SUITS
Statistical Analysis Plan

a randomized controlled trial to replace standard care with home monitoring for individuals with pulmonary fibrosis

Protocol id/Date:	11694/11 January 2023
SAP version number:	1.1
SAP version date:	15APRIL2026
Acronym:	SUITS
Sponsor name:	Erasmus MC
Address:	Rotterdam, Dep. of Respiratory Medicine, Postbus 2040, 3000 CA Rotterdam,

1. INTRODUCTION

This SAP describes the statistical analyses to be performed for the final analysis of the SUITS study. It is an expansion of the most recent version of the research protocol, version 1.1 of 11 January 2023.

The statistical analysis will process and present the results following the ICH standards, in particular the ICH-E3, ICH-E6, and ICH-E9 guidelines.

The aim of this study is to evaluate the impact of structurally replacing half of the outpatient clinic visits for patients with PF by home monitoring and remote consultations, also known as hybrid care, on patient activation, self-management and health(care) outcomes, compared to standard care.

1.1 STUDY OBJECTIVES AND ENDPOINTS

Objective	**Endpoint**
Primary	
To compare the effect of standard care and a home monitoring program integrated into standard care (hybrid care) on patients' self-management in patients with pulmonary fibrosis.	The change from baseline in PAM score from baseline to 12 months. *The PAM is a 13-item questionnaire assessing levels of patient activation in terms of managing their own health and healthcare. Total score ranges from 0-100, with higher scores indicating greater patient activation/better self-management.*
Secondary	
To compare the effect of standard care and a home monitoring program integrated into standard care (hybrid care) on healthcare utilization	The number of respiratory-related scheduled- and unscheduled outpatient clinic visits, remote consultations, emergency visits, and hospital admissions between baseline and 12 months.
To compare the effect of standard care and a home monitoring program integrated into standard care (hybrid care) on mortality and lung transplantation	The number of patients that deceased or received lung transplantation between baseline and 12 months.
To compare the effect of standard care and a home monitoring program integrated into standard care (hybrid care) on symptoms and impact on quality of life	The change in symptoms and quality of life in L-PF for the domains symptoms total, dyspnea, cough, energy and impacts from baseline to 12 months.

To compare the effect of standard care and a home monitoring program integrated into standard care (hybrid care) on health status.	The change in quality of life in K-BILD for the domains breathlessness and activities, chest, psychological wellbeing and overall score from baseline to 12 months.
To compare the effect of standard care and a home monitoring program integrated into standard care (hybrid care) on quality of life	The change in health status for EQ-5D-5L index and VAS from baseline to 12 months.
To compare the effect of standard care and a home monitoring program integrated into standard care (hybrid care) on pulmonary function in patients with pulmonary fibrosis.	The change in FVC and DLco from baseline to 12 months.
To evaluate the reliability and validity of home-based FVC measurements using a handheld spirometer in patients with pulmonary fibrosis, compared to standard in-hospital spirometry.	Reliability: intraclass correlation between home-based- and in-hospital FVC; Validity: mean difference, correlation, and limits of agreement between home-based and in-hospital FVC measurements.
Assess adherence to the home monitoring program in the hybrid care group.	Adherence to home spirometry, pulse-oximetry and questionnaires will be assessed from baseline to 12 months.
Explorative	
	Cost-effectiveness and cost-utility analyses from healthcare system and societal perspective.
	Healthcare professional- and patient satisfaction with hybrid care.
	Change in home-based FVC from baseline to 12 months.
	Time to detection of a decline of 10% in in-hospital versus home-based FVC.
	Evaluation of the effect of a home monitoring program on travel movements.

Legend: DLco, Diffusing Capacity of the Lungs for Carbon Monoxide; EQ-5D-5L, EuroQoL 5-Dimension 5-Level; FVC, Forced Vital Capacity; K-BILD, King's brief Interstitial Lung Disease; L-PF, Living with Pulmonary Fibrosis; PAM, Patient Activation Measure; VAS, Visual Analogue Scale.

1.2 STUDY DESIGN

This is a prospective, open-label, randomised clinical multicenter study to evaluate if home monitoring as part of a hybrid care pathway could effectively be implemented for patients with PF. Patients will be randomly assigned in a 1:1 ratio to receive either standard care or hybrid care, i.e. the home monitoring program replacing half of the annual outpatient clinic visits.

Standard care involves outpatient clinic visits every three months including in-hospital lung function testing. Patients in the hybrid care group will receive a home spirometer, pulse oximetry meter, and access to a smartphone or tablet app (SUITS app; see below for more details), where they can perform home based measurements and collect patient-reported outcome measurements (PROMs).

Furthermore, this app allows for direct communication with the hospital via eConsultation as well as video consultations. In the hybrid care group, half of the outpatient clinic visits will alternately be replaced by remote visits via video consultation.

FIGURE 1: SCHEMA FOR PARTICIPANTS ENTERING THE STUDY

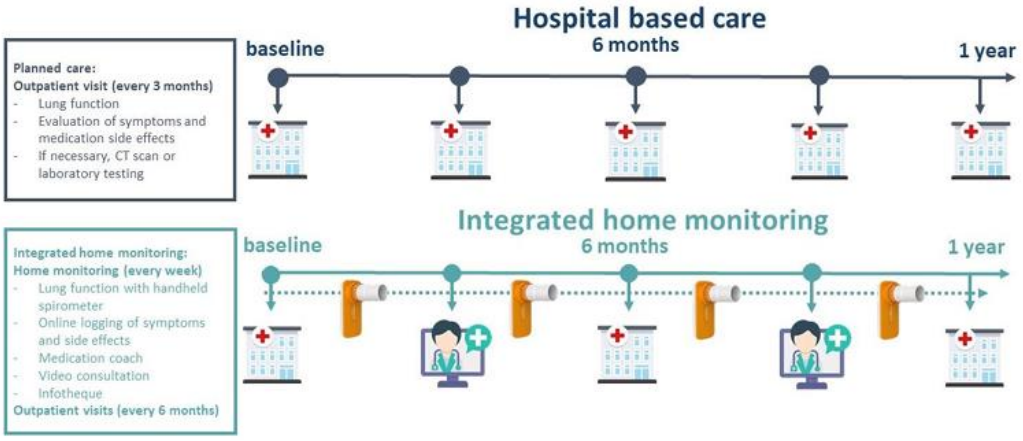

Table 1 shows the schedule of assessments for both groups. The total study duration for individual participants will be 12 months.

TABLE 1: SCHEDULE OF ASSESSMENTS

Outcome measurements		**Timing of measurements**					
		Baseline	**3 months**	**6 months**	**9 months**	**12 months**	**Full study period**
Questionnaires							
	PAM						
	Hybrid care group	X		X		X	
	Standard care group	X		X		X	
	K-BILD						
	Hybrid care group	X	X	X	X	X	

	Standard care group	X		X		X	
	L-PF						
	Hybrid care group	X		X		X	
	Standard care group	X		X		X	
	EQ-5D-5L+R						
	Hybrid care group	X		X		X	
	Standard care group	X		X		X	
	VAS symptoms						
	Hybrid care group	X	Weekly	Weekly	Weekly	Weekly	
	Standard care group	X		X		X	
	PESaM side-effects						
	Hybrid care group	X	Monthly	Monthly	Monthly	Monthly	
	Standard care group	X		X		X	
Cost-effectiveness							
	Formal healthcare costs, assessed with a questionnaire (iMCQ)* and patient file extraction						
	Hybrid care group		X	X	X	X	
	Standard care group		X	X	X	X	
	Informal healthcare costs, assessed with a questionnaire (iPCQ)*						
	Hybrid care group		X	X	X	X	
	Standard care group		X	X	X	X	
	Medication adjustments						
	Hybrid care group						X
	Standard care group						X
Physiology							
	In-hospital spirometry						
	Hybrid care group	X		X		X	
	Standard care group	X	X	X	X	X	
	Home spirometry						
	Hybrid care group	Weekly	Weekly	Weekly	Weekly	Weekly	

	Pulse-oximetry						
	Hybrid care group	Weekly	Weekly	Weekly	Weekly	Weekly	
Adoption							
	Number of patients actively performing home monitoring after 12 months						
	Hybrid care group						X
	Patient adherence to home spirometry, oximetry, and questionnaires						
	Hybrid care group						X
Healthcare utilization							
	Scheduled contact (number of scheduled video consultations and outpatient clinic visits, as well as time spent (in minutes) by the healthcare professional)						
	Hybrid care group						X
	Standard care group						X
	Unscheduled contact (number of respiratory-related unscheduled video consultations, phone calls, e-mails, digital messages, outpatient clinic visits, emergency ward visits, and hospital admissions, as well as time spent (in minutes) by the healthcare professional and whether the contact was initiated by an app alert)						
	Hybrid care group						X
	Standard care group						X
Mortality and lung transplantation							X

*Patient-reported healthcare utilization is assessed using an adapted version of the Medical Consumption Cost Questionnaire (iMCQ). This instrument captures healthcare utilization in terms of contacts, visits, and additional testing across different healthcare organizations (e.g. general practitioner). The Productivity Cost Questionnaire (iPCQ) assesses informal healthcare utilization, for example travel movements and the number of workdays missed by the patient due to illness.

1.3 EXPECTED SAMPLE SIZE

The sample size is targeted at 220 randomized participants.

The primary endpoint in our study is the between-group difference of the change in patient activation measure (PAM) score from baseline to 12 months. The sample size was calculated with data from a previous observational study in patients with IPF and sarcoidosis[17], where the standard deviation (SD) was 9.1 for patients with IPF, and 11.0 for patients with sarcoidosis. For this power calculation we used the mean SD of that study, which was 10.2. With 99 patients per group, the study will have 80% power

to detect a clinically significant between-group difference in the change in total PAM score with a confidence interval of 95%. To allow for 10% dropout, we aim to include 220 patients in our study.

1.4 RANDOMIZATION AND BLINDING

Patients will be randomly assigned in a 1:1 ratio to receive either standard care or the home monitoring program as part of a hybrid care pathway. Randomization will be done with a computer-generated schedule (ALEA Screening and Enrolment Application Software), using block-randomization, stratified for study site and use of anti-fibrotic treatment (yes or no). Participants, health-care providers, and research staff will not be masked to group allocation.

1.5 SOFTWARE

Data will be analysed using SPSS version 25.0 or later, R Studio version 4.5.2 or later and/or SAS version 9.4 or later (SAS Institute Inc., Cary, NC, USA).

2. GENERAL METHODOLOGY

2.1 ANALYSIS SETS

2.1.1 Analysis sets

The following analysis sets will be used:

- Modified intent-to-treat set (MITT): all randomized participants who had at least a complete baseline PAM assessment.
- Per protocol analysis set (PP): all randomized patients without major protocol deviations, defined as patients who completed the study for 12 months and used home spirometry for at least 80% of the time in the hybrid care group.

2.2 PHASES AND TIME POINTS

2.2.1 Phases

This study has one study phase starting at the time of randomisation (baseline day) and ends at visit month 12 or at date of earlier study termination.

2.2.2 Baseline and change from baseline

The last assessment performed on or before the day of randomisation will be used as baseline value. Changes from baseline will be based on these values. As this is a study without intervention assessments performed up to 14 days after the day of randomisation still be considered as baseline assessment.

2.2.3 Analysis visits

All assessments will be allocated to analysis visit windows. Tables and listings will present the analysis visit as defined below, not the eCRF visits. Allocations of assessments will be performed using Target day according to Table 2 below:

TABLE 2: ANALYSIS VISITS

Phase	Analysis visit	Target day	Lower limit day (-28)	Upper limit day (+28)
Treatment	baseline	1	-14	14
	Month 3	101	73	129
	Month 6	192	164	220
	Month 9	283	255	311
	Month 12	374	346	402

Per parameter and analysis visit, the non-missing value closest to the target day will be used in analysis, other values will only be listed. If more than one non-missing value is located at the same distance from the target, then the one latest in time will be selected for analysis. Partially missing assessment dates - disabling allocation to analysis visits - will not be imputed and thus these assessments will not be considered in the per timepoint analysis.

3. GENERAL CHARACTERISTICS ANALYSES

The following participant data will be tabulated for all participants.

3.1 PARTICIPANT DISPOSITION

- The number of participants and the number and percentage of prescreening failures (including reasons for prescreening failures);
- The number and percentage of participants for each analysis visit;
- Descriptive statistics of the study duration. Study duration is calculated as study end date – study start date;
- The number and percentage of participants who completed or discontinued the study and the number and percentage of participants for each study discontinuation reason and the number of ongoing participants.

3.2 PROTOCOL DEVIATIONS AND ELIGIBILITY

The number and percentage of participants with protocol deviations will be summarized overall, by important/non-important deviations, and by class of deviation using the SAF analysis set.

All information concerning important and nonimportant protocol deviations, violations on eligibility criteria, and participant compliance with hybrid care will be listed.

3.3 DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

3.3.1 Available data

The following parameters will be available:

- Demographics: sex at birth, age at informed consent, ethnicity (Asian, Black or African American, White, other: combining all remaining categories), height, body weight at baseline, year and month of birth, date of signing ICF
- Baseline characteristics:
    - Diagnosis at baseline
    - Disease-modifying treatment (antifibrotic medication, immunomodulatory medication)
    - Supplementary oxygen (yes/no)
    - Comorbidities n(%) (pulmonary arterial hypertension, lung cancer, obstructive sleep apnea, cardiac disease, depressive disorder)
    - Education level (university, higher vocational education, intermediate vocational education, secondary education, primary education)
- Baseline assessments
    - Pulmonary function (FVC % of predicted, FVC litres, DLco % of predicted, DLco (mmol/min/kPa))
    - Total PAM Score at baseline
    - L-PF score at baseline (symptoms total, dyspnea, cough, energy, impacts)
    - K-BILD score at baseline (total, breathlessness and activities, chest, psychological)
    - EQ-5D-5L score at baseline (index and EQ-VAS scale)

3.3.2 Presentation of results

Demographics will be summarized using descriptive statistics for age. Frequency tabulations will be provided for sex.

Baseline disease characteristics and assessments will be presented using descriptive statistics for diagnosis at baseline, education level, disease-modifying treatment, supplementary oxygen, comorbidities, pulmonary function, total PAM score, L-PF score at baseline, K-BILD scores at baseline, and EQ-5D-5L scores at baseline.

All demographic data and baseline disease characteristics will be listed.

4. EFFICACY AND QUALITY OF LIFE

The MITT analyses set will be used to analyse the efficacy and quality of life data.

4.1 PRIMARY ENDPOINT

The primary endpoint is the change in total PAM score from baseline to 12 months.

The Patient Activation Measure 13 (PAM-13) is a non-disease-specific measure that assesses patient knowledge, skills, and confidence for disease self-management. The PAM-13 consists of 13 items on a 4-point Likert scale (1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree). Item scores are summed up to a raw sum score resulting in theoretical values between 13 and 52, which are then transformed to a standardized metric ranging from 0 to 100. Higher scores indicate a greater patient activation.

PAM-13 scores can be categorized into four stages of activation, corresponding to the difficulty of the PAM-13 items:

- Level 1 (patients believe active role is important; items 1–2),
- Level 2 (patients have confidence and knowledge to take action; items 3–8),
- Level 3 (taking action; items 9–11) and
- Level 4 (staying on course under stress; items 12–13).

Levels are categorized according to the following cut-off thresholds: 1) level 1, ≤ 47; 2) level 2, 47.1–55.1; 3) level 3, 55.2–67; 4) level 4 ≥ 67.1).

For the present analysis, PAM scores will only be included if there are answers to at least seven items. In case of missing data, the total score was divided by the number of completed items and multiplied by 13 to get the sum raw score: 13*(sum of all answers/number of items answered).

A mixed model for repeated measures (MMRM) will be used to obtain the estimate of the group difference, based on changes from baseline, at month 12 using the data from month 6 and 12 visits. The model will include factors for study arm, visit, study arm by visit interaction, antifibrotic drug use, and center and baseline PAM score as covariate. An unstructured variance-covariance structure will be used.

4.1.1 Missing data

To assess the effect of missing data on the primary analysis outcome, the primary endpoint analyses will be repeated using multiple imputation (MI) analyses. These analyses will only be performed in case more than 10% of the patients miss their month 12 PAM assessment.

4.1.2 Per protocol Analyses

The primary endpoint analyses will also be repeated on the per protocol analyses set.

A mixed model for repeated measures (MMRM) in a similar way as in the MITT analysis will be used to obtain estimates of the group difference at month 12 using the data that will be collected at 0, 6 and 12 months in the per protocol population.

4.2 SECONDARY ENDPOINT

Secondary to the primary analysis:

1. Observed values and change from baseline in PAM score at baseline, 6, and 12 months.
2. Percentage change from baseline in total PAM score to 12 months using responders. A responder is defined as a participant who shows a percentage change from baseline in PAM score ≥ the minimal important change (MIC), defined as 4-points change.
3. Distribution of the total PAM score at each planned visit, using pre-defined level categories:
    a. Level 1 (low activation), score 0–47
    b. Level 2 (moderate activation), score 47.1–55.1
    c. Level 3 (average activation), score 55.2–67
    d. Level 4 (high activation), score 67.1–100

Other secondary endpoints:

1. The number of respiratory-related scheduled- and unscheduled outpatient clinic visits, remote consultations, emergency visits, and hospital admissions between baseline and 12 months.
2. The number of patients that deceased or received lung transplantation between baseline and 12 months.
3. Change from health status in EuroQoL 5-Dimensions 5-Level (EQ-5D-5L) index and VAS from baseline to 12 months.
4. Change in health status in King's brief Interstitial Lung Disease (K-BILD) in the domains breathlessness and activities, chest, psychological wellbeing, and overall score from baseline to 12 months.
5. Change in symptoms and quality of life in Living with Pulmonary Fibrosis (L-PF) in the domains symptoms total, dyspnea, cough, energy, and impacts from baseline to 12 months.
6. Change in in-hospital Forced vital capacity (FVC) from baseline to 12 months.
7. Change in Diffusing Capacity of the Lungs for Carbon Monoxide (DLco) from baseline to month 12.
8. Reliability and validity of home-based FVC measurements compared to in-hospital spirometry in the hybrid care group.
9. Adherence to the home monitoring program in the hybrid care group.

4.2.1 Analyses

In secondary analyses of the primary endpoint, we will analyse PAM scores at 0-, 6- and 12-months follow-up. In addition to the primary analysis of PAM scores based on mean values (group level), a responder analysis will be conducted using the proportional change in PAM scores from baseline to 12 months in both study groups based on the minimal important change (MIC) for PAM adjusted for the baseline PAM value. In addition, the distribution of the total PAM score will be analysed using set categories, including 1) scores from 0 to 47 are considered low activation; 2) scores from 47.1 to 55.1 are considered moderate activation; 3) scores from 55.2 to 67 are considered average activation; and 4) scores from 67.1 to 100 are considered high activation. The change in EQ-5D-5L VAS and index scores, K-BILD domain scores, and L-PF domain scores from baseline to 12 months follow-up will be compared between the two groups by using MMRM in a similar way as in the primary endpoint analysis. Healthcare utilization between baseline and 12 months will be assessed using descriptive

statistics to visualize the count of healthcare events. In addition, relative risks (RR) with 95% confidence intervals will be calculated to compare respiratory-related clinical events (visits to the emergency ward, unscheduled outpatient clinic visits, and hospital admissions) between groups. Change in in-hospital FVC and DLco from baseline to 12 months will be calculated for each participant. Changes in in-hospital FVC and DLco over time (baseline, 3, 6, 9, 12 months) will also be analyzed using a MMRM in a similar way as in the primary endpoint analysis.

To assess the reliability of home-based FVC, we will calculate the Intraclass Correlation Coefficient (ICC) between home-based and in-hospital FVC at each study visit. Validity will be assessed by comparing home-based and in-hospital FVC measurements at each study visit using mean differences and Bland-Altman analysis. The correlation between slopes from baseline to 12 months will be evaluated using a repeated measures model.

Adherence to home spirometry, pulse-oximetry and PROMs will be assessed by dividing the actual number of performed measurements by the expected number of measurements at 12 months and presented in %.

4.2.2 Presentation of results

Summary tables will only show analysis of the planned visits. All results will be listed.

Total PAM scores, EQ-5D-5L VAS and index scores, K-BILD breathlessness and activities, chest, psychological wellbeing, and overall scores, L-PF symptoms total, dyspnea, cough, energy, and impacts, in-hospital FVC (% of predicted and litres), and in-hospital DLco (% of predicted and (mmol/min/kPa)) will be summarized using descriptive statistics at each analysis visit. Actual values and change from baseline will be shown in the same table. Binary endpoints (responders) will be tabulated for total PAM scores using number of participants and proportions per arm based on the minimal important change (MIC) for PAM adjusted for the baseline PAM value. Frequency tabulations of the categories of total PAM total score will be presented by analysis visit, for both actual values and changes from baseline.

Healthcare utilization between baseline and 12 months will be presented in frequency tables, including relative risks of clinical events with 95% confidence intervals. Mortality and lung transplantation will be reported in the same frequency table.
Reliability results of home-based FVC will be reported as intraclass correlation coefficients (ICC). Validity results will be presented as mean differences, correlations, and Bland-Altman limits of agreement with confidence intervals.
Adherence to home spirometry, pulse oximetry, and questionnaire completion will be summarized descriptively.