35.8 percent of the patients having evidence of honeycombing. Pulmonary-artery enlargement was rare (1.9 percent of the patients). All abnormalities were most commonly found in the lower lung zones. There were no significant differences between the two groups in the extent of parenchymal abnormalities.

The frequency of positive findings on baseline examination of BAL fluid and the mean percentages of neutrophils and eosinophils are shown in Supplementary Appendix 7. Of 144 patients whose BAL-fluid results could be evaluated cytologically,

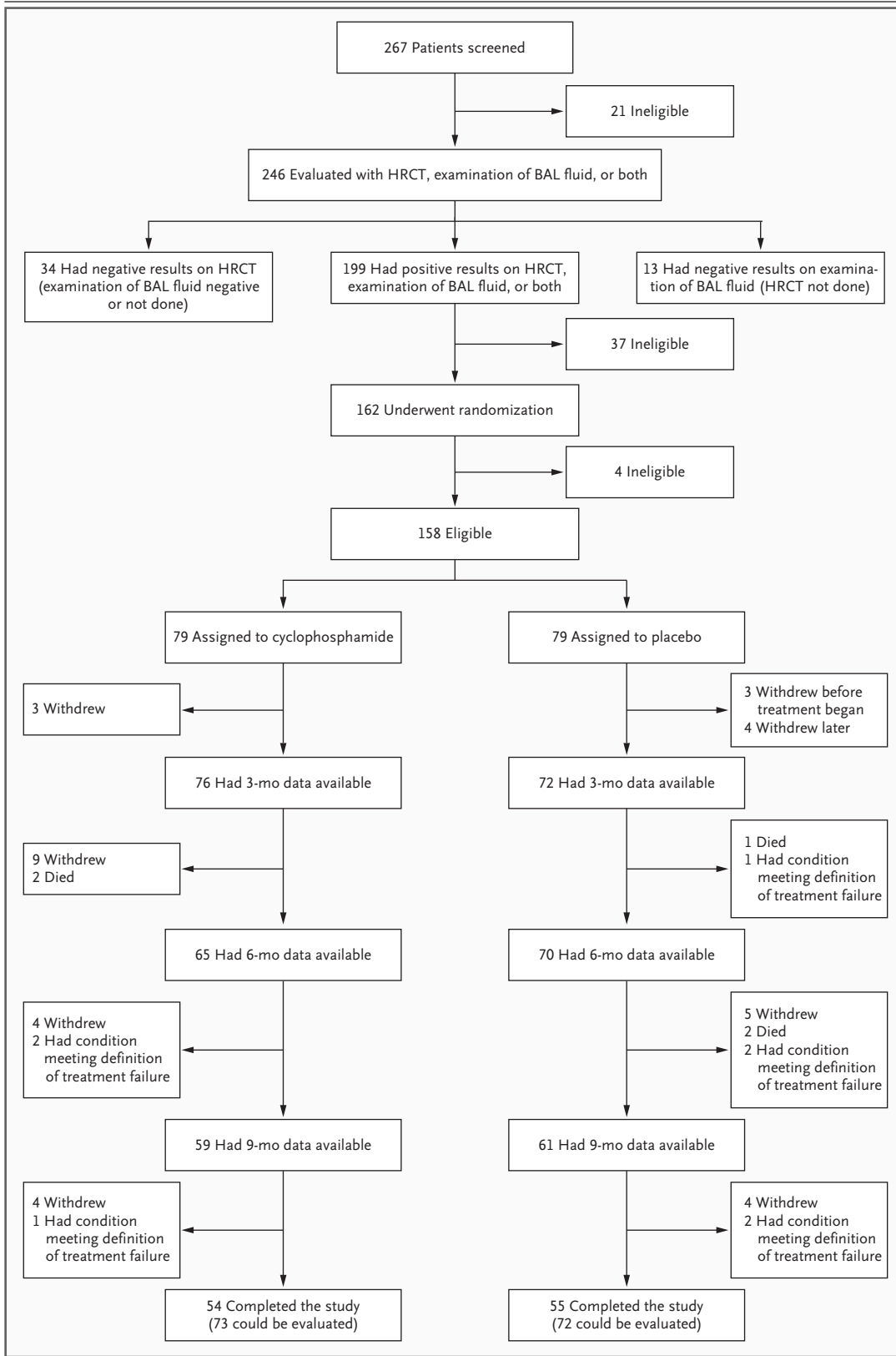
### Figure 1 (facing page). Screening, Enrollment, and Outcomes.

A positive result on high-resolution CT (HRCT) was defined by the presence of any ground-glass opacity, and a positive result on examination of BAL fluid was defined by the presence of neutrophilia of  $\geq 3$  percent, eosinophilia of  $\geq 2$  percent, or both on a cell differential count (Panel A). The results of spirometry for 1 patient in the placebo group were technically unsatisfactory and did not permit imputation of the FVC at 12 months; thus, 72 patients in the placebo group could be evaluated for the primary outcome measure.

70.8 percent had evidence of alveolitis according to the study criteria. The frequency and severity of findings on examination of BAL fluid were similar in the two groups. The proportions of eligible patients whose condition fulfilled the criteria for evidence of alveolitis on high-resolution CT, examination of BAL fluid, or both are given in Supplementary Appendix 8.

### RESPONSES TO TREATMENT

Of a total of 158 patients, 3 assigned to placebo and 1 assigned to cyclophosphamide withdrew before starting the study treatment and were not included in the analysis. A number of patients who received at least one dose of study medication withdrew from the study, had conditions that met the protocol definition of treatment failure, or died (Fig. 1). A total of 20 patients in the cyclophosphamide group and 13 in the placebo group withdrew within 12 months after randomization, most of them because of adverse events or serious adverse events. Three patients in the cyclophosphamide group had treatment failure, as compared with five in the placebo group. Five patients died during the first 12 months, two in the cyclophosphamide group and three in the placebo group. A total of 54 patients (68.4 percent) in the cyclophosphamide group and 55 (69.6 percent) in the placebo group completed all 12 months of treatment. However, a number of those who withdrew, had treatment failure, or died had completed at least the six-month visit, allowing their end-point results to be imputed and included in the analysis. One patient in the placebo group was not included because the results of spirometry were technically unsatisfactory. Thus, data on 145 patients (91.8 percent), including 73 in the cyclophosphamide group and 72 in the placebo group, were evaluated for the primary outcome. Baseline values and outcomes at month 12 for these 145 pa-



**Table 1. Baseline Characteristics of 158 Patients.\***

Characteristic	All Patients (N=158)	Cyclophosphamide Group (N=79)	Placebo Group (N=79)
Age (yr)			
Mean	47.9±1.0	48.2±1.4	47.5±1.4
Range	19.6–83.1	28.9–81.5	19.6–83.1
Female sex (% of patients)	70.3	75.6	64.6
Duration of scleroderma (yr)			
Mean	3.2±0.2	3.2±0.3	3.1±0.2
Range	0.04–12.0	0.04–12.0	0.2–6.8
Diffuse scleroderma-related disease (% of patients)	59.5	62.8	57.7
FVC (% of predicted)	68.1±1.0	67.6±1.5	68.6±1.5
FEV <sub>1</sub> :FVC (% of predicted)	82.8±0.6	82.8±1.0	82.8±0.8
Total lung capacity (% of predicted)	69.6±1.1	69.6±1.5	69.4±1.5
Functional residual capacity (% of predicted)	73.6±1.5	75.2±2.0	71.7±2.2
Residual volume (% of predicted)	70.7±2.1	72.3±3.1	69.0±2.9
D <sub>L</sub> CO (% of predicted)	47.2±1.1	47.0±1.6	47.4±1.6
MIP (cm of water)	88.7±2.7	84.1±3.5	92.5±4.1
MEP (cm of water)	91.8±3.2	87.9±4.5	95.2±4.6
Focal score for the Mahler Dyspnea Index	5.7±0.2	5.7±0.2	5.7±0.2
Visual-analogue score for breathing	28.4±2.1	27.2±2.8	29.5±3.1
Cough (% of patients)	69.5	71.1	68.0
SF-36 score			
Physical component	33.5±0.9	32.6±1.3	34.3±1.2
Mental component	49.8±0.9	48.6±1.2	50.7±1.2
Skin-thickening score			
All patients			
Mean	14.7±0.9	15.7±1.3	14.0±1.2
Range	0–51	2–51	0–40
Patients with diffuse scleroderma			
Mean	21.0±1.0	21.6±1.5	20.5±1.4
Range	6–51	7–51	6–40
Patients with limited scleroderma			
Mean	5.7±0.4	5.8±0.7	5.6±0.6
Range	2–16	2–16	2–14
Score for HAQ disability index	0.83±0.06	0.96±0.08	0.70±0.08†

\* Plus–minus values are means ±SE. FEV<sub>1</sub> denotes forced expiratory volume in one second, MIP maximum inspiratory pressure, and MEP maximum expiratory pressure. Scores for the Medical Outcomes Study 36-item Short-Form General Health Survey (SF-36) can range from 0 to 100, with lower scores indicating worse health status. Scores for the Mahler Dyspnea Index can range from 0 to 12, with lower scores indicating worse dyspnea. Scores for the visual-analogue scale for breathing can range from 1 to 100, with higher numbers indicating increasing difficulty breathing. Scores for skin thickening can range from 0 to 51, with higher scores indicating more severe thickening. Scores for the HAQ disability index can range from 1 to 3, with higher numbers indicating greater disability.

† P=0.02 for the comparison with the cyclophosphamide group.

tients, as well as mean ( $\pm$ SE) changes in these values, adjusted with the use of Huber estimation,<sup>29</sup> are shown in Table 2.

#### PRESPECIFIED PRIMARY OUTCOME

When the baseline FVC was used as the only covariate in the Huber model (primary analysis), the

adjusted mean absolute difference in FVC at 12 months between the cyclophosphamide group and the placebo group was 2.53 percent (95 percent confidence interval, 0.28 to 4.79 percent), favoring cyclophosphamide ( $P<0.03$ ). When both the baseline FVC and the worst score for fibrosis at baseline on high-resolution CT were included in the

**Table 2. Change in Values from Baseline to Month 12.\***

Characteristic	Baseline Value	Value at 12 Mo	Difference
<b>Cyclophosphamide group</b>			
FVC (% of predicted)	67.6 $\pm$ 1.3	66.6 $\pm$ 1.7	-1.0 $\pm$ 0.92 $\dagger$
Total lung capacity (% of predicted)	70.4 $\pm$ 2.1	70.5 $\pm$ 1.8	-0.3 $\pm$ 1.82 $\dagger$
D <sub>L</sub> CO (% of predicted)	47.2 $\pm$ 1.6	42.8 $\pm$ 1.7	-4.2 $\pm$ 1.16
Score on Mahler Dyspnea Index			
According to baseline instrument	5.6 $\pm$ 0.22		
According to transitional dyspnea index (focal score) $\dagger$		1.4 $\pm$ 0.23	
Cough (%)	72.5	66.2	
Score for HAQ disability index	0.94 $\pm$ 0.077	0.84 $\pm$ 0.08	-0.11 $\pm$ 0.05 $\dagger$
SF-36 score			
Physical component	33.0 $\pm$ 1.3	33.8 $\pm$ 1.3	0.7 $\pm$ 1.0
Mental component	48.7 $\pm$ 1.2	51.8 $\pm$ 1.3	2.9 $\pm$ 1.5
Skin-thickness score			
Diffuse	21.7 $\pm$ 10.1	15.9 $\pm$ 11.0	-5.3 $\pm$ 7.4
Limited	6.1 $\pm$ 3.6	5.0 $\pm$ 4.3	-0.8 $\pm$ 2.4
<b>Placebo group</b>			
FVC (% of predicted)	68.3 $\pm$ 1.5	65.6 $\pm$ 1.6	-2.6 $\pm$ 0.9
Total lung capacity (% of predicted)	67.9 $\pm$ 1.9	64.7 $\pm$ 1.9	-2.8 $\pm$ 1.2
D <sub>L</sub> CO (% of predicted)	47.9 $\pm$ 1.7	44.3 $\pm$ 2.1	-3.5 $\pm$ 1.0
Score on Mahler Dyspnea Index			
According to baseline instrument	5.6 $\pm$ 0.42		
According to transitional dyspnea index (focal score) $\ddagger$		-1.5 $\pm$ 0.43	
Cough (%)	55.9	67.2	
Score for HAQ disability index	0.70 $\pm$ 0.09	0.86 $\pm$ 0.10	0.16 $\pm$ 0.06
SF-36 score			
Physical component	35.1 $\pm$ 1.4	33.2 $\pm$ 1.4	-1.9 $\pm$ 1.2
Mental component	50.8 $\pm$ 1.4	50.9 $\pm$ 1.5	0.1 $\pm$ 1.5
Skin-thickness score			
Diffuse	20.2 $\pm$ 9.3	19.1 $\pm$ 11.2	-1.7 $\pm$ 6.9
Limited	5.5 $\pm$ 3.4	5.7 $\pm$ 4.2	0.2 $\pm$ 3.3

\* The analysis included 73 patients in the cyclophosphamide group and 72 patients in the placebo group. Plus-minus values are means  $\pm$ SE. Scores for the Medical Outcomes Study 36-item Short-Form General Health Survey (SF-36) can range from 0 to 100, with lower scores indicating worse health status. Scores for the Mahler Dyspnea Index can range from 0 to 12, with lower scores indicating worse dyspnea. Scores for the HAQ disability index can range from 1 to 3, with higher scores indicating greater disability. Scores for skin thickening can range from 0 to 51, with higher scores indicating more severe thickening.

$\dagger$   $P<0.05$ , after adjustment for baseline values, favoring cyclophosphamide.

$\ddagger$   $P<0.05$  favoring cyclophosphamide.

model (in the secondary analysis of the primary outcomes), the adjusted mean absolute difference in the FVC at 12 months between the two groups increased to 2.97 percent (95 percent confidence interval, 0.75 to 5.19 percent), favoring cyclophosphamide ( $P=0.009$ ). We used regression analysis to evaluate the effect of the baseline score for fibrosis on the FVC at 12 months (Fig. 2A). In the placebo group, the regression slope was significantly negative ( $-2.01$  percent of the predicted FVC per unit score for fibrosis,  $P=0.006$ ), indicating a greater decline in FVC over time among those who had more severe fibrosis at baseline. By comparison, the difference from 0 in the regression slope in the cyclophosphamide group (0.96 percent of the predicted FVC per unit score for fibrosis) was not significant ( $P=0.26$ ). The difference between the regression slopes in the two groups was significant ( $P<0.009$ ), suggesting that cyclophosphamide protected against the decrease in FVC in patients with fibrosis. The combined end point of the time to death plus the FVC at 12 months among the 145 patients evaluated also favored cyclophosphamide ( $P=0.04$  with the use of a modified Wilcoxon procedure).<sup>30</sup>

#### PRESPECIFIED SECONDARY OUTCOMES

A significant difference was also found between the cyclophosphamide group and the placebo group in total lung capacity at 12 months (absolute difference in the percent of the predicted value, 4.09 percent favoring cyclophosphamide; 95 percent confidence interval, 0.49 to 7.65 percent;  $P=0.026$ ). There were no significant differences between the two groups in either the DLCO (absolute difference in the percent of the predicted value,  $-1.04$ ;  $P=0.43$ ) at 12 months or in the DL:VA at 12 months (absolute difference in the percent of the predicted value, 0.0002 percent;  $P=0.92$ ). Breathlessness is the chief symptom related to scleroderma-related interstitial lung disease. The mean focal score according to the transitional dyspnea index showed a clinically meaningful improvement in dyspnea (by  $>1$  unit) in the cyclophosphamide group ( $+1.4\pm 0.23$ ), in contrast to a clinically meaningful worsening (by  $>1$  unit) in the placebo group ( $-1.5\pm 0.43$ ) (Table 2). There were significant differences between the two groups in the focal score ( $P<0.001$ ) and in the scores for the three-component transitional dyspnea index, which were incorporated into a mixed model ( $P<0.001$ ) (available in Supplementary Ap-

pendix 9). After adjustment for baseline HAQ scores, the adjusted mean HAQ disability scores at 12 months were significantly lower (indicating less disability) in the cyclophosphamide group than in the placebo group, with the difference between the two groups favoring cyclophosphamide ( $-0.16$ ; 95 percent confidence interval,  $-0.28$  to  $-0.04$ ;  $P=0.009$ ).

Among the scores for all the domains and summary components of the SF-36, only those for the vitality domain (7.99; 95 percent confidence interval, 2.18 to 13.8;  $P=0.007$ ) and the health-transition domain ( $-0.66$ ; 95 percent confidence interval,  $-1.02$  to  $-0.30$ ;  $P=0.003$ ) showed significant between-group differences, both favoring cyclophosphamide. In none of the domains did the scores show a significant improvement with placebo.

#### POST HOC ANALYSES

When pure ground-glass opacity, the presence or degree of neutrophilia, eosinophilia, or both on examination of BAL fluid, or the degree of honeycombing on the baseline high-resolution CT was added as a covariate to the Huber model, it did not significantly affect the FVC at 12 months. Figure 2B shows the distribution of changes in the FVC (in increments or decrements of 5 percent) in the cyclophosphamide group and the placebo group. A significantly greater percentage of patients in the cyclophosphamide group than in the placebo group had any improvement from baseline values in the FVC ( $P<0.01$  by Fisher's exact test). In 85 patients with diffuse disease, scores for skin thickness showed a significant difference between the two groups favoring cyclophosphamide ( $-3.06$ ; 95 percent confidence interval,  $-3.54$  to  $-0.52$ ;  $P=0.008$ ).

#### PRIMARY OUTCOME AT 24 MONTHS

Measurements of the FVC were available at 24 months for 113 patients and at 15, 18, and 21 months in an additional 3, 5, and 1 patients, respectively (a total of 122 patients were studied beyond 12 months, 62 in the cyclophosphamide group and 60 in the placebo group). In a longitudinal model made with the use of data collected from months 6 to 24, a significant treatment effect was observed, with a mean absolute difference between the two groups of 1.95 percent in the FVC (95 percent confidence interval, 1.2 to 2.6 percent;  $P<0.01$ ), favoring cyclophosphamide. No interaction between time and treatment was found.

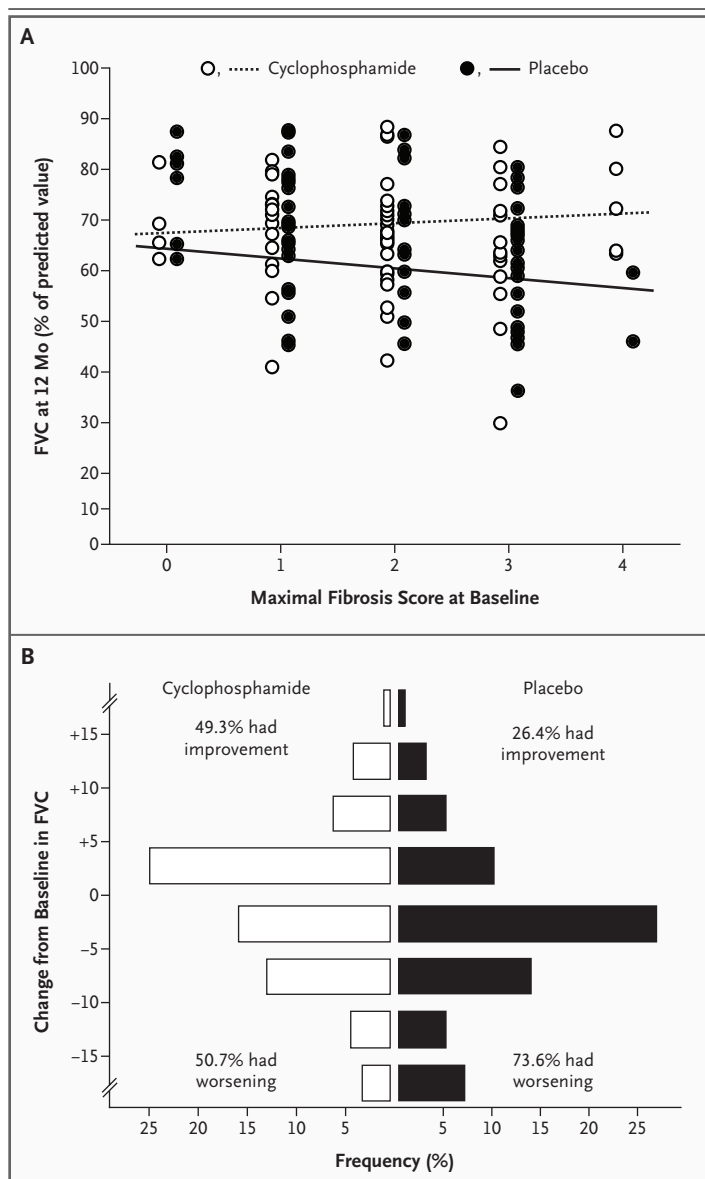


**Figure 2. The Percentage of Predicted FVC at 12 Months and Changes in the Percentage of Predicted FVC from Baseline to 12 Months.**

Panel A shows a scattergram of the FVC at 12 months (adjusted for the baseline FVC) according to treatment, as compared with the maximal fibrosis score as determined on baseline thoracic high-resolution CT in 69 patients in the cyclophosphamide group and 71 in the placebo group whose scans were scored by two independent radiologists who were unaware of the treatment assignments. The slope of the regression is significant in the placebo group ( $-2.01$  percent of the predicted FVC per unit score for fibrosis,  $P=0.006$ ) but not in the cyclophosphamide group ( $0.96$  percent of the predicted FVC per unit score for fibrosis,  $P=0.26$ ); the difference in the slopes between the two groups was significant ( $P=0.009$ ). In Panel B, the change in the FVC from baseline to month 12 in each of 145 patients who could be evaluated is displayed as a histogram according to treatment and absolute increments or decrements of 5 percent of the predicted value. All changes of 15 percent or more are grouped, and imputed values were used for patients who completed at least the 6-month evaluation but did not return for the 12-month visit. A significantly greater percentage of patients in the cyclophosphamide group than in the placebo group (49.3 percent vs. 26.4 percent) had any improvement in the FVC ( $P<0.01$ ); conversely, a significantly greater percentage of those in the placebo group than in the cyclophosphamide group (73.6 percent vs. 50.7 percent) had any worsening of the FVC ( $P<0.01$  by Fisher's exact test).

**ADVERSE EVENTS**

Details of adverse events that occurred during the one-year double-blind treatment period (year 1) and the subsequent 12 months of follow-up (year 2) are presented in Table 3. Hematuria, leukopenia, neutropenia, anemia, and pneumonia were more common among patients in the cyclophosphamide group than among those in the placebo group during year 1; the difference between the two groups was significant for leukopenia and neutropenia ( $P<0.05$  by Fisher's exact test). Few adverse events were recorded during year 2. The number of patients with serious adverse events and the total number of these events during year 1 were slightly, but not significantly, higher in the cyclophosphamide group than in the placebo group (17 vs. 11 patients had adverse events; 20 vs. 16 had serious adverse events). A similar trend was observed in year 2. During months 13 to 24 after randomization, four patients in the cyclophosphamide group and three in the placebo group died. In total, 12 patients died during the two years



of the study, but none of the deaths were considered to be related to the study treatment.

**DISCUSSION**

This double-blind, randomized, placebo-controlled clinical trial examined the efficacy of oral cyclophosphamide for the treatment of scleroderma-related interstitial lung disease. Our findings indicate a significant, albeit quite modest, treatment effect of cyclophosphamide on changes in the FVC and total lung capacity over a period of one year but no significant effect on measures of gas

**Table 3. Adverse Events and Serious Adverse Events.\***

Event	Cyclophosphamide Group		Placebo Group	
	Year 1	Year 2	Year 1	Year 2
	<i>number of patients</i>			
<b>Adverse event</b>				
Hematuria	9	1	3	2
Leukopenia†	19	0	0	0
Neutropenia†	7	0	0	0
Anemia	2	2	0	1
Pneumonia	5	1	1	0
<b>Serious adverse event‡</b>				
Probably related to treatment	2	4	0	0
Possibly related to treatment	3	4	2	5
Not related to treatment	15	19	14	17
Total	20	27	16	22
Death	2	4	3	3

\* During year 1 of the trial, patients received oral cyclophosphamide ( $\leq 2$  mg per kilogram per day) or matching placebo. Year 2 (month 13 to 24) was a follow-up period.

†  $P < 0.05$  by Fisher's exact test.

‡ According to the criteria of the Food and Drug Administration, a serious adverse event is an adverse drug experience occurring at any dose that results in death; life-threatening illness; persistent or clinically significant disability, incapacity, or both; hospitalization or prolongation of hospitalization; a congenital abnormality or birth defect; or cancer. Whether a serious adverse event was treatment-related was adjudicated by consensus of the executive committee of the study in a blinded fashion. During year 1, two serious adverse events occurred in the cyclophosphamide group (leukopenia with pneumonia and leukopenia with gastroenteritis) that were deemed to be probably related to treatment. Other possible treatment-related serious adverse events in this group included pneumonia and laryngospasm complicating cystoscopy. Serious adverse events that occurred in the placebo group included a questionable case of pneumonia and a documented case of pneumonia. During year 2, there were four serious adverse events in the cyclophosphamide group that were deemed to be probably treatment-related, including four separate episodes of hematuria in one patient leading to a urinary-diversion procedure. Serious adverse events in this group that were possibly treatment-related included carcinoma in situ of the bladder, squamous-cell carcinoma of the vulva, angiosarcoma of the scalp, and pneumonia (in one patient). Serious adverse events in the placebo group that were possibly treatment-related included sepsis with worsening pulmonary hypertension requiring hospitalization on four separate occasions in one patient and pneumonia in another patient.

transfer (DLCO or DLCO:VA). The clinical importance of this small treatment effect on lung physiology is supported by the additional findings that cyclophosphamide improved dyspnea (according to scores for the transitional dyspnea index), skin thickening, and other outcomes, including functional ability (according to the scores for the HAQ disability index) and some health-related measures of the quality of life (vitality and health

transition). For example, the transitional dyspnea index focal score improved by  $1.4 \pm 0.23$  in the cyclophosphamide group and worsened by  $1.5 \pm 0.43$  in the placebo group, a difference that is both statistically significant ( $P < 0.001$  with the use of a covariance model) and more than twice the minimal clinically significant level of 1.0.<sup>31</sup> Analysis of the primary end point at 24 months showed that the treatment effect on the FVC persisted. Although the limitations of this study included a high dropout rate and the potential for inadvertent unblinding, precautions were taken to minimize the effect of these limitations on the outcome of the trial (Supplementary Appendix 10).

An interesting finding was the relationship between fibrosis and the response to cyclophosphamide. Patients in the placebo group who had evidence of more severe fibrosis on baseline high-resolution CT had the greatest declines in FVC during the one-year treatment period; however, the baseline degree of fibrosis had no significant influence on the change in FVC in cyclophosphamide-treated patients in the same period. A significant interaction between fibrosis and treatment was found ( $P = 0.009$ ) (Fig. 2A). A possible conclusion from this result is that preexisting fibrosis in our cohort identifies a subgroup of patients with relatively early active systemic sclerosis alveolitis who were at greatest risk for progressive interstitial lung disease unless treated. The potential benefits of treatment with cyclophosphamide need to be weighed against its adverse effects. Although the average number of adverse events per patient during the one-year treatment period was only slightly higher in the cyclophosphamide group ( $4.0 \pm 0.49$ ) than in the placebo group ( $3.0 \pm 0.41$ ), this difference was significant on Poisson analysis ( $P = 0.01$ ). The only adverse events that occurred significantly more often among patients treated with cyclophosphamide were those known to be associated with this drug — namely, leukopenia and neutropenia — which often responded to appropriate dose adjustment. Although five serious adverse events that occurred in the cyclophosphamide group and one that occurred in the placebo group during the one-year treatment period were judged to be treatment-related, the overall number of serious adverse events and deaths was not significantly different between groups during the full two-year study period. Two days after the completion of the 12-month treatment period, however, severe hem-

orrhagic cystitis developed in one patient in the cyclophosphamide group and has required multiple hospitalizations and a urinary diversion procedure, underscoring the potential for cyclophosphamide to cause serious toxic effects. The potential long-term effects of cyclophosphamide, including its association with bladder cancer<sup>32</sup> and other malignant diseases,<sup>33,34</sup> were not evaluated in this study.

In summary, we found that one year of treatment with oral cyclophosphamide resulted in a small but significant improvement in the FVC and total lung capacity at 12 months but not in measures of gas transfer. The extent of fibrosis on baseline high-resolution CT, but not other high-resolution CT findings or cellularity on examination of BAL fluid, significantly influenced the primary treatment effect. The beneficial effect of cyclophosphamide on pulmonary function was paralleled by a significant improvement in dyspnea, functional ability, the health-related quality of life, and skin thickness. The favorable effect of cyclophosphamide on FVC was sustained for the two years of the study. Although treatment with

cyclophosphamide resulted in a greater number of adverse events than did placebo, the risk-benefit ratio appears to be favorable. Caution regarding the use of cyclophosphamide is still warranted, since potential long-term consequences were not evaluated.

Supported by a grant from the Public Health Service (UO1 HL 60587, to the University of California at Los Angeles Clinical Center) and by grants from the National Heart, Lung, and Blood Institute (UO1 HL 60587 and UO1 HL 60606, to the data coordinating center at the University of California at Los Angeles; UO1 HL 50550, to the University of Medicine and Dentistry of New Jersey; UO1 HL 60895, to the University of Illinois; UO1 HL 60682, to Boston University; UO1 HL 60750, to the Medical University of South Carolina; UO1 HL 60597, to Johns Hopkins School of Medicine; UO1 HL 60794, to Georgetown University; UO1 HL 60839, to the University of Texas at Houston; UO1 HL 60587, to the University of California at San Francisco; UO1 HL 60748, to the University of Alabama at Birmingham; UO1 HL 60587, to the University of Connecticut Health Center, Farmington; UO1 HL 60838, to Wayne State University; and UO1 HL 60823, to the Virginia Mason Research Center); by the National Institute of Arthritis and Musculoskeletal and Skin Diseases; and by a grant from the National Center for Research Resources, National Institutes of Health (5MO1 RR 00865, for the analysis of study data).

No potential conflict of interest relevant to this article was reported.

We are indebted to Bristol-Myers Squibb for supplying cyclophosphamide.

#### APPENDIX

The following persons and institutions participated in the Scleroderma Lung Study: *University of California at Los Angeles (UCLA), Los Angeles:* P.J. Clements, D.P. Tashkin, R. Elashoff, J. Goldin, M. Roth, D. Furst, K. Bulpitt, D. Khanna, W.-L.J. Chung, S. Viasco, M. Sterz, L. Woolcock, X. Yan, J. Ho, S. Vasunilashorn, I. da Costa; *University of Medicine and Dentistry of New Jersey, New Brunswick:* J.R. Seibold, D.J. Riley, J.K. Amorosa, V.M. Hsu, D.A. McCloskey, J.E. Wilson; *University of Illinois Chicago, Chicago:* J. Varga, D. Schraufnagel, A. Wilbur, D. Lapota, S. Arami, P. Cole-Saffold; *Boston University, Boston:* R. Simms, A. Theodore, P. Clarke, J. Korn, K. Tobin, M. Nuite; *Medical University of South Carolina, Charleston:* R. Silver, M. Bolster, C. Strange, S. Schabel, E. Smith, J. Arnold, K. Caldwell, M. Bonner; *Johns Hopkins School of Medicine, Baltimore:* R. Wise, F. Wigley, B. White, L. Hummers, M. Bohlman, A. Polito, G. Leatherman, E. Forbes, M. Daniel; *Georgetown University, Washington, D.C.:* V. Steen, C. Read, C. Cooper, S. Wheaton, A. Carey, A. Ortiz; *University of Texas at Houston, Houston:* M. Mayes, E. Parsley, S. Oldham, T. Filemon, S. Jordan, M. Perry; *University of California at San Francisco, San Francisco:* K. Connolly, J. Golden, P. Wolters, R. Webb, J. Davis, C. Antolos, C. Maynetto; *University of Alabama at Birmingham, Birmingham:* B. Fessler, M. Olman, C. Sanders, L. Heck, T. Parkhill; *University of Connecticut Health Center, Farmington:* N. Rothfield, M. Metersky, R. Cobb, M. Aberles, F. Ingenito, E. Breen; *Wayne State University, Detroit:* M. Mayes, K. Mubarak, J.L. Granda, J. Silva, Z. Injic, R. Alexander; *Virginia Mason Research Center, Seattle:* D. Furst, S. Springmeyer, S. Kirkland, J. Molitor, R. Hinke, A. Mondt; *Data Safety and Monitoring Board:* Harvard Medical School, Boston — T. Thompson; Veterans Affairs Medical Center, Brown University, Providence, R.I. — S. Rounds; Cedars Sinai—UCLA, Los Angeles — M. Weinstein; Clinical Trials Surveys, Baltimore — B. Thompson; *Mortality and Morbidity Review Committee:* UCLA, Los Angeles — H. Paulus, S. Levy; Johns Hopkins University, Baltimore — D. Martin.

#### REFERENCES

1. Steen VD, Conte C, Owens GR, Medsger TA Jr. Severe restrictive lung disease in systemic sclerosis. *Arthritis Rheum* 1994; 37:1283-9.
2. White B, Moore WC, Wigley FM, Xiao HQ, Wise RA. Cyclophosphamide is associated with pulmonary function and survival benefit in patients with scleroderma and alveolitis. *Ann Intern Med* 2000;132: 947-54.
3. Latsi PI, Wells AU. Evaluation and management of alveolitis and interstitial lung disease in scleroderma. *Curr Opin Rheumatol* 2003;15:748-55.
4. Silver RM, Warrick JH, Kinsella MB, Staudt LS, Baumann MH, Strange C. Cyclophosphamide and low-dose prednisone therapy in patients with systemic sclerosis (scleroderma) with interstitial lung disease. *J Rheumatol* 1993;20:834-44.
5. Akesson A, Scheja A, Lundin A, Wollheim F. Improved pulmonary function in systemic sclerosis after treatment with cyclophosphamide. *Arthritis Rheum* 1994;37: 729-35.
6. Steen VD, Lanz JK Jr, Conte C, Owens GR, Medsger TA Jr. Therapy for severe interstitial lung disease in systemic sclerosis: a retrospective study. *Arthritis Rheum* 1994;37:1290-6.
7. Davas EM, Peppas C, Maragou M, Alvanou E, Hondros D, Dantis PC. Intravenous cyclophosphamide pulse therapy for the treatment of lung disease associated with scleroderma. *Clin Rheumatol* 1999;18: 455-61.
8. Behr J, Vogelmeier C, Beinert T, et al. Bronchoalveolar lavage for evaluation and management of scleroderma disease of the lung. *Am J Respir Crit Care Med* 1996; 154:400-6.
9. Varai G, Earle L, Jimenez SA, Steinert

- RM, Varga J. A pilot study of intermittent intravenous cyclophosphamide for the treatment of systemic associated lung disease. *J Rheumatol* 1998;25:1325-9.
10. Pakas I, Ioannidis JP, Malagari K, Skopouli FN, Moutsopoulos HM, Vlachoyiannopoulos PG. Cyclophosphamide with low or high dose prednisolone for systemic sclerosis lung disease. *J Rheumatol* 2002;29:298-304.
11. Giacomelli R, Valentin G, Slasano F, et al. Cyclophosphamide pulse regimen in the treatment of alveolitis in systemic sclerosis. *J Rheumatol* 2002;29:731-6.
12. Preliminary criteria for the classification of systemic sclerosis (scleroderma). *Arthritis Rheum* 1980;23:581-90.
13. Medsger TA Jr. Classification, prognosis. In: Clements PJ, Furst DE, eds. Systemic sclerosis. 2nd ed. Philadelphia: Lippincott Williams & Wilkins, 2004:17-28.
14. BAL Cooperative Group Steering Committee. Bronchoalveolar lavage constituents in healthy individuals, idiopathic pulmonary fibrosis, and selected comparison groups. *Am Rev Respir Dis* 1990;141:S169-S202.
15. Crapo RO, Morris AH, Gardner RM. Reference spirometric values using techniques and equipment that meet ATS recommendations. *Am Rev Respir Dis* 1981;123:659-64.
16. Mahler DA, Weinberg DH, Wells CK, Feinstein AR. The measurement of dyspnea: contents, interobserver agreement, and physiologic correlates of two new clinical indexes. *Chest* 1984;85:751-8.
17. Crapo RO, Morris AH. Standardized single breath normal values of carbon monoxide diffusing capacity. *Am Rev Respir Dis* 1981;123:185-9.
18. American Thoracic Society. Standardization of spirometry, 1994 update. *Am J Respir Crit Care Med* 1995;152:1107-36.
19. American Association for Respiratory Care. AARC clinical practice guideline: body plethysmography. *Respir Care* 1994;39:1184-90.
20. Guidelines for the measurement of respiratory function: recommendations of the British Thoracic Society and the Association of Respiratory Technicians and Physiologists. *Respir Med* 1994;88:165-94.
21. American Thoracic Society. Single-breath carbon monoxide diffusing capacity (transfer factor): recommendations for a standard technique — 1995 update. *Am J Respir Crit Care Med* 1995;152:2185-98.
22. Black LF, Hyatt RE. Maximal respiratory pressures: normal values and relationship to age and sex. *Am Rev Respir Dis* 1969;99:696-702.
23. Clements PJ, Lachenbruch PA, Seibold JR, et al. Inter and intraobserver variability of total skin thickness score (modified Rodnan TSS) in systemic sclerosis. *J Rheumatol* 1995;22:1281-5.
24. Petty TL. The National Mucolytic Study: results of a randomized, double-blind, placebo-controlled study of iodinated glycerol in chronic obstructive bronchitis. *Chest* 1990;98:75-83.
25. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473-83.
26. Steen VD, Medsger TA Jr. The value of the Health Assessment Questionnaire and special patient-generated scales to demonstrate change in systemic sclerosis patients over time. *Arthritis Rheum* 1997;40:1984-91.
27. Clements PJ, Wong WK, Hurwitz EL, et al. The disability index of the Health Assessment Questionnaire is a predictor and correlate of outcome in the high-dose versus low-dose penicillamine in systemic sclerosis trial. *Arthritis Rheum* 2001;44:653-61.
28. Ferris BG. Epidemiology Standardization Project (American Thoracic Society). *Am Rev Respir Dis* 1978;118:1-120.
29. Huber PJ. Robust regression: asymptotics, conjectures, and Monte Carlo. *Ann Statist* 1973;1:799-821.
30. Finkelstein DM, Schoenfeld DA. Combining mortality and longitudinal measures in clinical trials. *Stat Med* 1999;18:1341-54.
31. Witek TJ Jr, Mahler DA. Minimal important difference of the transition dyspnoea index in multinational clinical trial. *Eur Respir J* 2003;21:267-72.
32. Talar-Williams C, Hijazi YM, Walther MM, et al. Cyclophosphamide-induced cystitis and bladder cancer in patients with Wegener granulomatosis. *Ann Intern Med* 1996;124:477-84.
33. Baltus JA, Boersma JW, Hartman AP, Vandenbroucke JP. The occurrence of malignancies in patients with rheumatoid arthritis treated with cyclophosphamide: a controlled retrospective follow-up. *Ann Rheum Dis* 1983;42:368-73.
34. Radis CD, Kahl LE, Baker GL, et al. Effects of cyclophosphamide on the development of malignancy and on long-term survival of patients with rheumatoid arthritis: a 20-year followup study. *Arthritis Rheum* 1995;38:1120-7.

Copyright © 2006 Massachusetts Medical Society.

#### CLINICAL TRIAL REGISTRATION

The *Journal* encourages investigators to register their clinical trials in a public trials registry. The members of the International Committee of Medical Journal Editors plan to consider clinical trials for publication only if they have been registered (see *N Engl J Med* 2004;351:1250-1). The National Library of Medicine's [www.clinicaltrials.gov](http://www.clinicaltrials.gov) is a free registry, open to all investigators, that meets the committee's requirements.