Nihtyanova Svetlana I. (Orcid ID: 0000-0001-5178-9038)
Denton Christopher P (Orcid ID: 0000-0003-3975-8938)

DYNAMIC PREDICTION OF PULMONARY HYPERTENSION IN SYSTEMIC SCLEROSIS USING LANDMARK ANALYSIS

Running head: Prediction of SSc-PH using landmark analysis

Authors:

Svetlana I. Nihtyanova, MSc MD[1,2]

Benjamin E. Schreiber, MA MD FRCP[3]

Voon H. Ong, PhD FRCP[1]

Athol U. Wells, PhD MD FRACP FRCP FRCR[4]

John G. Coghlan, MD FRCP[3]

Christopher P. Denton, PhD FRCP[1]

[1]Centre for Rheumatology and Connective Tissue Diseases, Royal Free Hospital, London, UK

[2]GSK, London, UK

[3]Pulmonary Hypertension Service, Royal Free Hospital, London, UK

[4]Interstitial Lung Disease Unit, Royal Brompton Hospital, London, UK

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/art.42349

Corresponding author: Professor Christopher P. Denton

UCL Centre for Rheumatology and Connective Tissue Diseases

2nd Floor – UCL Medical School Building

Royal Free Hospital

Rowland Hill Street

London NW3 2PF

E-mail: c.denton@ucl.ac.uk

Tel/Fax: +44 (0)20 7794 0432

Funding:

This work was funded in part by The Royal Free Charity through a generous bequest for Scleroderma Research from Richard King.

Disclosures:

SIN is an employee of and owns shares in GSK and has received consultancy fees from Roche (less than $10,000).

AUW reports consultancy fees from Boehringer Ingelheim, Roche and Veracyte

JGC reports grants and consultancy fees from Janssen Ltd and VicorePharma

CPD reports grants and personal fees (less than $10,000 each), all outside the submitted work, from Acceleron, Actelion, Arxx Therapeutics, Bayer, Boehringer Ingelheim, Corbus, CSL Behring, Galapagos, GlaxoSmithKline, Horizon, Inventiva, Roche, Servier, and Sanofi.

Other authors have no disclosures relevant to this work.

ABSTRACT

Objective

Pulmonary hypertension (PH) is a serious complication of systemic sclerosis (SSc). We explore prediction of short-term risk for PH using serial pulmonary function tests (PFTs) and other disease features.

Methods

Subjects with SSc, disease onset≥10 years prior to data retrieval, available autoantibody specificity and PFTs were included. Mixed effects modelling was used to describe change in PFTs over time. Landmarking was utilized to include serial assessments and stratified Cox proportional hazards regression analysis with landmarks as strata was used to develop the PH prediction models.

Results

We analysed 1247 SSc patients, 16.3% male, 35.8% with dcSSc. Anticentromere, anti-topoisomerase and anti-RNA polymerase antibodies were observed in 29.8%, 22.0% and 11.4% respectively and PH developed in 13.6%.

Over time diffusing capacity for carbon monoxide (DLco) and carbon monoxide transfer coefficient (Kco) declined in all SSc patients (up to 1.5%/year) but demonstrated much greater annual decline (up to 4.5% and 4.8% respectively) in the 5-7 years preceding PH diagnosis.

Comparison between multivariable models including either DLco, Kco or FVC/DLco ratio, demonstrated that both absolute values and change over preceding year in those measurements associate strongly with risk of PH (HR 0.93 and 0.76 for Kco and its change; HR

0.90 and 0.96 for DLco and its change; and HR 1.08 and 2.01 for FVC/DLco ratio and its change; p<0.001 for all). The Kco based model had the greatest discriminating ability (Harrell's C 0.903).

Conclusion

Our findings strongly support the importance of PFT trends over time in identifying patients at risk of PH.

Pulmonary hypertension (PH) is an important complication of systemic sclerosis (SSc), ultimately affecting 10-15% of all SSc patients. Unlike other SSc organ manifestations, which commonly develop early in the disease course, PH develops late in the disease. Hazard of PH is very low in the first three years of disease and thereafter is 1-2% per year (1-4). Although prognosis of SSc-PH is poor, with median survival of approximately 3 years (5, 6), emerging evidence suggests that earlier diagnosis and early treatment initiation despite mild symptoms is associated with a significant improvement in survival (6-10).

Potential screening tools for PH, separately or in combination, include clinical signs and symptoms, blood biomarkers, echocardiography, pulmonary function testing (PFT), cardiac MRI and cardiopulmonary exercise testing (11). Although several PH prediction models and algorithms exist, none utilize serial patient assessments and longitudinal data from PFTs over extended time periods (12-18).

PFTs are used routinely to screen for and monitor lung involvement in SSc. Annual PFT assessments are recommended for all SSc patients. Both spirometry and gas transfer should be assessed to evaluate for possible presence of restrictive or obstructive pulmonary disease or pulmonary vasculopathy (19). Interpretation of PFT results in isolation can be challenging in the context of SSc. Carbon monoxide transfer coefficient (Kco), the uptake of carbon monoxide per unit alveolar volume (VA), is relatively preserved in ILD, while serial decline in Kco is a specific measure of pulmonary vasculopathy (20). By contrast, both interstitial lung disease (ILD) and PH are associated with decline in diffusing capacity for carbon monoxide (DLco, the calculated product of measured Kco and measured VA). The ratio of forced vital capacity (FVC) and DLco (FVC/DLco) is also used in PH screening, with higher ratio predicting greater PH risk (21, 22).

Although FVC/DLco carries similar information as Kco, it is substantially more imprecise due to having the variability of FVC, Kco and VA from which it is calculated (20). In patients who develop both ILD and PH, the PFT changes are often mixed. In addition, smoking history and emphysema can further confound the interpretation of PFT measurements (23). The substantial intra- and interpatient variability in PFT measurements, DLco in particular, also means that a single PFT assessment may not be very informative and serial PFT results are needed for context.

Development of models for PH prediction in SSc is complicated by the interdependence of disease characteristics. For example, anti-centromere antibody (ACA) is associated with low risk of ILD (1), higher FVC and lower DLco compared to other antibodies, even in subpopulations where both PH and ILD have been excluded (24). Anti-topoisomerase I antibody (ATA) positive patients are at much higher risk of ILD and lower risk of PH than any other antibody group (1), but patients with severe ILD can develop group 3 PH. The different PFT measurements are also strongly intercorrelated, which may lead to multicollinearity when included in the same model. Time-varying disease characteristics used in prediction models, with potentially time-varying effects on outcome, make analysis and interpretation of longitudinal data challenging (25).

Landmark analysis was first described by Anderson et al. in 1983 (26). It was introduced as an unbiased approach to the comparison between survival time in responders and non-responders to chemotherapy among cancer patients. The methodology avoids the bias resulting from grouping patients based on a time-varying characteristic (response to treatment) that is assessed after the start of the survival time and permits detailed exploration of covariate interactions, while avoiding time-varying effects and confounding.

With this work we describe changes in serial PFT measurements (FVC, DLco and Kco) over time in a large single-centre cohort of patients with SSc and assess patterns of change that, together with other commonly assessed disease features, may predict short-term risk for PH. We utilize the landmarking methodology to include assessments that were done over a period of over two decades, at 12-month intervals, reflecting routine clinical care. We develop models that describe the associations between patient assessments and short-term risk of precapillary PH development.

METHODS

Patients and disease characteristics

All subjects had a confirmed diagnosis of SSc and fulfilled the 2013 ACR/EULAR SSc classification criteria. Patients were included in the study if they had disease onset (defined as time of first non-Raynaud's symptom) ≥10 years prior to data retrieval, had been tested for autoantibodies and had at least one PFT assessment. Cutaneous subset and organ complication definitions are included in the supplementary material.

As all patients were followed and diagnosed prior to the release of the proposed update to the hemodynamic definition of PH in 2019 (27), we used the Venice classification for PH, which specified mean pulmonary artery pressure (mPAP) ≥25 mmHg at rest and pulmonary artery wedge pressure (PAWP) ≤15 mmHg on right-sided cardiac catheterization (RHC) as criteria for PH. As it is often difficult to distinguish group 1 (connective tissue disease–associated pulmonary arterial hypertension) and group 3 (ILD–associated PH) in patients with moderate to severe ILD, we included both groups in the analysis and throughout the paper the term PH refers to precapillary PH.

Most PFTs were performed at the Royal Free Hospital and the Royal Brompton Hospital, London, UK, although some were done elsewhere, as many patients attending our specialist centre are not local to the hospital and are under shared care.

Statistical analysis

Descriptive statistics were used to summarise the cohort characteristics. To study serial FVC, DLco, Kco (% predicted) and FVC/DLco ratio we used random effects modelling. Possible non-linear associations between time and PFT results were explored using polynomials. For patients, who were not diagnosed with PH at the time of data extraction, the time variable was anchored at disease onset. In the PH patient group, time was centered at PH diagnosis to investigate changes in PFT results in the years preceding development of PH. As FVC/DLco ratio had a skewed distribution, this was log-transformed and modelled as log(FVC/DLco). Estimates were then exponentiated and reported as geometric means.

Mixed effects modelling and PFT trajectory prediction

Separate mixed effects models with only time as a predictor and FVC, DLco, Kco or log(FVC/DLco) as outcome were built in the whole dataset. For each measure we calculated patient-specific model-derived best linear unbiased predictions (BLUPs) of the random effects for the model parameters. For each subject we then calculated model-predicted values for FVC, DLco, Kco and log(FVC/DLco) at 12 month intervals over the follow-up, as a sum of the fixed-portion of the linear prediction and the predicted patient-specific random effects. The calculated values for log(FVC/DLco) were then back-transformed and included in the PH prediction models as FVC/DLco ratio.

Landmarking

The proposed method by Anderson *et al.* involves a choice of a time point during the follow-up (deemed a landmark) and determining patient characteristics at the landmark timepoint (26). Patients who have died or been lost to follow-up before that are excluded from the analysis and change in the characteristics after the landmark is ignored. Following this first publication, landmark analysis has become widely used (28). The methodology was developed further by van Houwelingen and Putter, who proposed the use of short-term prediction within a narrow "sliding window" (29). In this way, multiple short-term predictions could be made, using the values of time-varying patient characteristics at the landmark and predicting the cumulative incidence of an outcome within the pre-specified time window, starting from the landmark timepoint. The separate Cox models could further be combined into a stratified Cox "supermodel", using landmarks as strata.

Using landmark methodology, we can predict probability of PH development within a 12-month window from the landmark timepoint, accounting for the subject status at the landmark (29). To achieve this, we selected landmarks at yearly intervals, starting from 3 years from onset (Figure 1A). For each landmark we created a separate dataset, which included all patients who were alive and had not developed PH at that time point. Time-varying characteristics, including presence of organ complications and model-predicted PFT measures (absolute values and change over the preceding year) were recorded as of the landmark time point. Time from the landmark to the PH diagnosis was recorded, if this occurred within 12 months after the landmark, otherwise this was censored at 12 months (Figure 1B). For the time-to-event analysis

assessing predictors of PH development, all landmark datasets were stacked and analysed together.

Time-to-event analysis

Kaplan–Meier (KM) estimator was used to calculate survival. Cumulative incidence of PH over the whole follow-up period was calculated using 1-KM and the cumulative incidence function (CIF), adjusting for death as a competing risk (30). Comparison was made between 1-KM and CIF estimates to inform the need for competing risk consideration when developing the supermodels in the landmark dataset. Small differences in the estimates over short follow-up periods would indicate that death as a competing risk would have minimal effect on the estimation of PH risk. For the stacked dataset, we used stratified Cox proportional hazards regression analysis to assess predictors of PH, with each landmark dataset being analysed as a separate stratum (29). We tested the significance of interaction terms of the covariates and the landmark variable (treated as continuous) to assess for time-varying effects. Given the inherent associations between different PFT measurements, to avoid issues with multicollinearity, we developed multivariable models with either DLco, Kco or FVC/DLco ratio included. Model fit was tested using Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC). Predictive discrimination was assessed using Harrell's C statistic.

All analyses were performed using Stata 14.

RESULTS

Cohort description

A total of 1247 subjects with confirmed SSc diagnosis were included in the study. Demographic and clinical characteristics are summarised in Table 1. All subjects had disease onset between 10 and 22 years prior to data extraction with mean±SD follow-up of 12.6±5.4 years. Of the cohort, 133 (10.7%) were lost to follow-up (were not dead and had their last assessment more than 2 years prior to the year of data extraction). All subjects had at least one PFT over the follow-up, with 147 (11.8%) having two and 945 (75.8%) having three or more assessments. In total, 8165 PFT results were available and 6769 (82.9%) of those had results for FVC, DLco and Kco, while the remaining had missing data on one or two of the measurements. FVC was available in 8030 (98.4%), DLco in 7899 (96.7%) and Kco in 6838 (83.8%) of all PFTs. FVC/DLco ratio could be calculated from 7767 of the PFTs (95.1%). Mean time between PFTs was 15.3 months.

Over the follow-up, PH developed in 170 (13.6%) of the patients. KM and CIF estimates at 5, 10, 15 and 20 years from onset were 3.9%/3.8%, 9.4%/8.8%, 16.5%/14.8% and 23.0%/19.7% respectively. Approximately a third of the subjects had died at the time of data extraction with KM survival estimates at 5, 10, 15 and 20 years of 93.1%, 83.5%, 69.2% and 55.2%.

Pulmonary function test results

FVC

Out of the 1077 non-PH subjects, 1064 had FVC measurements, while among the 170 patients with PH, 141 had FVC assessments prior to PH diagnosis. In the non-PH group, FVC demonstrated little change over time. On average, FVC at baseline was 89% and over time there was a small, but significant increase of <1%/year with estimated average FVC at 10 and 20 years from onset of 93.6% and 93.0%. In the years preceding PH development, FVC showed

significant but clinically small non-linear decline of ≤1%/year, with estimated average FVC at PH diagnosis of 77.1%, while 5 and 10 years prior to that this was 81.9% and 84.0% respectively (Figures 2A & 2B, Supplementary Tables 1 & 2).

DLco

DLco was available for 1061/1077 non-PH and prior to PH diagnosis for 138/170 PH patients. In the non-PH group, DLco demonstrated small, but consistent annual decline of 0.8%-1.5%/year, with estimated average DLco at 1, 5, 10, 15 and 20 years of follow-up of 70.4%, 66.1%, 62.1%, 57.8% and 51.4% respectively. Modelling serial DLco measurements over the years preceding PH diagnosis revealed decline rates similar to those among non-PH patients up to approximately 7 years prior to PH diagnosis. After this, annual decline became greater than 1.5%, steadily increasing with every year closer to PH diagnosis (1.6%, 1.9%, 2.3%, 2.8%, 3.3%, 3.9% and 4.5% between year 7 and the time of PH confirmation on RHC (Figures 2C & 2D, Supplementary Tables 3 & 4). Model-estimated average DLco at 15, 10, 5 and 1 year before PH diagnosis was 64.6%, 59.7%, 52.9% and 40.7%.

Kco

Kco results were available for 983/1077 subjects without PH, while Kco measurements prior to PH diagnosis were available for 115/170 subjects with PH. Analysis of serial Kco was similar to that of DLco. Among the non-PH patients, Kco declined over time at a rate of 0.7-1.4%/year. Estimated average Kco at 1, 5, 10, 15 and 20 years of follow-up was 87.0%, 83.0%, 79.6%, 75.5% and 68.4%. Among those who developed PH, yearly decline increased steadily over the 6 to 7 years prior to PH diagnosis, with annual decline of 1.5%, 1.8%, 2.3%, 2.8%, 3.4%, 4.0% and 4.8% per year between year 7 and the time of RHC diagnosis. Model-estimated average Kco at 20,

15, 10, 5 and 1 year before PH diagnosis was 99.7%, 84.9%, 79.3%, 72.9% and 60.5% (Figures 3A & 3B, Supplementary Tables 5 & 6).

FVC/DLco

FVC/DLco ratio could be calculated in 1048/1077 non-PH subjects and was available prior to PH development in 137/170 PH patients. Among the non-PH patients, there was a gradual average increase in the FVC/DLco ratio over time by 0.02-0.04/year. At 1 year from onset, estimated average FVC/DLco ratio was 1.29, which increased to 1.42 at year 5, 1.54 at year 10, 1.67 at year 15 and 1.90 at year 20. In the PH group, there was similar gradual increase in the FVC/DLco ratio over the years preceding PH diagnosis, which varied between 0.02 and 0.04/year until around 7 years prior to PH diagnosis, when estimated FVC/DLco ratio was 1.5. Following that, we observed much faster annual increase, which became greater than 0.1/year in the final 3 years and average estimated FVC/DLco ratio was 1.6 at 5 years, 2.0 at 1 year before PH, and 2.2 at PH diagnosis (Figures 3C & 3D, Supplementary Tables 7 & 8).

Landmark analysis

Univariable analysis

Patient characteristics at each landmark are summarized in Supplementary Table 9. Associations between different variables and hazard of PH development from the univariable analyses are detailed in Table 2. There were significant associations between PH development within 12 months and all PFT measurements, including % predicted FVC, DLco, Kco and FVC/DLco ratio, as well as change in those over the preceding 12 months. Lower FVC, DLco and Kco, and decline in those increased the hazard of PH development. Conversely, higher FVC/DLco ratio and increase in the ratio over the preceding 12 months were associated with

higher risk of PH. There was evidence for interaction between change in PFTs and LM, suggesting that the increase in PH hazard for a unit drop in FVC, DLco, Kco or increase in FVC/DLco ratio became greater with longer disease duration (Supplementary Table 10).

Multivariable analysis

Multivariable models were developed with either DLco (with or without FVC), Kco (with or without FVC) or FVC/DLCO ratio and compared in terms of fit and discriminatory performance (Supplementary Table 11). The comparison suggested that FVC/DLco ratio is a poorer predictor of PH compared to DLco or Kco. The models including Kco performed better when adjusting for restrictive lung disease, either using FVC or ILD presence, while this was not the case when using DLCO. The three best models based on each PFT measure are summarized in Table 3.

DISCUSSION

With this work we describe the behavior of PFT measurements over a long period in a large, unselected cohort of SSc patients. We utilize landmark analysis to assess the association between FVC, DLco, Kco, FVC/DLco ratio and their annual rates of change, and short-term risk of PH development in patients followed for up to 22 years, using available serial assessments from that entire period. Our study expands findings from previous work, demonstrating the importance of PFTs, in particular gas transfer factor, in combination with autoantibodies, as predictors of PH development (18).

We confirm that over time DLco and Kco steadily decline in all SSc patients, irrespective of PH diagnosis, although much greater annual decline is observed in the 5-7 years preceding PH

diagnosis. While the average annual rates of change we report fall within the measurement variability of PFTs in individual patients, DLco and Kco trends over a two-year period or longer will show definite decline, exceeding measurement variability, which is a robust PFT signal that can easily be applied in clinical practice and could be used in PH risk stratification. It is important to acknowledge that the cohort average PFT measurements over time in non-PH patients may have been influenced by drop-out of patients due to severe ILD or death, which may explain the small increase in average FVC observed over time.

Most currently published models and algorithms for prediction of PH have been derived in subjects who are considered at risk of PH development (12-14). This restricts the applicability of such models and their extrapolation to the general SSc population could lead to incorrect predictions (31). In many models, predictions are based on cross-sectional data (14, 15, 17) or longitudinal data incorporated as a single summary value (for example, worst ever assessment or observation ever present over the entire follow up) (4). A more recent study from the Canadian Scleroderma Research Group registry utilized a more sophisticated methodological approach, including data from repeat visits prior to PH diagnosis, allowing for inclusion of both absolute PFT measurements and their change over the preceding year (16). Previous research in our centre, exploring predictors of pulmonary complications in SSc, used information available at baseline (within a window from disease onset) to identify patients at risk of PH development at any time during their follow-up (18). Of the multiple published screening tools for PH, the most robust to date is the DETECT score (17), which was derived in an enriched patient population with disease duration greater than 3 years and DLco<60%.

Although landmark analysis has been used extensively in oncology research, to our knowledge, so far only two studies have used this approach for prognostication in PH patients (32, 33). Mazurek et al. (32) utilized a single landmark, 1 year after baseline, to test if tricuspid annular plane systolic excursion assessed after 12 months of PAH therapy is predictive of survival. McLaughlin et al. (33) used three landmarks – at 3, 6 and 12 months, to investigate the association between PAH-related morbidity events and risk of death in the pooled data from the SERAPHIN and GRIPHON clinical trials (34, 35). For each of the landmarks, survival until the end of the study was compared between patients who had and had not experienced a PAH-related morbidity event before the landmark.

Using dynamic prediction with multiple landmarks, we develop models applicable to the entire length of the studied period and focus on prediction within a narrow time window, which reflects real life management of SSc-PH patients. We demonstrate how landmark analysis may be applied to a well characterized cohort of unselected SSc patients to define changes in PFT trajectory that could predict development of PH. This has relevance to current practice where routine interval assessment with PFT is incorporated into standard of care and there is added benefit in considering each measurement in the context of preceding values. It can also improve the selection of cases for further assessment and can define the broader profile of patients at risk to refine real world detection. In addition, the longitudinal dimension of this analysis and unselected nature of the cases included, make it more relevant to current practice than some previous analyses.

As it is often difficult to distinguish clinically group 1 and 3 PH, we take a pragmatic approach and include both in the analysis. This has become even more relevant with the recently

approved indication extension of inhaled treprostinil to include ILD-PH (36). The patients included in this study were seen and assessed prior to the publication of the proposed updated haemodynamic definition of pre-capillary PH (mPAP>20 mmHg, PAWP≤15 mmHg and pulmonary vascular resistance (PVR) ≥3 WU) and PH was defined by mPAP≥25mm Hg. Nevertheless, PVR≥3 WU remains a stringent threshold and several publications have presented evidence suggesting that the application of the new definition has limited effect on the overall PH diagnosis, identifying only a small number of additional cases (37-39). Consequently, it is unlikely our results would be substantially different if the new definition was used.

Similar to previous analyses of our SSc cohort, we found that of the SSc-specific autoantibodies, anti-U3RNP was the strongest predictor of PH. While this is a comparatively rare specificity among SSc patients, it is worth highlighting that it has higher frequency among non-Caucasians (40-42). ATA had the strongest negative association with PH development. In the multivariable analyses, ACA conveyed the second highest risk of PH development, which nevertheless was less than half the risk associated with anti-U3RNP.

Important confounders of the association between DLco or Kco and PH are smoking history and COPD/emphysema diagnosis. In this analysis smoking did not associate with PH, while history of COPD or emphysema was associated with an increased risk in the univariable analysis. The association did not hold consistently in the multivariable models, based on DLco or Kco possibly because the effect was conveyed through the PFT measurements, but remained a significant independent covariate in the model based on FVC/DLco ratio.

The best-performing models were based on Kco, adjusted for background restrictive lung disease (including either FVC or ILD as covariates). It is noteworthy that Kco is an alternative to

the FVC/DLco ratio and consideration should be given to the different PFT measures' variability. Compared to FVC and Kco, DLco has greater variability, as it is computed from measured Kco and VA. The FVC/DLco ratio then has the combined measurement variability of FVC, Kco and VA. As a result, in serial assessments of PFTs, a 10% decline in Kco predicts mortality among SSc-ILD patients, but a 20% increase in FVC/DLco ratio is required for the same effect on mortality (43). Consequently, the good fit of the models using Kco, especially when adjusting for FVC, would likely be at least partly due to Kco and FVC being less variable than DLco and FVC/DLco ratio. At the same time, even though DLco-based models show poorer fit than Kco-based ones, their discriminating ability is slightly better than that of Kco models that do not also adjust for FVC. One possible reason for this is because DLco captures both decline in Kco and VA (reflecting worsening ILD), which can also lead to PH. It is also possible that the poorer performance of FVC/DLco ratio is related to the inclusion of both Group 1 and Group 3 PH patients in the analysis.

Strengths of this study include the use of a robust dataset, based on a large number of patients followed for over two decades with serial clinical assessments done consistently in a single centre with low threshold for referral to RHC. Nevertheless, it is important to acknowledge limitations including that the exact timing of PH onset is not possible to determine, and we used the date of RHC-based diagnosis for the time-to-event analysis, which is likely to introduce noise in the data. Moreover, haemodynamic data were not complete for all patients as some of the diagnostic RHCs were performed prior to the use of electronic records or in other hospitals, therefore detailed exploration of associations between PFTs change and PH severity at diagnosis was not possible. Additionally, it was not possible to evaluate the role of other known

predictors of PH, such as echocardiogram-derived measurements and NT-proBNP levels, as those have not been measured routinely in all SSc patients in the past. A source of potential bias in the data is that patients with ILD and modest evidence for PH may have been less likely to be catheterized, while breathlessness and drop in DLco or increase in FVC/DLco will often prompt catheterization in the absence of other evidence of PH if no ILD is present. In addition, PFTs would likely be done more frequently in symptomatic patients with suspected PH development or ILD development or progression. The number and frequency of PFT assessments varied between patients, which necessitated the use of model-derived rather than observed values for the PFT measurements and their change and assumed data were missing at random. As a result of the PFTs being performed in several hospitals, part of the variability in results may be related to different approaches for calculation of % predicted values of PFT measurements. While PFT standardization is important, it is not possible in the context of prolonged observational studies. The data we analyse reflect real world clinical practice and the findings are therefore more broadly applicable.

In conclusion, we demonstrate that increase in FVC/DLco ratio and incremental decline in DLco and Kco precede PH development by more than 5 years and therefore could be used as early indicators of PH development. Prospective validation of our findings is needed for them to be applicable in practice, but they nevertheless strongly support the importance of PFT trends over time in identifying patients at risk of PH.

REFERENCES

1. Nihtyanova SI, Sari A, Harvey JC, Leslie A, Derrett-Smith EC, Fonseca C, et al. Using Autoantibodies and Cutaneous Subset to Develop Outcome-Based Disease Classification in Systemic Sclerosis. Arthritis Rheumatol. 2020;72(3):465-76.
2. Hachulla E, de Groote P, Gressin V, Sibilia J, Diot E, Carpentier P, et al. The three-year incidence of pulmonary arterial hypertension associated with systemic sclerosis in a multicenter nationwide longitudinal study in France. Arthritis Rheum. 2009;60(6):1831-9.
3. Vandecasteele E, Drieghe B, Melsens K, Thevissen K, De Pauw M, Deschepper E, et al. Screening for pulmonary arterial hypertension in an unselected prospective systemic sclerosis cohort. Eur Respir J. 2017;49(5).
4. Morrisroe K, Huq M, Stevens W, Rabusa C, Proudman SM, Nikpour M. Risk factors for development of pulmonary arterial hypertension in Australian systemic sclerosis patients: results from a large multicenter cohort study. BMC Pulm Med. 2016;16(1):134.
5. Ramjug S, Hussain N, Hurdman J, Billings C, Charalampopoulos A, Elliot CA, et al. Idiopathic and Systemic Sclerosis-Associated Pulmonary Arterial Hypertension: A Comparison of Demographic, Hemodynamic, and MRI Characteristics and Outcomes. Chest. 2017;152(1):92-102.
6. Lefèvre G, Dauchet L, Hachulla E, Montani D, Sobanski V, Lambert M, et al. Survival and prognostic factors in systemic sclerosis-associated pulmonary hypertension: a systematic review and meta-analysis. Arthritis Rheum. 2013;65(9):2412-23.
7. Humbert M, Yaici A, de Groote P, Montani D, Sitbon O, Launay D, et al. Screening for pulmonary arterial hypertension in patients with systemic sclerosis: clinical characteristics at diagnosis and long-term survival. Arthritis Rheum. 2011;63(11):3522-30.
8. White RJ, Vonk-Noordegraaf A, Rosenkranz S, Oudiz RJ, McLaughlin VV, Hoeper MM, et al. Clinical outcomes stratified by baseline functional class after initial combination therapy for pulmonary arterial hypertension. Respir Res. 2019;20(1):208.
9. Brown Z, Proudman S, Morrisroe K, Stevens W, Hansen D, Nikpour M. Screening for the early detection of pulmonary arterial hypertension in patients with systemic sclerosis: A systematic review and meta-analysis of long-term outcomes. Semin Arthritis Rheum. 2021;51(3):495-512.
10. Galiè N, Rubin L, Hoeper M, Jansa P, Al-Hiti H, Meyer G, et al. Treatment of patients with mildly symptomatic pulmonary arterial hypertension with bosentan (EARLY study): a double-blind, randomised controlled trial. Lancet (London, England). 2008;371(9630):2093-100.
11. Bruni C, De Luca G, Lazzaroni MG, Zanatta E, Lepri G, Airò P, et al. Screening for pulmonary arterial hypertension in systemic sclerosis: A systematic literature review. Eur J Intern Med. 2020;78:17-25.
12. Hsu VM, Chung L, Hummers LK, Wigley F, Simms R, Bolster M, et al. Development of pulmonary hypertension in a high-risk population with systemic sclerosis in the Pulmonary Hypertension Assessment and Recognition of Outcomes in Scleroderma (PHAROS) cohort study. Semin Arthritis Rheum. 2014;44(1):55-62.
13. Schreiber BE, Valerio CJ, Keir GJ, Handler C, Wells AU, Denton CP, et al. Improving the detection of pulmonary hypertension in systemic sclerosis using pulmonary function tests. Arthritis Rheum. 2011;63(11):3531-9.
14. Gladue H, Steen V, Allanore Y, Saggar R, Saggar R, Maranian P, et al. Combination of echocardiographic and pulmonary function test measures improves sensitivity for diagnosis of systemic sclerosis-associated pulmonary arterial hypertension: analysis of 2 cohorts. J Rheumatol. 2013;40(10):1706-11.

15. Meune C, Avouac J, Airò P, Beretta L, Dieudé P, Wahbi K, et al. Prediction of pulmonary hypertension related to systemic sclerosis by an index based on simple clinical observations. Arthritis Rheum. 2011;63(9):2790-6.
16. Semalulu T, Rudski L, Huynh T, Langleben D, Wang M, Fritzler MJ, et al. An evidence-based strategy to screen for pulmonary arterial hypertension in systemic sclerosis. Semin Arthritis Rheum. 2020;50(6):1421-7.
17. Coghlan JG, Denton CP, Grünig E, Bonderman D, Distler O, Khanna D, et al. Evidence-based detection of pulmonary arterial hypertension in systemic sclerosis: the DETECT study. Ann Rheum Dis. 2014;73(7):1340-9.
18. Nihtyanova SI, Schreiber BE, Ong VH, Rosenberg D, Moinzadeh P, Coghlan JG, et al. Prediction of pulmonary complications and long-term survival in systemic sclerosis. Arthritis Rheumatol. 2014;66(6):1625-35.
19. UK NPHCot, Ireland. Consensus statement on the management of pulmonary hypertension in clinical practice in the UK and Ireland. Heart (British Cardiac Society). 2008;94 Suppl 1:i1-41.
20. Hughes JM, Pride NB. Examination of the carbon monoxide diffusing capacity (DL(CO)) in relation to its KCO and VA components. Am J Respir Crit Care Med. 2012;186(2):132-9.
21. Steen V, Medsger TA, Jr. Predictors of isolated pulmonary hypertension in patients with systemic sclerosis and limited cutaneous involvement. Arthritis Rheum. 2003;48(2):516-22.
22. Steen VD, Graham G, Conte C, Owens G, Medsger TA, Jr. Isolated diffusing capacity reduction in systemic sclerosis. Arthritis Rheum. 1992;35(7):765-70.
23. Antoniou KM, Margaritopoulos GA, Goh NS, Karagiannis K, Desai SR, Nicholson AG, et al. Combined Pulmonary Fibrosis and Emphysema in Scleroderma-Related Lung Disease Has a Major Confounding Effect on Lung Physiology and Screening for Pulmonary Hypertension. Arthritis Rheumatol. 2016;68(4):1004-12.
24. Gunn J, Pauling JD, McHugh NJ. Impact of anti-centromere antibodies on pulmonary function test results in patients with systemic sclerosis without established or suspected pulmonary disease. Clin Rheumatol. 2014;33(6):869-71.
25. Daniel RM, Cousens SN, De Stavola BL, Kenward MG, Sterne JA. Methods for dealing with time-dependent confounding. Statistics in medicine. 2013;32(9):1584-618.
26. Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response. J Clin Oncol. 1983;1(11):710-9.
27. Simonneau G, Montani D, Celermajer DS, Denton CP, Gatzoulis MA, Krowka M, et al. Haemodynamic definitions and updated clinical classification of pulmonary hypertension. Eur Respir J. 2019;53(1).
28. Dafni U. Landmark analysis at the 25-year landmark point. Circ Cardiovasc Qual Outcomes. 2011;4(3):363-71.
29. van Houwelingen H, Putter H. Dynamic Prediction in Clinical Survival Analysis. CRC Press. 2012.
30. Pepe MS, Mori M. Kaplan-Meier, marginal or conditional probability curves in summarizing competing risks failure time data? Statistics in medicine. 1993;12(8):737-51.
31. Lau EM, Humbert M, Celermajer DS. Early detection of pulmonary arterial hypertension. Nature reviews Cardiology. 2015;12(3):143-55.
32. Mazurek JA, Vaidya A, Mathai SC, Roberts JD, Forfia PR. Follow-up tricuspid annular plane systolic excursion predicts survival in pulmonary arterial hypertension. Pulm Circ. 2017;7(2):361-71.
33. McLaughlin VV, Hoeper MM, Channick RN, Chin KM, Delcroix M, Gaine S, et al. Pulmonary Arterial Hypertension-Related Morbidity Is Prognostic for Mortality. Journal of the American College of Cardiology. 2018;71(7):752-63.
34. Sitbon O, Channick R, Chin KM, Frey A, Gaine S, Galiè N, et al. Selexipag for the Treatment of Pulmonary Arterial Hypertension. 2015;373(26):2522-33.

35. Pulido T, Adzerikho I, Channick RN, Delcroix M, Galiè N, Ghofrani HA, et al. Macitentan and morbidity and mortality in pulmonary arterial hypertension. N Engl J Med. 2013;369(9):809-18.
36. Waxman A, Restrepo-Jaramillo R, Thenappan T, Ravichandran A, Engel P, Bajwa A, et al. Inhaled Treprostinil in Pulmonary Hypertension Due to Interstitial Lung Disease. N Engl J Med. 2021;384(4):325-34.
37. Yamamoto K, Tanabe N, Takahashi Y, Naito A, Sekine A, Suda R, et al. Characteristics of patients meeting the new definition of pre-capillary pulmonary hypertension (Nice 2018) in a single Japanese pulmonary hypertension center. BMC Pulm Med. 2021;21(1):260.
38. Xanthouli P, Jordan S, Milde N, Marra A, Blank N, Egenlauf B, et al. Haemodynamic phenotypes and survival in patients with systemic sclerosis: the impact of the new definition of pulmonary arterial hypertension. Ann Rheum Dis. 2020;79(3):370-8.
39. Jaafar S, Visovatti S, Young A, Huang S, Cronin P, Vummidi D, et al. Impact of the revised haemodynamic definition on the diagnosis of pulmonary hypertension in patients with systemic sclerosis. Eur Respir J. 2019;54(2).
40. Steen V, Domsic RT, Lucas M, Fertig N, Medsger TA, Jr. A clinical and serologic comparison of African American and Caucasian patients with systemic sclerosis. Arthritis Rheum. 2012;64(9):2986-94.
41. Mejia Otero C, Assassi S, Hudson M, Mayes MD, Estrada YMR, Pedroza C, et al. Antifibrillarin Antibodies Are Associated with Native North American Ethnicity and Poorer Survival in Systemic Sclerosis. J Rheumatol. 2017;44(6):799-805.
42. Mendes C, Viana VST, Pasoto SG, Leon EP, Bonfa E, Sampaio-Barros PD. Clinical and laboratory features of African-Brazilian patients with systemic sclerosis. Clin Rheumatol. 2020;39(1):9-17.
43. Goh NS, Hoyles RK, Denton CP, Hansell DM, Renzoni EA, Maher TM, et al. Short-Term Pulmonary Function Trends Are Predictive of Mortality in Interstitial Lung Disease Associated With Systemic Sclerosis. Arthritis Rheumatol. 2017;69(8):1670-8.

ACKNOWLEDGEMENTS

We are grateful to Professor Bianca De Stavola for her invaluable advice on the statistical aspects of this project and in particular, landmark analysis.

FIGURE LEGENDS

Figure 1. Landmarking of the dataset. A) Landmarks were selected at yearly intervals, between years 3 and 22 from disease onset and separate dataset was created for each landmark. B) Dataset 7 as an example. This included all patients who were alive and had not developed PH at 7 years from onset. The values of time-varying characteristics, including presence of organ complications and model-predicted PFT measures (absolute values and change over the preceding year) were recorded as they were at 7 years from onset. PH events were recorded if they developed between years 7 and 8 from diagnosis. Time from year 7 to PH diagnosis was calculated and recorded, if this was <12 months, otherwise time was censored at 12 months. For the time-to-event analysis assessing predictors of PH development, all landmark datasets were stacked and analysed together.

Figure 2. Serial FVC and DLco measurements in the SSc cohort. Each thin line represents multiple measurements over time in a single patient; blue dots represent measurements in patients who had FVC or DLco assessed only once; thick green line is the model-predicted mean FVC or DLco measure. A) Observed serial FVC assessments over the entire follow-up in subjects who had not been diagnosed with PH at the time of data extraction. B) Observed serial FVC assessments in the years preceding PH diagnosis among patients who had been diagnosed with PH at the time of data extraction. C) Observed serial DLco assessments over the entire follow-up in subjects who had not been diagnosed with PH at the time of data extraction. D) Observed serial DLco assessments in the years preceding PH diagnosis among patients who had been diagnosed with PH at the time of data extraction.

Figure 3. Serial Kco and FVC/DLco ratio in the SSc cohort. Each thin line represents multiple measurements over time in a single patient; blue dots represent measurements in patients who had Kco and FVC/DLco ratio assessed only once; thick green line is the model-predicted mean Kco and FVC/DLco ratio. A) Observed serial Kco assessments over the entire follow-up in subjects who had not been diagnosed with PH at the time of data extraction. B) Observed serial Kco assessments in the years preceding PH diagnosis among patients who had been diagnosed with PH at the time of data extraction. C) Observed serial FVC/DLco ratio over the entire follow-up in subjects who had not been diagnosed with PH at the time of data extraction. D) Observed serial FVC/DLco ratio in the years preceding PH diagnosis among patients who had been diagnosed with PH at the time of data extraction. FVC/DLco outliers (values > 4; 43 observations from 8 patients) were excluded from the figure.

Table 1. Cohort characteristics

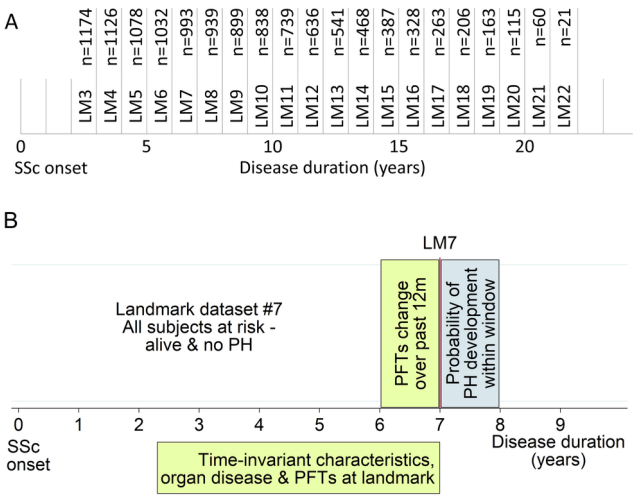
	n	(%)
Total number	1247	(100.0)
Male	203	(16.3)
Age at onset (years), mean ± SD	46.6	± 13.4
Emphysema/COPD history	92/1245	(7.4)
Smoking history	529/1175	(45.0)
Diffuse cutaneous subset	446	(35.8)
Overlap syndromes	250	(20.0)
Autoantibodies		
Anti-centromere	371	(29.8)
Anti-topoisimerase I	274	(22.0)
Anti-RNA polymerase	142	(11.4)
Anti-U3RNP	53	(4.3)
Anti-PmScl	55	(4.4)
Other, including anti-nRNP, hnRNP, rRNP, Th/To, SL, Ku, Jo1, Ro, La, XR, PL4, PL7, PL12, Sm	201	(16.1)
ANA+ ENA-	178	(14.3)
ANA negative	52	(4.2)
Organ complications		
Pulmonary fibrosis, any	564	(45.2)
Clinically significant pulmonary fibrosis	521	(41.8)
Pulmonary hypertension (Group 1 and Group 3)	170	(13.6)
Pulmonary arterial hypertension (Group 1)	132	(77.7)
Cardiac scleroderma	58	(4.7)
Scleroderma renal crisis	87	(7.0)
Death	404	(32.4)

Table 2. Univariable analysis

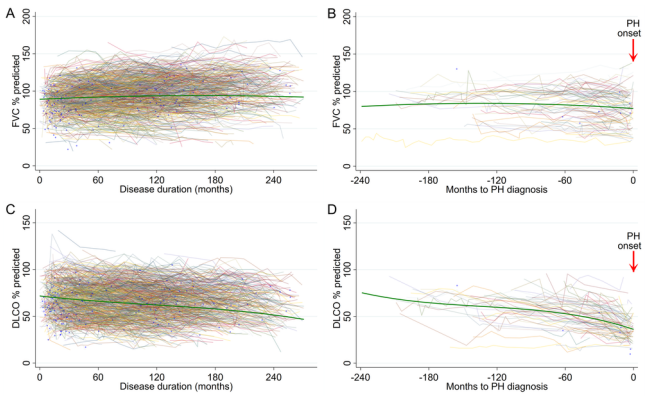
	HR	95% CI		p-value
Age at onset	1.02	1.01	1.04	0.001
Male	1.80	1.23	2.64	0.003
Smoking history	1.23	0.88	1.71	0.233
COPD/emphysema	1.92	1.22	3.01	0.005
Overlap	0.53	0.34	0.83	0.006
Diffuse cutaneous subset	1.04	0.73	1.47	0.830
Autoantibodies (ref. Anticentromere)				
Anti-topoisomerase 1	0.67	0.39	1.14	0.143
Anti-RNA polymerase	1.34	0.78	2.30	0.295
U3RNP	2.60	1.41	4.80	0.002
Anti-PmScl	0.61	0.22	1.71	0.350
ANA+ENA-	1.15	0.67	1.96	0.613
Other antibodies	1.54	0.97	2.44	0.065
Characteristics at landmark				
Age	1.02	1.01	1.04	0.001
Clinically significant ILD	1.96	1.42	2.71	<0.001
Cardiac scleroderma	0.92	0.29	2.88	0.883
Scleroderma renal crisis	2.01	1.11	3.63	0.020
PFTs at LM				
FVC%	0.97	0.96	0.98	<0.001
DLco%	0.91	0.90	0.92	<0.001
Kco%	0.92	0.91	0.93	<0.001
(FVC%/DLco%)*10	1.12	1.11	1.13	<0.001
PFTs change over 1 year preceding LM				
ΔFVC%	1.04	0.82	1.32	0.74
ΔFVC%*LM	0.95	0.92	0.97	<0.001
ΔDLco%	0.67	0.57	0.79	<0.001
ΔDLco%*LM	0.98	0.96	1.00	0.053
ΔKco%	0.75	0.63	0.89	0.001
ΔKco%*LM	0.96	0.94	0.98	0.001
Δ(FVC%/DLco%)*10	2.13	1.63	2.80	<0.001
Δ(FVC%/DLco%)*10*LM	1.07	1.02	1.11	0.001

Table 3. Multivariable analysis - models based on DLco, Kco or FVC/DLco ratio

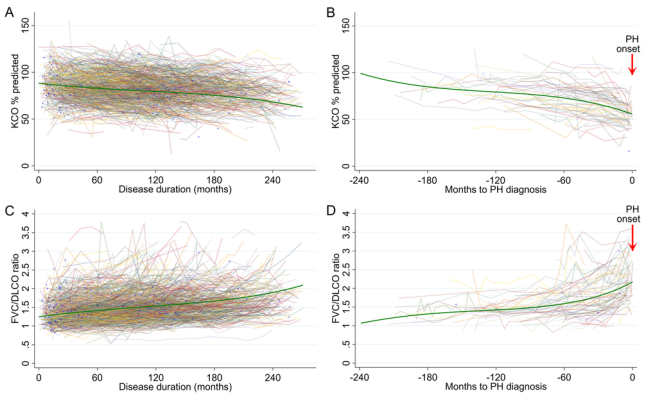
	HR	95% CI		p-value	HR	95% CI		p-value	HR	95% CI		p-value
Autoantibodies (ref. Anticentromere)												
Anti-topoisomerase 1	0.24	0.12	0.47	<0.001	0.18	0.10	0.32	<0.001	0.43	0.24	0.79	0.007
Anti-RNA polymerase	0.75	0.37	1.51	0.419	0.68	0.36	1.26	0.222	1.08	0.60	1.93	0.793
U3RNP	2.78	1.44	5.36	0.002	2.58	1.39	4.82	0.003	2.64	1.38	5.07	0.004
Anti-PmScl	0.37	0.13	1.09	0.072	0.36	0.13	1.01	0.053	0.58	0.20	1.65	0.307
ANA+ENA-	0.44	0.23	0.85	0.014	0.48	0.27	0.84	0.01	0.83	0.46	1.49	0.532
Other antibodies	0.61	0.35	1.06	0.082	0.53	0.33	0.86	0.011	0.74	0.44	1.25	0.258
Age	1.03	1.01	1.04	0.002								
Scleroderma renal crisis	2.57	1.21	5.48	0.014	2.42	1.21	4.82	0.012				
Clinically significant ILD									2.62	1.77	3.89	<0.001
COPD/emphysema									0.37	0.17	0.77	0.008
PFTs at LM												
KCO%	0.93	0.92	0.94	<0.001								
ΔKCO%	0.76	0.67	0.88	<0.001								
FVC%	0.97	0.96	0.98	<0.001								
ΔFVC%*LM	0.97	0.95	0.98	<0.001								
DLCO%					0.90	0.89	0.91	<0.001				
ΔDLCO%*LM					0.96	0.95	0.97	<0.001				
(FVC%/DLCO%)*10									1.08	1.05	1.12	<0.001
Δ(FVC%/DLCO%)*10									2.01	1.48	2.74	<0.001



ART_42349_Figure 1-600dpi.tif



ART_42349_Figure 2-600dpi.tif



ART_42349_Figure 3-600dpi.tif