EXTENDED REPORT

A prediction model for progressive disease in systemic sclerosis

Jessica Meijs,[1] Anne A Schouffoer,[1,2] Nina Ajmone Marsan,[3] Theo Stijnen,[4] Hein Putter,[4] Maarten K Ninaber,[5] Tom W J Huizinga,[1] Jeska K de Vries-Bouwstra[1]

ABSTRACT

Objective: To develop a model that assesses the risk for progressive disease in patients with systemic sclerosis (SSc) over the short term, in order to guide clinical management.

Methods: Baseline characteristics and 1 year follow-up results of 163 patients with SSc referred to a multidisciplinary healthcare programme were evaluated. Progressive disease was defined as: death, $\geq$10% decrease in forced vital capacity, $\geq$15% decrease in diffusing capacity for carbon monoxide, $\geq$10% decrease in body weight, $\geq$30% decrease in estimated-glomerular filtration rate, $\geq$30% increase in modified Rodnan Skin Score (with $\Delta\geq$5) or $\geq$0.25 increase in Scleroderma Health Assessment Questionnaire. The number of patients with progressive disease was determined. Univariable and multivariable logistic regression analyses were used to assess the probability of progressive disease for each individual patient. Performance of the prediction model was evaluated using a calibration plot and area under the receiver operating characteristic curve.

Results: 63 patients had progressive disease, including 8 patients who died $\leq$18 months after first evaluation. Multivariable analysis showed that friction rubs, proximal muscular weakness and decreased maximum oxygen uptake as % predicted, adjusted for age, gender and use of immunosuppressive therapy at baseline, were significantly associated with progressive disease. Using the prediction model, the predicted chance for progressive disease increased from a pretest chance of 37% to 67–89%.

Conclusions: Using the prediction model, the chance for progressive disease for individual patients could be doubled. Friction rubs, proximal muscular weakness and maximum oxygen uptake as % predicted were identified as relevant parameters.

Key messages

What is already known about this subject?
▶ Few studies have described algorithms on an individualised basis to predict mortality after 2 to 15 years of follow-up in systemic sclerosis (SSc).

What does this study add?
▶ A prediction model assessing the chance for progressive disease for individual patients at short term was currently developed.

How might this impact on clinical practice?
▶ Using the prediction model, the predicted chance for progressive disease could be doubled from a pretest chance of 37% to 67-89%.
▶ Maximum oxygen uptake as measured by CPET is identified as biomarker for progressive SSc.

INTRODUCTION

Individualised management and treatment is one of the most important challenges in medicine. Systemic sclerosis (SSc) is a rare multisystem disease which is highly heterogeneous in presentation and disease course. Recent evidence suggests that earlier initiation of adequate treatment based on regular screening for organ involvement contributes to improved survival.[1] The availability of new treatment options, such as autologous haematopoietic stem cell transplantation (HSCT), offers the chance for prolonged event-free survival.[2] For optimal efficacy of this treatment, careful timing in the disease course is of pivotal importance. Given the associated treatment-related mortality during the first year, this treatment option underlines the need to identify patients with a high risk of severe organ involvement in the short term.

Numerous attempts have been made to identify predictors for severe organ involvement and mortality in SSc.[3–11] Only a few studies have described algorithms to predict outcome on a more individualised basis.[12–16] The aforementioned studies defined outcome of interest after 2–15 years of follow-up. In contrast, a recently published study containing observational data from the EUSTAR database described a model identifying, among patients with diffuse cutaneous SSc, those with a 44% chance of skin fibrosis progression during the first year, as compared

to 9.7% in the whole cohort.[17] Ideally, in order to guide individualised management of patients with SSc, a model combining outcome parameters for several organ systems and mortality predicting disease course in the short term should be available. Whether it is possible to reliably identify patients at risk using such a model, given the heterogeneous nature of SSc, remains to be determined.

The present study aimed to develop a model that predicts progressive disease in the short term, defined by either deterioration of organ functions, or mortality, in patients with SSc. The derived prediction model is evaluated for discriminative performance, and a cut-off value is determined in order to evaluate utility in clinical practice.

PATIENTS AND METHODS
Study design
This study is performed using data from a prospective cohort study in patients with SSc who participated in an annual 2-day multidisciplinary healthcare programme aiming to structure screening for organ involvement and to provide multidisciplinary care for patients with SSc. Ethical approval was obtained from the Institutional Review Board of the Leiden University Medical Centre (LUMC). All participants gave written informed consent.

Patients
Data from all patients referred to the multidisciplinary healthcare programme between April 2009 and January 2014 were collected. Patients were included if they had a diagnosis of SSc according to the American Rheumatism Association,[18] the LeRoy criteria[19] or the ACR/EULAR 2013 classification criteria.[20]

On the basis of the degree of skin involvement, three subtypes of patients were classified:
1. Diffuse cutaneous SSc (DcSSc) with skin involvement proximal to the elbows and knees.
2. Limited cutaneous SSc (LcSSc) with skin involvement distal to the elbows and knees.
3. Limited non-cutaneous SSc (LSSc) without skin involvement.

Skin scores were all performed by experienced rheumatologists (AS, JVB, AAS). Patients were classified on the basis of their maximum skin score ever. For example, if a patient had had a skin score of 30 and underwent HSCT after which the skin score decreased to 6, the patient was still classified as DcSSc.

For the current analysis, selected patients had to have participated in the care programme at least twice, with the second visit 1 year after the baseline visit (range 10–23 months).

Multidisciplinary healthcare programme
All patients participated in the healthcare programme that combines annual extensive organ screening with multidisciplinary team care. Cardiopulmonary investigations included: high-resolution CT (HRCT) of the thorax, pulmonary function tests (including analyses of forced vital capacity (FVC) and diffusing capacity for carbon monoxide (DLCO)), cardiopulmonary exercise test (CPET; including analyses of maximum heart rate, maximum wattage, maximum oxygen consumption (VO2) and maximum ventilation), echocardiography and 24 h Holter electrocardiography (ECG). Furthermore, laboratory investigations including measurement of autoantibodies and nailfold videocapillaroscopy were performed.

In addition, the patients completed the Scleroderma Health Assessment Questionnaire (SHAQ) for assessment of physical functioning and the Short Form-36 (SF-36) for quality of life.

Diagnosis of interstitial lung disease (ILD) was determined on the basis of the presence of a non-specific interstitial pneumonia pattern or usual interstitial pneumonia pattern on the HRCT-thorax,[21] as reported by the radiologist.

The systolic pulmonary artery pressure (SPAP) was estimated using echocardiography by an experienced cardiologist and elevated pulmonary pressure is defined using a cut-off value of 35 mm Hg.[22] Left ventricular end-diastolic and end-systolic volumes were measured using the biplane modified Simpson's rule. Left ventricular ejection fraction was calculated as left ventricular end-diastolic—LV end-systolic volume/left ventricular end-diastolic volume. Furthermore, the presence of pericardial effusion was noted. The presence of arrhythmias was defined as the presence of multiform ventricular extrasystole >100 per day, couplets or runs of ventricular tachycardia or supraventricular tachycardia of at least 30 s on 24 h Holter ECG monitoring.[23] Conduction abnormalities were defined as a complete left bundle branch block or right bundle branch block, atrioventricular block (first, second or third degree) or pacemaker rhythm for sinus node dysfunction.[23]

Change of treatment
Initiation of new immunosuppressive treatment is mainly considered in case of extensive and/or progressive skin involvement, relevant decline in VC and/or DLCO (without using an absolute threshold) in combination with the presence of non-specific interstitial pneumonia or usual interstitial pneumonia on HRCT. Autologous HSCT is applied according to the inclusion criteria and treatment regimen as described in the ASTIS trial (24). Azathioprine (AZA) is prescribed in case of primary biliary cirrhosis and hydroxychloroquine (HCQ) in case of SSc overlap syndrome with rheumatoid arthritis (RA). Rituximab is given as part of a randomised placebo-controlled clinical trial (RITIS), registered at https://www.clinicaltrialsregister.eu/ EudraCT Number: 2008-007180-16.

Progressive disease
Since we aimed to define risk for progressive disease in general, in order to guide clinical management, several

variables were chosen, each reflecting a different organ system. Cut-offs for these variables were based on reported values for minimal important difference (MID). Selected variables were: (1) death before the second visit; (2) decrease of ≥10% in FVC (percentage of predicted);[16] (3) decrease of ≥15% in DLCO (percentage of predicted);[16] (4) decrease of ≥10% in body weight;[24] (5) decrease of ≥30% in estimated-glomerular filtration rate (eGFR);[25] (6) increase of ≥30% in modified Rodnan Skin Score (mRSS) with a minimum of Δ5;[26 27] or 7 ≥0.25 increase in SHAQ.[27] Overall progressive disease was defined as the occurrence of at least one of the above prespecified outcomes during 1 year of follow-up.

Statistical analysis

Associations between baseline variables and the presence of progressive disease were evaluated and expressed as ORs with the 95% CIs and p values.

Missing values of variables used to define overall progressive disease and baseline predictors were replaced by multiple imputation using multiple regression modelling by the multiple imputations by chained equations procedure as implemented in SPSS.[28] Missing CPET were considered missing not at random, as an inability to perform CPET most likely reflects severe impaired cardiopulmonary performance status. Therefore, VO2 max was not imputed and the missing-indicator method was used.[28] When CPET was missing, an indicator variable, with value 1 if the CPET was missing and 0 if the result was present, was created.

Univariable logistic regression analysis was used to determine the independent association between baseline characteristics and overall progressive disease after 1 year of follow-up.

Possible correlations between all variables which were significantly contributing in the univariable logistic regression analyses were checked for multicollinearity using a variance inflation factor (VIF) of 10.[29] In case of multicollinearity between variables (VIF >10), the most significant variable was selected for further analysis.

For the multivariable model, all predictor variables with a p value smaller than 0.05 in the univariable analysis, indicating an important association with progressive disease, were selected using a predictor selection approach (forward selection). Univariable and multivariable logistic regression analyses were always adjusted for previous and current immunosuppressive therapy (including cyclophosphamide, methotrexate and HSCT) at baseline. Multivariable logistic regression analysis was adjusted for age and gender.

The predicted probability of progressive disease was calculated for every patient.

The predicted probabilities were compared with the observed percentage of patients with progressive disease. The positive predictive value (PPV) and negative predictive value (NPV) were determined for several cut-off values of the predicted probability.

The predictive performance of the model was assessed by examining measures of calibration and discrimination. Calibration refers to how close predicted progressive disease agrees with observed progressive disease and was assessed with a calibration plot.[30] Since progressive disease is a binary outcome, a loess algorithm was used as a smoothing technique to estimate the observed probability.[31]

The discrimination of the prediction model was assessed by receiver operating characteristic (ROC) curve analysis.

For internal validation, a bootstrap procedure was performed for control for overfitting.[32]

All statistical analyses were executed using SPSS V.20.0 software (SPSS Inc, Chicago, USA), except that bootstrap validation was performed by using R V.3.1.1.

RESULTS
Patient population

By January 2014, 163 patients with SSc had had a second evaluation after a mean period of 13.5 months (SD 2.5). Eight patients died before the second visit could have been performed. Baseline characteristics of the 171 included patients are presented in table 1. The patients were mostly women (80%), Caucasian (70%) and, on average, 53 years (SD 14). Patients had a median disease duration of 2 years. The disease subset at baseline was classified as DcSSc in 61 patients, LcSSc in 75 patients and LSSc in 28 patients.

At baseline, 63 (39%) patients were treated with immunosuppressive medication, including 55 (32%) patients who were previously treated with one or more immunosuppressive medications including autologous HSCT (n=13), cyclophosphamide (n=18), corticosteroids (n=28), methotrexate (MTX; n=28) and AZA (n=2), HCQ (n=2) and 60 (35%) patients currently being treated with immunosuppressive medication, including mycophenolate mofetil (MMF; n=6), corticosteroids (n=24), MTX (n=22), AZA (n=5) and HCQ (n=7). In total, 65 (38%) patients were previously, or are currently, treated at baseline evaluation with cyclophosphamide, MTX or HSCT.

Change of treatment

On the basis of the findings during the multidisciplinary healthcare programme, new immunosuppressive treatment (one or more medications) was started at baseline in 37 patients (22%). Newly prescribed treatment included autologous HSCT (n=2), cyclophosphamide (n=10), MMF (n=5), corticosteroids (n=4), MTX (n=7), AZA (n=1), HCQ (n=2) and rituximab/placebo (n=8). In none of the patients with previous HSCT was new immunosuppressive medication started.

Mortality

Within 1 year after the first visit, eight patients (mean age 62.6 years) died, including four patients with DcSSc,

Table 1 Baseline characteristics of the systemic sclerosis population with a baseline visit and 1 year follow-up*

	Patients N=171	DcSSc N=64	LcSSc N=79	LSSc N=28
Sociodemographics				
Age, years, mean (SD)	53.2 (14.3)	50 (14)	56.2 (13.6)	51.8 (15.8)
Female, N (%)	135 (78.9)	44 (68.8)	65 (82.3)	26 (92.9)
Caucasian origin, N (%)†	118 (69)	44 (68.8)	54 (68.4)	20 (71.4)
Disease characteristics, N (%)				
Disease duration, years, median (IQR)	2 (0–10)	3 (1–8)	4 (0.8–12)	0 (0–1)
DU	43 (25.1)	14 (21.9)	24 (30.4)	5 (17.9)
Telangiectasia	100 (58.5)	32 (50)	55 (69.6)	13 (46.4)
Synovitis	18 (10.5)	6 (9.4)	9 (11.4)	3 (10.7)
Friction rubs	6 (3.5)	5 (7.8)	1 (1.3)	0 (0)
Calcinosis	26 (15.2)	6 (9.4)	16 (20.3)	4 (14.3)
Proximal muscular weakness	9 (5.3)	8 (12.5)	1 (1.3)	0 (0)
Renal crisis	6 (3.5)	3 (4.7)	3 (3.8)	0 (0)
MRSS, median (IQR)	2.5 (0–6)	6 (2–19)	2.5 (2–4)	0 (0–0)
Autoantibodies, N (%)				
ANA‡	155 (90.6)	56 (87.5)	75 (94.9)	24 (85.7)
Anti-Scl-70†	39 (22.8)	28 (43.8)	10 (12.7)	1 (3.6)
Anticentromere‡	64 (37.4)	4 (6.3)	40 (50.6)	20 (71.4)
Cardiopulmonary investigations				
FVC % of predicted, mean (SD)§	99.1 (22.7)	88.4 (20.7)	103.4 (21.1)	110.8 (22.4)
DLCO % of predicted, mean (SD)§	64 (17.6)	59.7 (18.6)	64.3 (16)	72.9 (17)
ILD, N (%)	82 (48)	40 (62.5)	36 (45.6)	6 (21.4)
SPAP≥35 mm Hg, N (%)	28 (16.4)	12 (18.8)	14 (17.7)	2 (7.1)
LVEF %, mean (SD)	60.3 (7.7)	60.8 (9)	59.9 (6.7)	60.5 (56–65.8)
Pericardial fluid, N (%)	5 (2.9)	3 (4.7)	1 (1.3)	1 (3.6)
SSc pattern on capillaroscopy¶	52 (94.5)	9 (90)	28 (96.6)	15 (88.2)
Immunosuppressive therapy, N (%)				
Current	55 (32.2)	22 (34.4)	24 (30.4)	9 (32.1)
Previous	60 (35.1)	38 (59.4)	19 (24.1)	3 (10.7)
SHAQ (0–3), median (IQR)‡	0.63 (0.13–1.00)	0.63 (0.22–1.13)	0.5 (0.13–0.88)	0.63 (0.13–0.88)
SF-36, median (IQR)†				
Physical component summary scale	39.7 (31.5–48.7)	41.4 (28.4–49.6)	39.6 (32.5–48)	39.1 (31.8–49.5)
Mental component summary scale	52.5 (41–56.6)	53.2 (46–58.4)	51.3 (38.8–55)	51.5 (37.7–57)

*Including 8 patients who died within 1 year of follow-up.
†<15% missing.
‡<10% missing.
§<5% missing.
¶Capillaroscopy is performed in 55 patients (32%) at baseline.
ANA, antinuclear antibody; Anti-Scl-70, anti-topoisomerase; DcSSc, diffuse cutaneous SSc; DLCO, diffusing capacity for carbon monoxide; DU, digital ulcers; FVC, forced vital capacity; ILD, interstitial lung disease; LcSSC, limited cutaneous SSc; LSSC, limited non-cutaneous SSc; LVEF, left ventricle ejection fraction; MRSS, modified Rodnan Skin Score; SF-36, Short-Form-36; SHAQ, Scleroderma Health Assessment Questionnaire; SPAP, systolic pulmonary arterial pressure.

three patients with LcSS and one patient with LSSc. Three patients died due to ILD, one due to PAH, one due to adenocarcinoma of the lung (in a non-smoker), one due to cardiac failure and one due to cytomegalovirus pneumonitis after allogeneic SCT. In one patient (died at age 87 years after suffering from renal disease), the exact cause of death could not be determined. All patients were classified as those with 'progressive disease', since for none of the patients could an association between SSc and death be ruled out with absolute certainty.

Progressive disease

Sixty-three patients showed overall progressive disease at follow-up evaluation according to the predefined criteria, including eight patients who died (table 2). Overall progressive disease was found in 25 (39%) patients with DcSSc, in 30 (38%) with LcSSc and in eight (29%) with LSSc. One patient with LcSSC and one patient with LSSc progressed to DcSSc within 1 year of follow-up. Three patients with LSSc evolved to LcSSc based on development of sclerodactyly. Progressive disease in patients with LSSc was primarily based on a decrease of pulmonary function; FVC decreased in four patients and DLCO in one patient.

The organ systems and number of prespecified outcomes contributing to overall progressive disease are demonstrated in online supplementary figure 1. The majority of the patients (71%) had overall progressive disease based on one event, while in 13% of the patients

Table 2 Incidence of progressive disease in SSc according to SSc subtype Progressive disease, N (%)

	Total N=171	DcSSc N=64	LcSSc N=79	LSSc N=28
Deceased	8 (5)	3 (5)	4 (5)	1 (4)
≥10% decrease in body weight	10 (6)	5 (8)	2 (3)	3 (11)
≥30% increase with minimum of Δ5 in MRSS*	15 (9)	7 (11)	7 (9)	1 (4)
≥30% decrease in eGFR	6 (4)	3 (5)	3 (4)	0 (0)
≥10% decrease in FVC*	13 (8)	3 (5)	6 (8)	4 (14)
≥15% decrease in DLCO*	12 (7)	5 (8)	6 (8)	1 (4)
≥0.25 increase in SHAQ†	11 (6)	3 (5)	8 (10)	0 (0)
Total patients with progressive disease	63 (37)	25 (39)	30 (38)	8 (29)

*<5% missing.
†<15% missing.
DcSSc, diffuse cutaneous systemic sclerosis; DLCO, diffusing capacity for carbon monoxide; eGFR, estimated glomerular filtration rate; FVC, forced vital capacity; LcSSc, limited cutaneous systemic sclerosis; MRSS, modified Rodnan Skin Score; SHAQ, Scleroderma Health Assessment Questionnaire.

two events and in 3% of the patients three events contributed to overall progressive disease.

Missing values
Age, gender, disease subset, SSc-related autoantibodies, friction rubs, proximal muscular weakness, eGFR, ESR and body weight were available for all patients. The following baseline variables were missing and imputed: mRSS (N=3), FVC (N=3) and DLCO (N=4), urine protein (N=7), SHAQ (N=4) and physical component summary score (according to SF-36 (PCSS); N=12).

CPET was not performed in 11 patients due to an inability to cycle based on bad physical performance (N=4) and musculoskeletal disability (N=2).

At follow-up, the following outcome parameters were missing and imputed: mRSS (N=2), FVC (N=3), DLCO (N=5) and SHAQ (N=22).

Prediction of progressive disease
Univariable analyses
Table 3 shows results of the univariable logistic regression analysis, adjusted for previous and current immunosuppressive therapy. After adjusting for immunosuppressive therapy, friction rubs, proximal muscular weakness, pulmonary crackles, mRSS, DLCO, VO2 max, SHAQ and PCSS according to SF-36 were significantly associated with progressive disease after 1 year of follow-up. FVC and gender were borderline significant (p value <0.10).

No multicollinearity was found between friction rubs, proximal muscular weakness, pulmonary crackles, mRSS, DLCO, VO2 max, SHAQ and PCSS.

Prediction in subpopulations with DcSSc and LcSSc
No significant predictors for progressive disease in patients with DcSSc were identified. In patients with LcSSc, VO2 max % predicted and PCSS of SF-36 were identified as significant predictors for progressive disease. However, the multivariable logistic regression analysis for patients with LcSSc did not identify significant predictors for progressive disease (see online supplementary material file).

Multivariable analyses and derivation of the prediction model
In the multivariable logistic regression analysis, independent predictive variables for progressive disease were friction rubs, proximal muscular weakness, VO2 max % predicted and immunosuppressive therapy. OR for progressive disease increased with 12.462 (95% CI 1.253 to 123.905) in the presence of friction rubs and 5.550 (95% CI 1.000 to 30.796) in the presence of proximal muscular weakness, and decreased with 0.979 (95% CI 0.964 to 0.995) per unit increase in VO2 max % predicted. A missing VO2 max was accompanied by its corresponding missing indicator variable. The coefficients for the prediction model are listed in table 4.

The multivariable model remained unchanged when adding the borderline significant variable in the univariable analysis: FVC was excluded from the final model resulting from the forward selection.

Predictive performance of the prediction model
Table 5 shows the predictive performance of several cut-off values for the predicted probability and the number of observed patients with progressive disease. Using cut-off values 0.25 and 0.75, 80% of the patients who had a score of <0.25 did not develop progressive disease (NPV 80%, 95% CI 69% to 91%), and 89% of the patients who had a score of >0.75 did develop progressive disease (PPV 89%, 95% CI 84% to 94%). The calibration plot of the prediction model is shown in the online supplementary file (figure 2). The prediction model showed a reliable calibration, predicting progressive disease in agreement with the observed progressive disease. The calibration plot showed that for predicted probabilities smaller than 0.55, the prediction model is overestimating the observed overall disease progression. For the probabilities higher than 0.60, the model slightly underestimates the chance for progressive disease.

The discriminative ability of the model was evaluated with an ROC curve (see online supplementary file, figure 3), showing an AUC of 0.707 (95% CI 0.682 to 0.732). The optimum cut-off value of the prediction model, as verified by the ROC curve, showed a sensitivity of 60% and a specificity of 85%, corresponding to a

Table 3 Baseline characteristics of patients with progressive and stable disease

	Progressive disease N=63	Stable disease N=108	Adjusted* Univariable OR (95% CI)	p Value
Sociodemographics				
Age, years, mean (SD)	54.6 (15.5)	52.3 (13.7)	1.010 (0.988 to 1.032)	0.395
Female, N (%)	46 (73)	89 (82.4)	1.928 (0.896 to 4145)	0.093
Body weight, kg, mean (SD)	68.1 (14.7)	69.8 (12.9)	0.994 (0.970 to 1.019)	0.633
Disease characteristics, N (%)				
Friction rubs	5 (7.9)	1 (0.9)	11.229 (1.247 to 101.139)	0.031
Proximal muscular weakness	7 (11.1)	2 (1.9)	7.012 (1.393 to 35.313)	0.018
Pulmonary crackles	21 (33.3)	22 (20.4)	2.181 (1.057 to 4.501)	0.035
MRSS, median (IQR)	4 (2–10)	2 (0–6)	1.014 (1.002 to 1.082)	0.039
Laboratory investigations				
eGFR, mL/min, median (IQR)	89 (68.7–133.3)	87.3 (70.6–104.7)	1.006 (0.996 to 1.015)	0.264
Cardiopulmonary investigations				
FVC % of predicted, mean (SD)	95.5 (24.7)	101.1 (21.4)	0.985 (0.971 to 1.000)	0.053
DLCO % of predicted, mean (SD)	59.4 (19.8)	66.6 (15.7)	0.974 (0.955 to 0.993)	0.008
Maximum VO2% of predicted, mean (SD)	83 (26.4)	94.2 (24.3)	0.984 (0.974 to 0.994)	0.002
SF-36, median (IQR)				
Physical Component Summary Scale	37.9 (27.4–46.8)	42.6 (32.5–49.6)	0.961 (0.931 to 0.991)	0.012
SHAQ (0–3), median (IQR)†	0.75 (0.25–1.13)	0.5 (0.13–0.88)	1.788 (1.010 to 3.095)	0.044

*Adjusted for previous or current treatment on baseline evaluation with cyclophosphamide, methotrexate or autologous haemopoietic stem cell transplantation.
DLCO, diffusing capacity for carbon monoxide; eGFR, estimated-glomerular filtration rate; FVC, forced vital capacity; MRSS, modified Rodnan Skin Score; SF-36, Short Form-36; SHAQ, Scleroderma Health Assessment Questionnaire; VO2, volume oxygen uptake.

cut-off value of 0.38 for the predicted probability, with a PPV of 42% and an NPV of 76%.

INTERNAL VALIDATION
The AUC of the bootstrap predictions equalled the AUC value of the prediction model (0.72, 95% CI 0.64 to 0.81), indicating that overfitting was not a problem.

DISCUSSION
This study is the first attempt to develop a clinical model to assess the chance for overall progressive disease in the short term in patients with SSc in order to guide clinical management. Our study shows that even in a cohort of patients with SSc not selected for disease duration or subtype, overall disease progression is frequently observed. By applying the model, the expected chance for progression could be increased from 37% to 67–89%, depending on the chosen cut-off value, indicating that improved discrimination of patients is a reasonable possibility. The current prediction model is not externally validated and should therefore be validated in other cohorts. However, internal validation showed that overfitting was not a problem, and results seem to be robust.

A broad set of variables was available for evaluation of association with overall disease and its role in pulmonary involvement. Friction rubs, proximal muscular weakness and VO2 max % predicted as determined by CPET were the relevant predicting variables included in the model, after correction for age, gender and immunosuppressive treatment. This suggests that friction rubs, proximal muscular weakness and VO2 max are relatively sensitive variables in measuring overall progressive disease. Recently, a EUSTAR study concerning predictors of progressive disease has been published,[4] identifying joint synovitis and tendon friction rubs as parameters independently associated with disease progression after 2 years of follow-up. In our population, we did not find an association between synovitis and progressive disease; however, friction rubs were significantly associated with progressive disease, confirming the relevance of this finding. As compared to this EUSTAR study, we used a different definition of progressive disease, generally identifying patients at an earlier disease stage.

CPET evaluates patients during exercise, so it is likely to detect cardiopulmonary abnormalities not measurable at rest. Different aspects of pulmonary involvement can be evaluated during CPET (lung parenchymal damage as well as vascular abnormalities). Furthermore, it is highly reproducible, non-invasive and operator-independent. The role of the CPET in organ involvement screening is a relatively new finding. Most studies focused on the role of CPET in identifying PAH[33 34] Very few studies have investigated CPET as a possible biomarker for active/progressive disease in SSc. Cuomo et al[35] suggested including CPET in screening

Table 4 Independent predictive variables for progressive disease based on multivariable logistic regression analysis

Predictors	B	OR	95% CI	p Value
Age, years	0.024	1.024	0.999 to 1.050	0.057
Female	0.376	1.457	0.627 to 3.384	0.381
Previous or current immunosuppressive therapy	−0.864	0.422	0.196 to 0.909	0.027
Friction rubs	2.523	12.462	1.253 to 123.905	0.031
Proximal muscular weakness	1.714	5.550	1.000 to 30.796	0.050
Maximum VO2, % of predicted*	−0.021	0.979	0.964 to 0.995	0.009
Missing indicator variable CPET	−1.067	0.344	0.049 to 2.432	0.285

*β is 0 if CPET is missing.
CPET, cardiopulmonary exercise test; VO2, volume oxygen uptake.

programmes for severity of SSc. They found that an impaired maximum oxygen uptake was present in 93% of the patients and independently associated with the severity of lung involvement. Our study confirmed the importance of CPET in identifying patients at risk for progressive disease in general, including progressive disease based on other parameters than pulmonary involvement or PAH.

Our study has several limitations which should be taken into account.

First, no validated definition of progressive disease is available, and therefore overall progressive disease was defined as a combination of MIDs as used in randomised clinical trials in SSc. Whether this definition is a useful outcome parameter should be evaluated in future studies. We have chosen to define outcome parameters reflecting the different organ systems (skin, lungs, kidneys), as well as parameters reflecting health in general (weight loss, mortality, functional ability). The cut-off values have been based on defined MIDs, as our intention is to select patients with a high risk for significant deterioration for more stringent annual follow-up. Second, the accuracy of our model is moderate. However, it is in line with other prediction models.[12–15 17] We hypothesise that the poor discriminative ability is at least partially caused by our decision to define a broad outcome parameter describing several organ systems in the relatively short term.

Third, while other prediction models specifically have focused on patients with DcSSc,[14 17] we were interested in predicting progressive disease in the whole patient population, as regular follow-up of all patients with SSC is being advocated.[36 37] In line with this, recently developed diagnostic criteria also aim at classifying patients earlier in the disease course,[20 38] in order to identify significant organ involvement at an earlier time point. As a matter of interest, we did evaluate possible predictive factors in subpopulations with DcSSc and LcSSc, but no significantly strong predictors were identified. This can possibly be explained by the low number of patients when only selecting either DcSSc (n=64) or LcSSc (n=79). In addition, since most patients with DcSSc had long-standing disease (mean disease duration of 5.1 ±6 years), discrimination of those patients who still progress is even more difficult, and larger patient groups are needed. Including patients with L(c)SSc and patients with DcSSc with longer disease duration naturally decreased the overall percentage of patients with progressive disease in our cohort as patients presenting with early DcSSc have a different natural history.[39] Within 1 year of follow-up, progression to DcSSc was found in 2% of patients with L(c)SSc. Strikingly, although all subpopulations of SSc, including LSSc, LcSSc and DcSSc, were included, mRSS was not selected for the final prediction model. We believe that this is explained by the fact that part of the patients with DcSSc had been treated successfully before the baseline visit and had stable low skin scores during the time frame under study. However, since our intention was to develop a prediction rule which can be used to guide

Table 5 Predictive performance of several cut-off values for predicted probability of progressive disease

Cut-off values	Observed number of patients*	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
0.25	49	84	36	43	80
0.32	84	69	60	49	76
0.45	125	46	84	46	73
0.50	138	35	90	67	70
0.60	150	27	96	81	69
0.75	162	13	99	89	66

*Below cut-off value. For example, 138 of our patients had a probability of progressive disease below 0.5. Using this cut-off value, 67% of the patients are correctly classified as getting progressive disease, while 70% are correctly classified as not at risk for progressive disease.
No, number; NPV, negative predictive value; PPV, positive predictive value.

clinical practice, for all patients with SSc currently in follow-up irrespective of disease duration or previous treatment, we explicitly chose to also include these patients.

Lastly, 38% of the patients were previously or currently treated with cyclophosphamide, MTX or autologous HSCT at baseline evaluation, which can have influenced our findings. Therefore, all logistic regression analyses were repeated including only patients who had not been treated before. Multivariable logistic regression analyses in these patients identified VO2 max % predicted as the only significant predictor (data not shown). Since our aim was to develop a prediction rule which can be used to guide clinical practice, also in patients who have been treated before, we explicitly chose to develop our model based on a population including untreated patients and treated patients.

This study explored the possibility of prediction of progressive disease in the short term in a heterogeneous population with SSc. The advantages of our study are that the data were prospectively derived from a single centre cohort of patients with SSc. The number of missing values was very low. Since all patients fulfilling SSc classification criteria that visit in the healthcare programme are scheduled for a follow-up visit, independent of disease duration and subtype, the study population reflects the whole population present in a tertiary care centre and therefore the risk of selection bias is low. As compared to other cohorts with SSc, the sociodemographic characteristics, disease severity and functional status of our cohort are comparable to those of other cohorts.[40 41]

In conclusion, our study shows that individualised management in patients with SSc is a reasonable possibility. Using the developed prediction model, the chance for progressive disease could be increased from 39 to 67–89%, advocating annual stringent follow-up at least in patients with friction rubs, proximal muscle weakness and low maximum oxygen uptake at baseline. Future studies are needed to further optimise prediction of disease progression for the individual patient. In addition, maximum oxygen uptake as measured by CPET was identified as a possible new biomarker for progressive disease in SSc. This finding should be replicated in different cohorts of patients with SSc.

Author affiliations
[1]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[2]Haga Hospital, The Hague, The Netherlands
[3]Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands
[4]Department of Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands
[5]Department of Pulmonology, Leiden University Medical Center, Leiden, The Netherlands

REFERENCES

1. Nihtyanova SI, Tang EC, Coghlan JG, et al. Improved survival in systemic sclerosis is associated with better ascertainment of internal organ disease: a retrospective cohort study. *QJM* 2010;103:109–15.
2. van Laar JM, Farge D, Sont JK, et al. Autologous hematopoietic stem cell transplantation vs intravenous pulse cyclophosphamide in diffuse cutaneous systemic sclerosis: a randomized clinical trial. *JAMA* 2014;311:2490–8.
3. Assassi S, Sharif R, Lasky RE, et al. Predictors of interstitial lung disease in early systemic sclerosis: a prospective longitudinal study of the GENISOS cohort. *Arthritis Res Ther* 2010;12:R166.
4. Avouac J, Walker UA, Hachulla E, et al., the EUSTAR collaborators*; the EUSTAR collaborators. Joint and tendon involvement predict disease progression in systemic sclerosis: a EUSTAR prospective study. *Ann Rheum Dis* 2014. Published Online First. doi:10.1136/annrheumdis-2014-205295
5. Campo A, Mathai SC, Le Pavec J, et al. Hemodynamic predictors of survival in scleroderma-related pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2010;182:252–60.
6. Hesselstrand R, Wildt M, Ekmehag B, et al. Survival in patients with pulmonary arterial hypertension associated with systemic sclerosis from a Swedish single centre: prognosis still poor and prediction difficult. *Scand J Rheumatol* 2011;40:127–32.
7. Launay D, Sitbon O, Hachulla E, et al. Survival in systemic sclerosis-associated pulmonary arterial hypertension in the modern management era. *Ann Rheum Dis* 2013;72:1940–6.
8. Meune C, Avouac J, Airò P, et al. Prediction of pulmonary hypertension related to systemic sclerosis by an index based on simple clinical observations. *Arthritis Rheum* 2011;63:2790–6.
9. Moore OA, Goh N, Corte T, et al. Extent of disease on high-resolution computed tomography lung is a predictor of decline and mortality in systemic sclerosis-related interstitial lung disease. *Rheumatology (Oxford)* 2013;52:155–60.
10. Nihtyanova SI, Schreiber BE, Ong VH, et al. Prediction of pulmonary complications and long-term survival in systemic sclerosis. *Arthritis Rheumatol* 2014;66:1625–35.
11. Tyndall AJ, Bannert B, Vonk M, et al. Causes and risk factors for death in systemic sclerosis: a study from the EULAR Scleroderma Trials and Research (EUSTAR) database. *Ann Rheum Dis* 2010;69:1809–15.
12. Beretta L, Santaniello A, Cappiello F, et al. Development of a five-year mortality model in systemic sclerosis patients by different analytical approaches. *Clin Exp Rheumatol* 2010;28(2 Suppl 58): S18–27.
13. Bryan C, Knight C, Black CM, et al. Prediction of five-year survival following presentation with scleroderma: development of a simple model using three disease factors at first visit. *Arthritis Rheum* 1999;42:2660–5.

14. Domsic RT, Nihtyanova SI, Wisniewski SR, *et al*. Derivation and validation of a prediction rule for two-year mortality in early diffuse cutaneous systemic sclerosis. *Arthritis Rheumatol* 2014;66:1616–24.
15. Fransen J, Popa-Diaconu D, Hesselstrand R, *et al*. Clinical prediction of 5-year survival in systemic sclerosis: validation of a simple prognostic model in EUSTAR centres. *Ann Rheum Dis* 2011;70:1788–92.
16. Goh NS, Desai SR, Veeraraghavan S, *et al*. Interstitial lung disease in systemic sclerosis: a simple staging system. *Am J Respir Crit Care Med* 2008;177:1248–54.
17. Maurer B, Graf N, Michel BA, *et al*., EUSTAR co-authors. Prediction of worsening of skin fibrosis in patients with diffuse cutaneous systemic sclerosis using the EUSTAR database. *Ann Rheum Dis* 2015;74:1124–31.
18. [No authors listed]. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. *Arthritis Rheum* 1980;23:581–90.
19. LeRoy EC, Medsger TA Jr. Criteria for the classification of early systemic sclerosis. *J Rheumatol* 2001;28:1573–6.
20. van den Hoogen F, Khanna D, Fransen J, *et al*. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2013;72:1747–55.
21. American Thoracic Society; European Respiratory Society. American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias. This joint statement of the American Thoracic Society (ATS), and the European Respiratory Society (ERS) was adopted by the ATS board of directors, June 2001 and by the ERS Executive Committee, June 2001. *Am J Respir Crit Care Med* 2002;165:277–304.
22. Rudski LG, Lai WW, Afilalo J, *et al*. Guidelines for the echocardiographic assessment of the right heart in adults: a report from the American Society of Echocardiography endorsed by the European Association of Echocardiography, a registered branch of the European Society of Cardiology, and the Canadian Society of Echocardiography. *J Am Soc Echocardiogr* 2010;23:685–713.
23. Vacca A, Meune C, Gordon J, *et al*. Cardiac arrhythmias and conduction defects in systemic sclerosis. *Rheumatology (Oxford)* 2014;53:1172–7.
24. Baron M, Bernier P, Cote LF, *et al*. Screening and therapy for malnutrition and related gastro-intestinal disorders in systemic sclerosis: recommendations of a North American expert panel. *Clin Exp Rheumatol* 2010;28(2 Suppl 58):S42–6.
25. Penn H, Howie AJ, Kingdon EJ, *et al*. Scleroderma renal crisis: patient characteristics and long-term outcomes. *QJM* 2007;100:485–94.
26. Clements P, Lachenbruch P, Siebold J, *et al*. Inter and intraobserver variability of total skin thickness score (modified Rodnan TSS) in systemic sclerosis. *J Rheumatol* 1995;22:1281–5.
27. Gazi H, Pope JE, Clements P, *et al*. Outcome measurements in scleroderma: results from a Delphi exercise. *J Rheumatol* 2007;34:501–9.
28. van der Heijden GJ, Donders AR, Stijnen T, *et al*. Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: a clinical example. *J Clin Epidemiol* 2006;59:1102–9.
29. O'Brien RM. A caution regarding rules of thumb for variance inflation factors. *Quality Quantity* 2007;41:673–90.
30. Steyerberg EW, Vickers AJ, Cook NR, *et al*. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010;21:128–38.
31. Austin PC, Steyerberg EW. Graphical assessment of internal and external calibration of logistic regression models by using loess smoothers. *Stat Med* 2014;33:517–35.
32. Steyerberg EW, Harrell FE Jr, Borsboom GJ, *et al*. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* 2001;54:774–81.
33. Dumitrescu D, Oudiz RJ, Karpouzas G, *et al*. Developing pulmonary vasculopathy in systemic sclerosis, detected with non-invasive cardiopulmonary exercise testing. *PLoS ONE* 2010;5:e14293.
34. Walkey AJ, Ieong M, Alikhan M, *et al*. Cardiopulmonary exercise testing with right-heart catheterization in patients with systemic sclerosis. *J Rheumatol* 2010;37:1871–7.
35. Cuomo G, Santoriello C, Polverino F, *et al*. Impaired exercise performance in systemic sclerosis and its clinical correlations. *Scand J Rheumatol* 2010;39:330–5.
36. Nihtyanova SI, Ong VH, Denton CP. Current management strategies for systemic sclerosis. *Clin Exp Rheumatol* 2014;32(2 Suppl 81):156–64.
37. Valentini G. The assessment of the patient with systemic sclerosis. *Autoimmun Rev* 2003;2:370–6.
38. Minier T, Guiducci S, Bellando-Randone S, *et al*. Preliminary analysis of the very early diagnosis of systemic sclerosis (VEDOSS) EUSTAR multicentre study: evidence for puffy fingers as a pivotal sign for suspicion of systemic sclerosis. *Ann Rheum Dis* 2014;73:2087–93.
39. Medsger TA Jr. Natural history of systemic sclerosis and the assessment of disease activity, severity, functional status, and psychologic well-being. *Rheum Dis Clin North Am* 2003;29:255–73.
40. Vanthuyne M, Smith V, De Langhe E, *et al*. The Belgian Systemic Sclerosis Cohort: correlations between disease severity scores, cutaneous subsets, and autoantibody profile. *J Rheumatol* 2012;39:2127–33.
41. Walker UA, Tyndall A, Czirják L, *et al*. Clinical risk assessment of organ manifestations in systemic sclerosis: a report from the EULAR Scleroderma Trials And Research group database. *Ann Rheum Dis* 2007;66:754–63.