BMJ Open

PRODIGY-ILD: Data driven predicted outcomes in interstitial lung disease.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2024-088271
Article Type:	Protocol
Date Submitted by the Author:	01-May-2024
Complete List of Authors:	Gunne, Emer; University College Dublin, School of Medicine Holden, Sinead; University College Dublin, Franciosi, Alessandro; St Vincent's University Hospital, Respiratory Medicine Keane, Michael; University College Dublin, McCarthy, Cormac; University College Dublin, School of Medicine; Saint Vincent's University Hospital, Doran, Peter; University College Dublin School of Medicine and Medical Science, School of Medicine
Keywords:	Interstitial lung disease < THORACIC MEDICINE, Observational Study, Patient Reported Outcome Measures, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS

Title:

PRODIGY-ILD: Data driven predicted outcomes in interstitial lung disease.

Corresponding Author:

Emer Gunne[1]

Email: emer.gunne@ucdconnect.ie

ORCID-ID 0000-0002-6939-1586

Other Authors:

Sinead Holden[2],

ORCID-ID 0000-0002-7314-2152

Alessandro N Franciosi[3],

ORCID-ID 0000-0002-4241-8718

Michael P. Keane[1,3],

ORCID-ID 0000-0003-4507-0742

Cormac McCarthy[1,3],

ORCID-ID 0000-0003-2896-5210

Peter Doran[1,3]

ORCID-ID 0000-0003-2064-5335

Affiliations:

1. University College Dublin, School of Medicine, Dublin, Ireland
2. University College Dublin, School of Public Health, Physiotherapy & Sports Science, Dublin, Ireland
3. St. Vincent University Hospital, Elm Park, Dublin, Ireland

Word count:

Abstract: words 297

Main document text: 3,450 words

Tables & Figures:

1 numbered Figure

4 numbered Tables

ABSTRACT:

Introduction:

Interstitial lung disease (ILD) is a heterogeneous group of non-infectious, predominantly diffuse and usually chronic respiratory disorders. Patients may develop a progressive phenotype characterised by progressive pulmonary fibrosis. Whilst these conditions are invariably life-limiting, wide variations in the clinical course of the disease have made it difficult to predict outcomes such as rate of progression, onset of acute exacerbations and mortality. New methods to predict the clinical course of ILD are needed for both treatment planning and clinical trial design

Advances in digital health technologies has facilitated the ability to collect "real-time" data to monitor diseases. The objective of this study is to develop prediction models for patients with progressive fibrotic ILD using comprehensive datasets captured using digital health technology in addition to clinical record information. Creating a virtual representation of a patient generated by integrating decentralised data sources. Prediction modeling datasets have the potential to be used as cost-effective synthetic control arms for clinical trials.

Methods and analysis:

This study is a prospective cohort study with 50 participants.

Patient inclusion criteria: Age 18 years or older with a diagnosis of Idiopathic Pulmonary Fibrosis or Progressive Pulmonary Fibrosis and the ability to provide written informed consent.

Patient exclusion criteria: Age under 18 years or unwilling to wear an Apple watch for the duration of the study.

Participants will be provided with an Apple watch to passively collect physiological and activity data. This data will be combined with clinical history and course, in addition to a set of patient reported outcome measures. Participants will be followed for 3 years to assess rate of disease progression, occurrence of acute exacerbations and mortality allowing comparison of actual outcomes with outcomes predicted from both dynamic prediction models and machine learning approaches. Interim analysis will be conducted at 6 monthly intervals.

ARTICLE SUMMARY

Strengths and limitations of this study

- Real-time passively collected data will give deeper insights to patient physiological measures, while not over burdening the participant to provide active measures.
- Study devices, training on device use and ongoing support will be provided,

overcoming potential sources of bias and barriers to participation in terms of access to technology and tech literacy.

- Interstitial lung disease is a rare disease therefore our sample size is small at 50 participants; however, we will collect a large amount of data points.
- Interstitial Lung Disease Outpatient Clinic attendees in Ireland are not an ethnic diverse population, which may limit the generalizability of the results.

MAIN BODY

Introduction

Interstitial lung disease (ILD) is a heterogeneous group (over 200 disorders) of non-infectious, predominantly diffuse and usually chronic respiratory disorders[1]. The disease affects the interstitium as well as the alveolar and airway architecture[2]. Some interstitial lung diseases are characterised by progressive pulmonary fibrosis, such as in Idiopathic Pulmonary Fibrosis (IPF) and Progressive Pulmonary Fibrosis (PPF)[3]. IPF is a chronic, fibrosing interstitial pneumonia of unknown cause, associated with radiological and histologic features of usual interstitial pneumonia[3].

In 2017, the incidence of ILD in Ireland was 7.66 per 100,000 population for men and 4.2 per 100,000 for women[4]. Mortality rates were 5.18 per 100,000 and 2.73 per 100,000 respectively[4], this is in line with the median incidence in Europe in 2017[4]. However, these figures could be underestimated, as the British Lung Foundation have reported 6,000 people diagnosed with IPF annually. IPF occurs primarily in the older adult population, with onset of disease in the sixth or seventh decade[6]. More men have been reported with IPF than women, and most patients have a history of past cigarette smoking[7]. Specialist care for ILD patients is delivered by respiratory multidisciplinary care teams at 8 clinical centres across Ireland.[8]

Clinical manifestations include chronic exertional dyspnea and cough. The disease is characterised by progressive worsening of dyspnoea and lung function, progressive fibrosis on high resolution computed topography, acute respiratory decline with a median survival of 3-5 years[2,3]. Patients with IPF are susceptible to abrupt declines in lung function (acute exacerbations) which typically develop in less than 1 month, are accompanied by new radiologic abnormalities on high resolution computed topography[9], and are a major indicator of morbidity and mortality, with acute exacerbation preceding approximately 40 percent of IPF deaths. Median survival following an acute exacerbation is approximately three to four months[9]. A significant challenge is the fact that IPF and other ILDs often follow an unpredictable clinical course, which impedes physicians' ability to predict rate of disease progression, acute exacerbations, and survival[10]. This impacts new therapeutic development,

as the absence of clear signals of progression impacts on trial design. Given ILDs are rare diseases, clinical investigations require additional precision.

Numerous efforts have been made to develop reliable prediction models for ILD patients[11]. The upgraded CRP (Clinical, Radiological, and physiological[12]) scoring system incorporates several parameters: age, smoking history; clubbing; extent of profusion of interstitial opacities, and presence or absence of pulmonary hypertension on the chest radiograph; percent predicted total lung capacity; and the partial pressure of oxygen in arterial blood at the end of maximal exercise. However, it failed to take into account the significance of gender in predicting survival in patients with IPF. Wells et al[13] proposed the composite physiological index CPI, this index only included pulmonary function test results, overlooking radiology findings in predicting prognosis. Du Bois et al[14] 2011 and Richards et al[15] 2012 respectively developed predictive systems based on IPF diagnostic criteria and biomarker predictive models. However, these models have been difficult to use, lack validation, focus predominantly on IPF and have not been widely adopted. The GAP model suggested by Ley et al in 2012[16] which uses four variables; gender, age and two pulmonary physiological parameters, has proved more straight forward to use however a validation study for prognoses of patients with IPF for each GAP score suggested a need to refine the model in terms of groups included in stage 1[11]. Another limitation is its overestimation of risk in lower-risk groups[16]. A modified ILD-GAP index added an ILD subtype variable to the prediction model, both GAP and ILD-GAP have been widely used in the clinical setting to help predict mortality. More recently studies to take account of co-morbidities suggest combining ILD-GAP with the Charlson Comorbidity index score to better predict ILD-related events[17]. GAP6 suggests adding the 6-minute walk test to include functional capacity to the model[18]. These suggested additions to the model have not yet been externally validated.

Recent attention has focused on how monitoring using digital health technologies can be leveraged to understand the course and progression of the disease with evidence that home monitoring trials have shown good adherence[19-28]. Wearable devices can passively collect continuous biometric data, including heart rate[29], blood oxygen levels and objective measures such as activity levels, which may provide insights into global cardio-pulmonary status and longitudinal trends. Pairing this with mobile patient reported symptoms and other clinical measures may provide opportunities to better understand and predict the course of IPF.

Study objectives.

The overarching objective of this study is to develop a rich discovery dataset (deep phenotype) to identify and explore prediction models built on passively collected physiological data and

actively collected patient reported outcome data which can predict short term and long term clinically meaningful outcomes for patients living with IPF. Wearables to collect physiological and activity measures and remote data approaches to collect electronic patient reported outcomes will be used. This real-time data will be combined with clinic data to provide a rich dataset (deep phenotype) for each patient. These data will be used to predict outcomes for ILD patients using both traditional prediction modeling approaches and more recent machine learning based strategies. Patient's actual outcomes will be recorded and compared to predicted outcomes to determine the utility of the prediction models.

Primary objective

The primary objective of this study is to develop prediction models for clinical outcomes by looking at patterns in data types and the significance of those patterns for individual participants. It is envisaged that we will produce three prediction models for the following clinical outcomes for ILD patients:

1. Disease progression prediction model
2. Acute exacerbation prediction model
3. Mortality prediction model

Secondary objective

To produce a prediction model dataset for each participant which could be used as a synthetic arm for an N=1 clinical trial.

Methods and analysis

Design

This will be a prospective, cohort study in patients with ILD using wearable devices (smart watch) and electronic patient reported outcome measures (phone app) combined with patient healthcare record data to predict clinical outcomes.

Study Setting

Patients diagnosed with Idiopathic Pulmonary fibrosis or Progressive Pulmonary Fibrosis who have been referred to St Vincent's University Hospital Interstitial Lung Disease Clinic will be invited to join the study by the specialist team. The specialist team at the study site, is made up of Respiratory Physicians, Nurse Specialists, Respiratory Physiotherapists, Rheumatologists, Radiologists and Pathologists.

Inclusion and exclusion criteria

Patients will be considered eligible for enrolment to this study if they fulfil the inclusion criteria and none of the exclusion criteria, as defined below.

Patient inclusion criteria:

1. Age 18 years or older.
2. Diagnosis of Idiopathic Pulmonary Fibrosis or Progressive Pulmonary Fibrosis.
3. Ability to provide written informed consent.

Patient exclusion criteria:

1. Age under 18 years.
2. Patients who are unwilling to wear a smart watch for the duration of the study.
3. Cognitive impairment or inability to understand and follow instruction which would limit the patient's understanding of the project or the measurement.

Enrolment

Patients will be recruited from ILD outpatient clinics at St Vincent's University Hospital. Participants will be identified as suitable candidates by the specialist team and will be offered information about participation in the study. They will be given an information leaflet and will be directed to the study website (https:// prodigy-ild.ie) for further explanation of the study and what will be expected of them. Once participants have had time to consider if they would like to take part in the study, they will be asked to sign the Informed consent form and enrolled to the study. It is anticipated that recruitment to this study will take place over a six-month period concluding in Quarter 4 2024. Participants will be followed for 3 years.

Sample size.

A target recruitment of 50 participants to this cohort will ensure sufficient data are available for analysis.

Study Procedures

Baseline visit

Once a participant has been on-boarded to the study a research record will be created using demographics, clinical history, diagnostic and disease relevant information. At baseline participants will complete a questionnaire to collect outstanding data on demographic information, medical history, medication and smoking status. Diagnostic radiology, pulmonary function and lab results will be collected directly from the patient record.

Digital health technology

Study data will be collected and managed using REDCap electronic data capture hosted at UCD Clinical Research Centre[30,31]. REDCap (Research Electronic Data Capture) is a secure, web-based software platform designed to support data capture for research studies. As part of enrolment patients will be assigned a unique study identification number, in the REDCap study database. The unique study identification number will be used to name the participants wearable device and to anonymously track their data throughout the study. The study team will provide each participant with a wearable device (Apple watch series 6 or above) and an iPhone series 8 or above if required. Training and on-boarding will take place at the baseline visit. In addition, MyCap, a companion App to RedCap research software, will be used to

collected Patient reported outcome data electronically, to enable comprehensive remote PROM collection over time[32]. The MyCap App will be downloaded to the participants phone and participants will be enrolled to the app by scanning a QR code they are presented with on their study App. Participants will receive a prompt "You have a MyCap activity due today" at 8am on the day a questionnaire is scheduled to be completed. In the case of a non-response to questionnaires participants will be sent a message to respond to the questionnaire. A MyCap on screen prompt will flash up to say, "You have a secure message waiting". When participants open the MyCap app they can check their messages and complete the relevant questionnaire which is due. The questionnaire will not expire therefore follow up messages for outstanding questionnaires can be sent to the participant. If there is still no response, a phone call will be made to ensure participants are aware of when to engage with the study questionnaires. A further escalation if there is still no response to engage with study questionnaires will be to call out to the participant to troubleshoot issues, they are having with the study devices.

Patient Reported Outcome Measures (PROMs)

Participants will be required to complete the following survey instruments at enrollment and at regular intervals, as per Table 1.

Table 1: Schedule of patient questionnaires

PROMS	Domain	Frequency
mMRC - Breathlessness	Dyspnoea	Baseline & Weekly
VAS - cough severity	Cough	Baseline & Weekly
Leicester Cough Questionnaire	Cough	Baseline & Monthly
FAS - Fatigue	Fatigue	Baseline & Monthly
K-BILD -Kings Brief Interstitial Lung Disease	Health Related Quality of Life	Baseline & 3 Monthly

1.Breathlessness Questionnaire – mMRC[33,34], the modified Medical Research Council dyspnoea scale is a self-rating tool used to assess the degree of baseline functional disability due to breathlessness on a scale from 0 to 4. Participants are asked to rate their breathlessness as follows.

- 0, I only get breathless with strenuous exercise.
- 1, I get short of breath when hurrying on level ground or walking up a slight hill.
- 2, On level ground, I walk slower than people of my age because of breathlessness, or I have to stop for breath when walking at my own pace on the level.
- 3, I stop for breath after walking about 100 yards or after a few minutes on level ground.
- 4, I am too breathless to leave the house, or I am breathless when dressing/undressing.

2.*Cough Severity Questionnaire – VAS*[35], A Visual analogue scale for cough severity where the patient indicates the severity of their cough over the last two weeks on a scale of 0 to 100. This score is familiar to clinicians, clinically meaningful and can document longitudinal observations. It has been reported that a ≥ 30-mm reduction in cough severity VAS was estimated as a clinically meaningful change threshold for clinical trials in chronic cough[34].

3.*Cough Quality of Life Questionnaire – LCQ*[36], The Leicester Cough Questionnaire is the gold standard cough questionnaire to assess the impact of cough in the previous two weeks. This 19-item questionnaire comprises three health domains: physical, psychological, and social takes less than 5 minutes to complete. Participants are presented with a 7-point (Likert) response scale and are asked to select the appropriate response, ranging from 1 = all of the time to 7 = none of the time. Domain scores range from 1-7 and total scores range from 3-21, a higher score indicates better health status.

4.*Fatigue Questionnaire – FAS Fatigue*[37,38], The Fatigue assessment scale is a 10-item self-report scale (1-5) evaluating symptoms of chronic fatigue. It measures both physical (5 questions) and mental (5 questions) fatigue. The total FAS score can be calculated by summing the scores on all questions (recoded scores for questions 4 and 10). The total score ranges from 10 to 50. A total FAS score < 22 indicates no fatigue, a score ≥ 22 indicates fatigue. A mental and physical fatigue score is also calculated.

5.*King's Brief Interstitial Lung Disease – KBILD*[39,40], is a self-completed 15-item validated ILD-specific measure of health-related quality of life. Each question has a seven-point Likert response scale. It consists of three domains Physiological (3,5,6,8,12,14), Breathlessness and activities (1,4,11,13) and chest symptoms (2,7,9). It takes 5-7 minutes to complete. The KBILD domain and total score ranges are 0-100; 100 represents best health status.

Wearable Devices

Study participants will be provided with an Apple watch (series 6 or later) wearable device and charging cable. The Apple Watch series 6 pairs with the iPhone series 8 or above and collected data can be viewed in the iPhone Health App. The watch will passively collect continuous real-time data on Heart Rate, blood oxygen levels, activity levels and 6-minute walk distance (Table 2). All data collected by the smart watch automatically synchronizes to the Health App on the participants iPhone when participants have internet connection. The Health App has a function to export all health data collected. This anonymous file will be exported from the participant Health App while participants are attending clinic or during a home visit if participants do not attend a clinic appointment within 6 months of their baseline date. The export file will be transferred to a research MacBook using Apple Airdrop secure file transfer, from there it will be uploaded to the secure UCD research google drive where it will be stored for later analysis.

Participants will be required to wear the study wearable device for at least 20 hours per day and will be responsible to keep the device charged to allow this. Training materials will be provided in hard copy or can be emailed to the participant in electronic format. Participants will be shown how the watch must be worn snug to the wrist, to ensure correct readings are recorded while in use.

Table 2: Apple Watch Continuous Monitoring Data

Variable	Data recorded
Heart Rate (HR) (BPM)	Heart Rate (Beats Per Minute): Apple watch records HR approximately every 5 minutes. Resting Heart Rate: The average heart beats per minute measured when the wearer has been inactive or relaxed for several minutes. Heart Rate Variability: A measure of the variation in the time interval between beats. Apple watch calculates HRV by using standard deviation of beat-to-beat measurements which are captured by the heart rate sensor. Average HRV per day is recorded. Walking Heart Rate Average: the average heart beats per minute measured by Apple Watch during walks at a steady pace throughout the day.
Respiratory Rate (breaths/min)	Apple watch records Respiratory Rate while the wearer is sleeping
Blood Oxygen (SpO2)	Apple watch attempts to record oxygen saturation every 30 minutes during sleep. It will also opportunistically record oxygen saturation during the day while arm is at rest. Also, a participant can opt to do a recording themselves at their convenience.
Six-Minute Walk estimate	A weekly estimate of how far the wearer can walk on flat ground in six minutes based on recent motion and workout data. The Apple Watch can give a predicted six-minute walk distance up to 500 meters.
Step Count (count)	Step count is the number of steps you take throughout the day.
Distance Walking Running (km)	Calculates an average distance you have walked and run over the last 7 days
Exercise minutes	Measure of how many minutes of brisk activity you do
Sleep analysis	Time In Bed Time Asleep (Core, Deep, REM) Time Awake
Wrist Temperature	Wrist temperature is a measurement related to your body's temperature taken by Apple Watch while you are sleeping. Each value is an average of several measurements taken during sleep.

Study procedures

Follow up visit.

Participants follow up visits will coincide with clinic visits every four months. It is therefore anticipated that there will be three patient visits per year. If a participant is unable to attend a clinic appointment a home visit will be scheduled to collect the participants export file. Patients

will be followed for 3 years, giving potentially a total of 9 visits per participant. Routine clinic visits include vital signs, spirometry and six-min walk test checking for desaturation. Additional periodic tests will include diffusion capacity of the lungs for carbon monoxide ($DLCO_2$), high resolution computed tomography (HRCT) and blood samples for biomarker studies.

All patient record data collected at clinic appointments will be captured in REDCap database for integration with MyCap PROMS and wearable datasets in our PRODIGY-ILD data warehouse (Figure 1).

Table 3: Study schedule of events and ILD Outpatient Clinic

Type of Visit	Screening	Baseline	Weekly	Monthly	Every 3 Months	ILD Clinic visits as per standard clinical care	Annually as per standard clinical care
Visit Number	1	2					
Week Number	-2	0					
Inclusion/Exclusion Criteria	✓						
Informed consent	✓	✓					
Medical history	✓						
Medications	✓					✓	
Acute exacerbations	✓					✓	
Spirometry	✓					✓	
DLCO[1]	✓					✓	
Six Minute Walk Test	✓					✓	
Serology	✓	✓					✓
HRCT[2]	✓	✓					✓
Dispensing of Apple watch and iPhone		✓					
Training on PROMS[3]		✓					
Breathlessness (PROMS mMRC))[4]		✓	✓				
Cough (PROMS VAS)[5]		✓	✓				
Cough (PROMS LCQ)[6]		✓		✓			
Fatigue (PROMS FAS)[7]				✓			
HRQoL (PROMS K-BILD)[8]					✓		
Apple Health Export Biometric and activity data						✓	

[1] DLCO: Diffusion capacity of the lungs for carbon monoxide, [2]HRCT: High Resolution Computed Topography, [3]PROMS: Patient reported outcome measures, [4]mMRC: Modified Medical Research Council dyspnoea scale, [5]VAS: Visual analogue scale for cough severity, [6]LCQ: Leicester Cough Questionnaire,[7]FAS: Fatigue Assessment Scale,[8]K-BILD: King's Brief Interstitial Lung Disease.

Clinical outcomes (disease progression, onset of acute exacerbations and mortality) will be assessed at clinic appointments, through changes in imaging (Computed Tomography (CT) of the chest), pulmonary function tests, 6-minute walk test distance and patient symptoms. Follow up data will be captured in the patient's research record in REDCap.

Data analysis

The overall objective of this study is to leverage the data collected to generate predictive models of outcome. Specifically, we will seek to develop individual patient level models, of

1. Disease progression
2. Acute exacerbation
3. Mortality

Models will be developed using dynamic prediction in addition to machine learning approaches.

Six-month patient data will be collected and analysed. Outcomes predicted will be compared with actual outcomes recorded at month twelve, allowing refinement, enhancement, and validation of the developed models. Data analysis will proceed as follows:

Step 1: Data quality control and validation:

Individual patient level data will be reviewed for completeness. Given the lack of previous studies integrating clinical, physiological, activity and patient reported data in this population, we will employ a conservative approach to data completeness with a requirement for at least 70% completeness for key variables. Missing data will be summarised and dealt with through Case deletion, where the 70% threshold is not reached. Rational substitution will be employed where possible. For missing data at random, arithmetic imputation methods will be employed including worst case imputation for dropouts and interpolation/extrapolation where prior and after data is available.

Step 2: Descriptive analysis: Data from activity and physiological measurements will be summarised as per Table. 4. Data summarisation will involve providing a concise overview of key characteristics of the dataset including central tendency values, variance to give insights into the spread of the data and frequency distributions, assessing for normality, while highlighting any notable patterns or peaks in the dataset.

Step 3: Feature selection, potential predictors of progression, exacerbation and mortality will be reviewed by specialist respiratory clinicians. The rationale behind the feature selection process will be clearly documented and the chosen predictors will be rigorously validated to ensure clinical relevance to the target outcomes. In parallel, activity and physiological measurements (individual and composite, single point and time-rolling average trends, raw values and individual-normalised) will be analysed to identify potential features, predictive of clinical outcome.

Step 4: Predicted outcomes, the prediction model will be simplified by the elimination of unnecessary variables. Methods such as correlation analysis will be performed to identify potential redundancies. Statistical tests will be used to rank features and their association with the target outcomes.

Step 5: Validation: we will evaluate the predictive model's effectiveness by comparing its performance to actual events per predictor both within individual and across individuals. By

doing this we will gain insights into the practical relevance of individual predictors, and we will refine the model accordingly.

Prediction models for clinical outcomes, acute exacerbations, progression, and mortality will be developed. Traditional regression analysis along with dynamic prediction methods will be employed to develop a series of models whose predictive performance will be compared to actual recorded outcomes, allowing performance of prediction models to be evaluated.

Table 4: Baseline characteristics from wearable

Variable	Min	Max	Mean (Standard Deviation)	Coefficient of variation	Weekly Δ max - min	Days with data	Hours worn	Number of recordings
Oxygen Saturation (%)								
Heart Rate (count/min)								
Heart Rate Variability (ms)								
Resting Heart Rate (count/min)								
Walking Heart Rate Average (count/min)								
Respiratory Rate (count/min)								
StepCount (count)								
Distance Walking Running (km)								
Six Minute Walk Test Distance (m)								
Sleep Analysis								
Wrist Temperature								

Patient and Public involvement

This study has been designed together with input from patients and patient advocates. Specifically, the Irish Lung Fibrosis patient and public involvement (PPI) Research Advisory Group and their volunteers have provided input into the relevance of wearables and research for this population, as well as recruitment and onboarding procedures to ensure these are designed with the end users in mind. In addition, we have co-developed the patient facing resources (including the website) to ensure their relevance, comprehensibility, and accessibility for this patient population.

Discussion

The primary objective of this study is to develop robust prediction models for clinical outcomes in ILD using a combination of clinical, physiological, activity and patient reported data fields. Through the comprehensive collection and analysis of these data sets, we will identify better predictors of disease progression, acute exacerbations, and mortality. Predictions will be compared to actual outcomes to validate prediction models. These predictions will be of significant value in clinical management by providing clinicians with an improved tool for

prognostication. Critically these models will be utilized as synthetic controls for clinical trials in these rare diseases, thereby enabling additional investigation of new therapeutic strategies.

The strength of this study lies in the ability of the wearable to passively collect a large quantity of data without burdening the patient to record active measures. Patients in advanced disease states may not wish to be reminded of declining health measures and are not required to look at their health data as it is collected passively.

This study method could be successfully applied to other disease groups. This could be particularly useful for diseases in which clinical course varies by individual and rare disease populations where it is difficult to recruit large numbers for clinical trials of new treatments.

Declarations

Acknowledgements

Health Research Board – Trials Methodology Research Network (HRB-TMRN), Galway, Ireland and the UCD Clinical Research Centre for their support of this study. Lindsay Brown Advanced Nurse Practitioner for assistance with PROMS selection. Dylan Keagan PIL/ICF. Irish Lung Fibrosis Association PPI Nicola Cassidy, Paula Jenkins, and Robert Hurley for their input to the design of the Website and Patrick McKay, Advantage Point Creative Consultancy for Website design.

Author Contributions

Conception and design of the study: PD, MK, CMcC, ANF, SH, EG. Drafting the manuscript: SH, EG. Revising the manuscript for important intellectual content: PD, MK, CMcC, ANFSH, EG. Approval of the version of the manuscript to be published: PD, CMcC, MK.

Funding

The grant support for this research study was provided by the Health Research Board – Trials Methodology Research Network (HRB-TMRN), The funding body had no role in the design of this study protocol, and will not be involved in the collection, analysis, and interpretation of data or manuscript preparation.

Availability of data and materials

Grouped data generated or analysed during this study will be published. No individual data

will be made available to protect the recognition of individual participants.

Ethics approval and consent to participate.

Full ethics approval was granted for this research study by the St. Vincent's University Hospital Research Ethics Committee, Dublin, Ireland; Ref no: RS23-023. Written informed consent will be obtained from all participants.

References

1. Kumar, Abbas, Fausto, Mitchell. Robbins Basic Pathology, Chapter 13 The Lung, Diffuse Interstitial Lung Diseases, pg. 792 8th Edition Elsevier 2007.
2. King T.E, Flaherty K.R and Dieffenbach P. Approach to the adult with interstitial lung disease: Clinical evaluation. UpToDate Wolters Kluwer Literature review current through Feb 2023. Accessed March 2023.
3. Raghu G et al. Idiopathic Pulmonary Fibrosis (an Update) and Progressive Pulmonary Fibrosis in Adults: An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. Am J Respir Crit Care Med. 2022 May 1; 205(9): e18-e47. Published online 2022 May1. Doi:10.1164/rccm.202202-0399ST
4. Salciccioli J.D, Marshall D.C, Goodall R, Crowley C et al. Interstitial lung disease incidence and mortality in the UK and the European Union: an observational study, 2001-2017.ERJ Open Research 2022 8: 00058-2022; DOI: 10.1183/23120541.00058-2022
5. Irish Thoracic Society Position statement on the management of Idiopathic Pulmonary Fibrosis. The Irish Thoracic Society Interstitial Lung Disease Group. August 2018 https://irishthoracicsociety.com/wp-content/uploads/2018/08/ITS-Pulmonary-Fibrosis-Position-Statement.pdf
6. Raghu G, Collard HR, Egan JJ, et al. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. Am J Respir Crit Care Med 2011; 183: 788–824.
7. American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. international consensus statement. American Thoracic Society (ATS), and the European Respiratory Society (ERS). *Am J Respir Crit Care Med.* 2000;161(2 Pt 1):646-664. doi: 10.1164/ajrccm.161.2.ats3-00.
8. Irish Lung Fibrosis Association. Available at https://ilfa.ie/ Accessed on the 8th of December 2022
9. Collard HR et al. Acute Exacerbation of Idiopathic Pulmonary Fibrosis. An International Working Group Report. *Am J Respir Crit Care Med.* 2016 Aug 1;194(3):265-75. doi: 10.1164/rccm.201604-0801CI
10. Kim JH., Perlman D. and Tomic R. Natural history of pulmonary fibrosis. Respiratory Medicine (2015) 109, 661-670. http://dx.doi.org/10.1016/j.rmed.2015.02.002

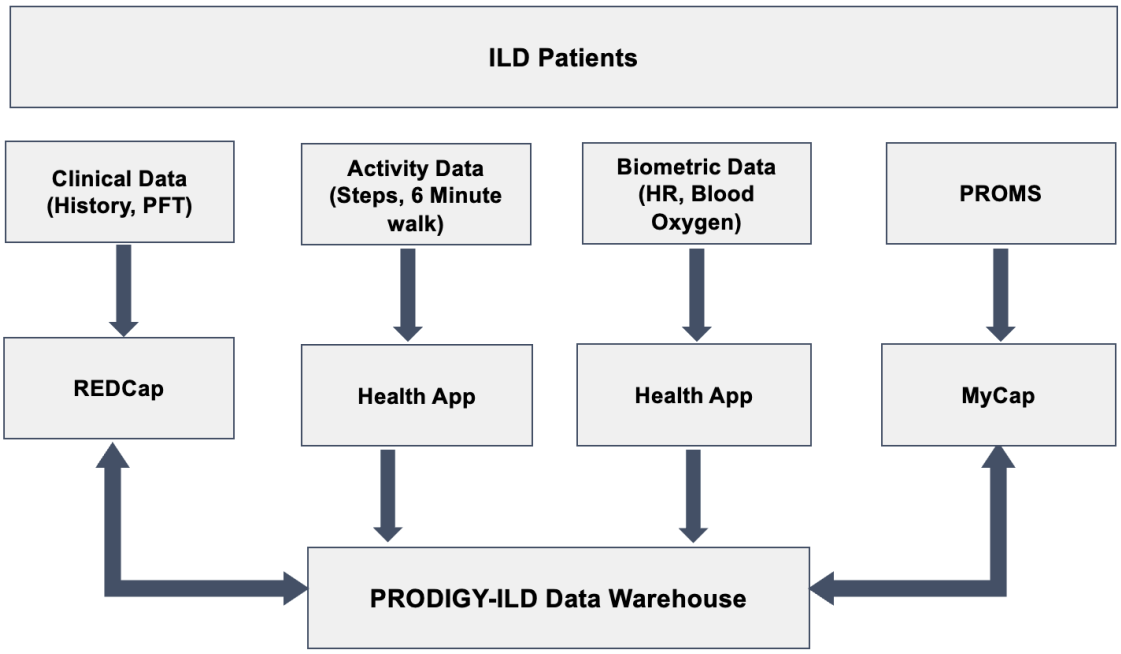
11. Lee SH, Kim SY, Kim DS et al. Predicting survival of patients with idiopathic pulmonary fibrosis using GAP score: a nationwide cohort study. Respiratory Research 2016; 17:131 doi: 10.1186/s12931-016-0454-0
12. King TE, Tooze JA, Schwarz MI, Brown KR, Cherniack RM. Predicting survival in idiopathic pulmonary fibrosis: Scoring system and survival model. Am J Respir Crit Care Med. 2001; 164:1171 – 81.
13. Wells AU, Desai SR, Rubens MB, Goh NSL, Cramer D, Nicholson AG, Colby TV, du Bois RM, Hansell DM. Idiopathic pulmonary fibrosis - A composite physiologic index derived from disease extent observed by computed tomography. Am J Respir Crit Care Med. 2003; 167:962–9.
14. du Bois RM, Weycker D, Albera C, Bradford WZ, Costabel U, Kartashov A, Lancaster L, Noble PW, Raghu G, Sahn SA, et al. Ascertainment of individual risk of mortality for patients with idiopathic pulmonary fibrosis. Am J Respir Crit Care Med. 2011; 184:459–66.
15. Richards TJ, Kaminski N, Baribaud F, Flavin S, Brodmerkel C, Horowitz D, Li K, Choi J, Vuga LJ, Lindell KO, et al. Peripheral blood proteins predict mortality in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med. 2012; 185:67–76.
16. Ley B, Ryerson CJ, Vittinghoff E, Ryu JH, Tomassetti S, Lee JS, Poletti V, Buccioli M, Elicker BM, Jones KD, et al. A Multidimensional Index and Staging System for Idiopathic Pulmonary Fibrosis. Ann Intern Med. 2012; 156:684–U658.
17. Fujii H, Hara Y, Saigusa Y, Tagami Y, et al.ILD-GAP Combined with the Charlson Comorbidity Index Score (ILD-GAPC) as a Prognostic Prediction Model in Patients with Interstitial Lung Disease. Can Respir J. 2023; 2023: 5088207.Published online 2023 Feb 8. Doi: 10.1155/2023/5088207
18. Lee J.H., Jang J.H., Jang H-J., Kim S.Y, et al. New prognostic scoring system for mortality in idiopathic pulmonary fibrosis by modifying the gender, age, and physiology model with desaturation during the six-minute walk test. Front Med (Lausanne). 2023 Jan 25:10:1052129. Doi:10.3389/fmed.2023.1052129
19. Althobiani M, et al. Variables used to monitor patients with Interstitial Lung Disease. BMJ Open Resp Res 2021; 8: e001088
20. Althobiani M., Alqahtani JS, Hurst JR, Russell AM, and Porter J. Telehealth for patients with interstitial lung diseases (ILD): results of an international survey of clinicians.
21. Althobiani M, Evans R, Alqahtani J, et al. Home monitoring of physiology and symptoms to detect interstitial lung disease exacerbations and progression: a systematic review. ERJ Open Res. 2021 Dec 20;7(4): 00441-2021. Doi:101183/23120541.00441-2021
22. Catharina C Moor, Rémy L M Mostard, Jan C Grutters, Paul Bresser, Joachim G J V Aerts, Niels H Chavannes, Marlies S Wijsenbeek. Home Monitoring in Patients with Idiopathic Pulmonary Fibrosis. A Randomized Controlled TrialAm J Respir Crit Care Med 2020 Aug 1;202(3):393-401. Doi: 10.1164/rccm.202002-0328OC.
23. Imre Noth, Vincent Cokn, Nazia Chaudhuri, Tamera J Corte, Kerri A Johannson, Marlies Wijsenbeek, Stephane Jouneau, Andreas Michael, Manuel Quaresma, Klaus B Rohr, Anne-Marie Russell, Susanne Stowasser, Toby M Maher, INMARK trial investigators. Home

spirometry in patients with idiopathic pulmonary fibrosis: data from the INMARK trial Eur Respir J. 2021 Jul 8;58(1): 2001518. Doi:10.1183/13993003.01518-2020. Print 2021 Jul

24. Marlies S Wijsenbeek, Elisabeth Bendstrup, Claudia Valenzuela, Michael T Henry, Catharina C Moor, Stéphane Jouneau, Alessandro G Fois, Onofre Moran- Mendoza, Syed Anees, Mirela Mirt, Monica Bengus, Frank Gilberg, Klaus-Uwe Kirchgaessler, Carlo Vancheri. Disease Behaviour During the Peri-Diagnostic Period in Patients with Suspected Interstital Lung Disease: The STARLINER Study Adv Ther. 2021 Jul;38(7):4040-4056. doi: 10.1007/s12325-021-01790-y. Epub 2021 Jun 11.
25. S. Walsh, T. Cahill, C. Edwards, E. Costello, J. Walsh, A.-M. Russell, A.W. O'Regan. PaGent-Reported Monitoring of Symptoms and Spirometry Via the patientMpower Plaform in Idiopathic Pulmonary Fibrosis. American Journal of Respiratory and CriGcal Care Medicine 2018; 197: A4933
26. Wim A. Wuyts, * Marlies Wijsenbeek, Benjamin Bondue, Demosthenes Bouros, Paul Bresser, Carlos Robalo Cordeiro, Ole Hilberg, Jesper Magnusson, Effrosyni D. Manali, António Morais, Spyridon Papiris, Saher Shaker, Marcel Veltkamp, and Elisabeth Bendstrupm. Idiopathic Pulmonary Fibrosis: Best Practice in Monitoring and Managing a Relentless Fibrotic Disease. Respiration. 2020 Jan; 99(1): 73–82. Published online 2019 Dec12. Doi: 10.1159/000504763
27. Toby M Maher 1 2, Courtney Schiffman 3, Michael Kreuter 4 5, Catharina C Moor 6, Steven D Nathan 7, Judit Axmann 8, Paula Belloni 3, Monica Bengus 8, Frank Gilberg 8, Klaus-Uwe Kirchgaessler 8, Marlies S Wijsenbeek 6. A review of the challenges, learnings, and future directions of home handheld spirometry in interstitial lung disease. Respir Res 2022 Nov 11;23(1):307. Doi: 10.1186/ s12931-022-02221-4.
28. Marlies S Wijsenbeek, Catharina C Moor, Kerri A Johannson, Peter D Jackson, Yet H Khor, Yasuhiro Kondoh, Sujeet K Rajan, Gabriela C Tabaj, Brenda E Varela, Pieter van der Wal, Richard N van Zyl-Smit, Michael Kreuter, Toby M Maher. Home monitoring in interstitial lung diseases. Lancet Respir Med. 2022 Oct 4; Doi:10.1016/ S2213-2600(22)00228-4
29. Apple Heart Study Protocol: Evaluation of the use of Apple Watch faeatures for identification of cardiac arrhythmias. NCT03769207 Available at https://Clinicaltrials.gov Accessed March 2023.
30. PA Harris, R Taylor, R Thielke, J Payne, N Gonzalez, JG. Conde, Research electronic data capture (REDCap) – A metadata-driven methodology and workflow process for providing translational research informatics support, J Biomed Inform. 2009 Apr;42(2):377-81.
31. PA Harris, R Taylor, BL Minor, V Elliott, M Fernandez, L O'Neal, L McLeod, G Delacqua, F Delacqua, J Kirby, SN Duda, REDCap Consortium, The REDCap consortium: Building an international community of software partners, J Biomed Inform. 2019 May 9 [doi: 10.1016/j.jbi.2019.103208]
32. PA Harris, J Swafford, E S Serdoz, J Eidenmuller, G Delacqua, V Jagtap, R J Taylor, A Gelbard, A C Cheng, S N Duda, "MyCap: a flexible and configurable platform for mobilizing the participant voice", JAMIA Open, Volume 5, Issue 2, July 2022, ooac047. https://doi.org/10.1093/jamiaopen/ooac047

33. Mahler DA, Wells CK. Evaluation of clinical methods for rating dyspnea. Chest 1988; 93: 580–586. [PubMed] [Google Scholar]
34. Bestall JC, Paul EA, Garrod R, et al. Usefulness of the medical research council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. Thorax 1999; 54: 581–586. [PMC free article] [PubMed] [Google Scholar]
35. Martin Nguyen A, Bacci ED, Vernon M, et al. Validation of a visual analog scale for assessing cough severity in patients with chronic cough. Ther Adv Respir Dis 2021 Oct 26. doi: 10.1177/17534666211049743.
36. Birring S, Prudon B, Carr A, Singh S, Morgan M, Pavord I. Development of a symptom specifc health status measure for patients with chronic cough: Leicester Cough Questionnaire (LCQ). Thorax 2003; 58:339–43
37. FAS: Fatigue Assessment Scale and how to use FAS. Accessed June 2023, Available at: https://www.ildcare.nl/index.php/condition-use-the-fas-fatigue-assessment-scale/ https://www.ildcare.nl/index.php/how-to-use-the-fas-fatigue-assessment-scale/
38. Hendriks C, Drent M, Elfferich M, De Vries J. The Fatigue Assessment Scale (FAS): quality and availability in sarcoidosis and other diseases. Curr Opin Pulm Med 2018; 24 (5): 495-503. https://pubmed.ncbi.nlm.nih.gov/29889115
39. Patel AS, Siegert RJ, Brignall K, et al. The development and validation of the King's Brief Interstitial Lung Disease (K-BILD) health status questionnaire. BMJ Thorax 2011 doi: 10.1136/thoraxjnl-2012-201581
40. Claire M. Nolan, Surinder S. Birring, Matthew Maddocks, Toby M. Maher, Suhani Patel, Ruth E. Barker, Sarah E. Jones, Jessica A. Walsh, Stephanie C. Wynne, Peter M. George, William D-C. Man. King's Brief Interstitial Lung Disease questionnaire: responsiveness and minimum clinically important difference. European Respiratory Journal 2019, 54:1900281; DOI: 10.1183/13993003.00281-201

Figure 1: Data Warehouse of data collected from wearable, app, and patient record.



BMJ Open

PRODIGY-ILD: A protocol for a prospective cohort study to predict outcomes in patients with interstitial lung disease using digital health technologies.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2024-088271.R1
Article Type:	Protocol
Date Submitted by the Author:	19-Feb-2025
Complete List of Authors:	Gunne, Emer; University College Dublin, School of Medicine Holden, Sinead; University College Dublin, Franciosi, Alessandro; St Vincent's University Hospital, Respiratory Medicine Keane, Michael; University College Dublin, McCarthy, Cormac; University College Dublin, School of Medicine; Saint Vincent's University Hospital, Doran, Peter; University College Dublin School of Medicine and Medical Science, School of Medicine
Primary Subject Heading:	Respiratory medicine
Secondary Subject Heading:	Research methods, Health informatics
Keywords:	Interstitial lung disease < THORACIC MEDICINE, Observational Study, Patient Reported Outcome Measures, Wearable Electronic Devices, Machine Learning, Digital Technology

Title:

PRODIGY-ILD: A protocol for a prospective cohort study to predict outcomes in patients with interstitial lung disease using digital health technologies.

Corresponding Author:
Emer Gunne[1]
Email: emer.gunne@ucdconnect.ie
ORCID-ID 0000-0002-6939-1586

Other Authors:
Sinead Holden[2],
ORCID-ID 0000-0002-7314-2152

Alessandro N Franciosi[3],
ORCID-ID 0000-0002-4241-8718

Michael P. Keane[1,3],
ORCID-ID 0000-0003-4507-0742

Cormac McCarthy[1,3],
ORCID-ID 0000-0003-2896-5210

Peter Doran[1,3]
ORCID-ID 0000-0003-2064-5335

Affiliations:

1. University College Dublin, School of Medicine, Dublin, Ireland
2. University College Dublin, School of Public Health, Physiotherapy & Sports Science, Dublin, Ireland
3. St. Vincent University Hospital, Elm Park, Dublin, Ireland

Word count:
Abstract: words 300
Main document text: 3,529 words

Tables & Figures:
1 numbered Figure
3 numbered Tables

ABSTRACT:

Introduction

Interstitial lung disease (ILD) patients may develop a progressive phenotype usually characterised by progressive pulmonary fibrosis. Whilst this condition is life-limiting, wide variations in its clinical course have made it difficult to predict rate of disease progression, onset of acute exacerbations and mortality. New approaches are needed to predict the clinical course of ILD, to enable both treatment planning, evaluation and clinical trial design

Advances in digital health technologies has facilitated the ability to collect "real-time" data to monitor diseases. These data, including physiologic measures, activity indices and patient reported outcomes may be useful as components of new outcome predictors. The objective of this study is to firstly deploy comprehensive data collection enabling deep profiling of patients with ILD and to use these data to develop better predictors of outcome. Finally, these predictions will be evaluated based on real observed outcomes for individual patients.

Methods and analysis

This study is a prospective cohort study with 50 participants.

Inclusion criteria: Age 18 years or older with a diagnosis of ILD and the ability to provide written informed consent.

Exclusion criteria: Age under 18 years or unwilling to wear a smartwatch for the duration of the study.

Participants will be provided with a smartwatch to passively collect biometric data. This data will be combined with clinical history and course, in addition to a set of patient reported outcome measures. Participants will be followed for 3 years to assess rate of disease progression, occurrence of acute exacerbations and mortality. Initial data will be used to develop clinical prediction models. These models will be further evaluated for accuracy using regular follow up data.

Ethics and dissemination

Approved by the St. Vincent's University Hospital Research Ethics Committee, Dublin, Ireland; Reference number: RS23-023. Results will be presented at medical conferences and disseminated via peer-reviewed journals.

ARTICLE SUMMARY

Strengths and limitations of this study

- Real-time passively collected data will give deeper insights to patient physiological measures, while not over burdening the participant to provide active measures.

- Study devices, training on device use and ongoing support will be provided, overcoming potential sources of bias and barriers to participation in terms of access to technology and tech literacy.
- Interstitial lung disease is a rare disease therefore our sample size is small at 50 participants, which may limit the robustness of prediction modeling. However, we will collect a large volume of data points and anticipate a high event rate in this population over the three-year follow-up period.
- By using the rule of thumb of 10 events per variable, we expect to have sufficient information for the development of the model; though we acknowledge a larger dataset and patient population will likely be required for the further validation of the predictor.
- Interstitial Lung Disease Outpatient Clinic attendees in Ireland are not an ethnic diverse population, which may limit the generalizability of the results.

MAIN BODY

Introduction

Interstitial lung disease (ILD) is a heterogeneous group (over 200 disorders) of non-infectious, predominantly diffuse and usually chronic respiratory disorders[1]. The disease affects the interstitium as well as the alveolar and airway architecture[2]. Some interstitial lung diseases are characterised by progressive pulmonary fibrosis, such as in Idiopathic Pulmonary Fibrosis (IPF) and Progressive Pulmonary Fibrosis (PPF)[3]. IPF is a chronic, fibrosing interstitial pneumonia of unknown cause, associated with radiological and histologic features of usual interstitial pneumonia[3].

In 2017, the incidence of ILD in Ireland was 7.66 per 100,000 population for men and 4.2 per 100,000 for women[4]. Mortality rates were 5.18 per 100,000 and 2.73 per 100,000 respectively[4], this is in line with the median incidence in Europe in 2017[4]. However, these figures could be underestimated, as the British Lung Foundation have reported 6,000 people diagnosed with IPF annually. IPF occurs primarily in the older adult population, with onset of disease in the sixth or seventh decade[6]. More men have been reported with IPF than women, and most patients have a history of past cigarette smoking[7]. Specialist care for ILD patients is delivered by respiratory multidisciplinary care teams at 8 clinical centres across Ireland.[8]

Clinical manifestations include chronic exertional dyspnea and cough. The disease is characterised by progressive worsening of dyspnoea and lung function, progressive fibrosis on high resolution computed topography, acute respiratory decline with a median survival of 3-5 years[2,3]. Patients with IPF are susceptible to abrupt declines in lung function (acute exacerbations) which typically develop in less than 1 month, are accompanied by new

radiologic abnormalities on high resolution computed topography[9], and are a major indicator of morbidity and mortality, with acute exacerbation preceding approximately 40 percent of IPF deaths. Median survival following an acute exacerbation is approximately three to four months[9]. A significant challenge is the fact that IPF and other ILDs often follow an unpredictable clinical course, which impedes physicians' ability to predict rate of disease progression, acute exacerbations, and survival[10]. This impacts new therapeutic development, as the absence of clear signals of progression impacts on trial design. Given ILDs are rare diseases, clinical investigations require additional precision.

Numerous efforts have been made to develop reliable prediction models for ILD patients[11]. The upgraded CRP (Clinical, Radiological, and physiological[12]) scoring system incorporates several parameters: age, smoking history; clubbing; extent of profusion of interstitial opacities, and presence or absence of pulmonary hypertension on the chest radiograph; percent predicted total lung capacity; and the partial pressure of oxygen in arterial blood at the end of maximal exercise. However, it failed to take into account the significance of gender in predicting survival in patients with IPF. Wells et al[13] proposed the composite physiological index CPI, this index only included pulmonary function test results, overlooking radiology findings in predicting prognosis. Du Bois et al[14] 2011 and Richards et al[15] 2012 respectively developed predictive systems based on IPF diagnostic criteria and biomarker predictive models. However, these models have been difficult to use, lack validation, focus predominantly on IPF and have not been widely adopted. The GAP model suggested by Ley et al in 2012[16] which uses four variables; gender, age and two pulmonary physiological parameters, has proved more straight forward to use however a validation study for prognoses of patients with IPF for each GAP score suggested a need to refine the model in terms of groups included in stage 1[11]. Another limitation is its overestimation of risk in lower-risk groups[16]. A modified ILD-GAP index added an ILD subtype variable to the prediction model, both GAP and ILD-GAP have been widely used in the clinical setting to help predict mortality. More recently studies to take account of co-morbidities suggest combining ILD-GAP with the Charlson Comorbidity index score to better predict ILD-related events[17]. GAP6 suggests adding the 6-minute walk test to include functional capacity to the model[18]. These suggested additions to the model have not yet been externally validated.

Recent attention has focused on how monitoring using digital health technologies can be leveraged to understand the course and progression of the disease with evidence that home monitoring trials have shown good adherence[19-28]. Wearable devices can passively collect continuous biometric data, including heart rate[29], blood oxygen levels and objective measures such as activity levels, which may provide insights into global cardio-pulmonary status and

longitudinal trends. Pairing this with mobile patient reported symptoms and other clinical measures may provide opportunities to better understand and predict the course of IPF.

Study objectives.

The overarching objective of this study is to develop a rich discovery dataset (deep phenotype) to identify and explore prediction models built on passively collected physiological data and actively collected patient reported outcome data which can predict short term and long term clinically meaningful outcomes for patients living with IPF. Wearables to collect physiological and activity measures and remote data approaches to collect electronic patient reported outcomes will be used. This real-time data will be combined with clinic data to provide a rich dataset (deep phenotype) for each patient. These data will be used to predict outcomes for ILD patients using both traditional prediction modeling approaches and more recent machine learning based strategies. Patient's actual outcomes will be recorded and compared to predicted outcomes to determine the utility of the prediction models.

Primary objective

The primary objective of this study is to develop prediction models for clinical outcomes by looking at patterns in data types and the significance of those patterns for individual participants. It is envisaged that we will produce three prediction models for the following clinical outcomes for ILD patients:

1. Disease progression prediction model
2. Acute exacerbation prediction model
3. Mortality prediction model

Secondary objective

To produce a prediction model dataset for each participant which could be used as a synthetic arm for an N=1 clinical trial.

Methods and analysis

Design

This will be a prospective, cohort study in patients with ILD using wearable devices (smart watch) and electronic patient reported outcome measures (phone app) combined with patient healthcare record data to predict clinical outcomes.

Study Setting

Patients diagnosed with Idiopathic Pulmonary fibrosis or Progressive Pulmonary Fibrosis who have been referred to St Vincent's University Hospital Interstitial Lung Disease Clinic will be invited to join the study by the specialist team. The specialist team at the study site, is made up of Respiratory Physicians, Nurse Specialists, Respiratory Physiotherapists, Rheumatologists, Radiologists and Pathologists.

Inclusion and exclusion criteria

Patients will be considered eligible for enrolment to this study if they fulfil the inclusion criteria and none of the exclusion criteria, as defined below.

Patient inclusion criteria:

1. Age 18 years or older.
2. Diagnosis of Idiopathic Pulmonary Fibrosis or Progressive Pulmonary Fibrosis.
3. Ability to provide written informed consent.

Patient exclusion criteria:

1. Age under 18 years.
2. Patients who are unwilling to wear a smart watch for the duration of the study.
3. Cognitive impairment or inability to understand and follow instruction which would limit the patient's understanding of the project or the measurement.

Enrolment

Patients will be recruited from ILD outpatient clinics at St Vincent's University Hospital. Participants will be identified as suitable candidates by the specialist team and will be offered information about participation in the study. They will be given an information leaflet and will be directed to the study website (https:// prodigy-ild.ie) for further explanation of the study and what will be expected of them. Once participants have had time to consider if they would like to take part in the study, they will be asked to sign the Informed consent form (supplementary material) and enrolled to the study. It is anticipated that recruitment to this study will take place over a six-month period concluding in Quarter 4 2024. Participants will be followed for 3 years.

Sample size.

The primary objective of this study is to develop prediction models for clinical outcomes (disease progression, acute exacerbations and mortality) by looking at patterns in data types and the significance of those patterns for individual participants. The development of a prediction model requires firstly a developmental dataset containing the likely predictor values, which are used to develop the prediction model. The sample size for the development data set must be sufficiently large to enable a model to be developed which can subsequently be tested. In this study we are utilizing the 10 events per variable rule of thumb to estimate the sample size for the developmental dataset (Peduzzi et al[30,31]). Using this approach, we estimate that 50 participants will provide sufficient events to enable the development of the prediction modeling. It should be noted that given the progressive nature of this disease, and the proposed 3 year follow up, we anticipate that all patients will progress with exacerbations, hospitalisations and death occurring within the study time frame, in most participants. Depending on the predictor variable ultimately uitilised, a larger

sample size for the validation set will be enrolled. The sample size was determined based on the need to achieve reliable and generalizable results for this objective. Sampling for rare diseases is inherently challenging due to limited patient populations. Given the prevalence of ILD (rare disease) a sample size of 50 is feasible within the available time frame and resources.

Study Procedures

Baseline visit

Once a participant has been on-boarded to the study a research record will be created using demographics, clinical history, diagnostic and disease relevant information. At baseline participants will complete a questionnaire to collect outstanding data on demographic information, medical history, medication and smoking status. Diagnostic radiology, pulmonary function and lab results will be collected directly from the patient record.

Digital health technology

Study data will be collected and managed using REDCap electronic data capture hosted at UCD Clinical Research Centre[32,33]. REDCap (Research Electronic Data Capture) is a secure, web-based software platform designed to support data capture for research studies. As part of enrolment patients will be assigned a unique study identification number, in the REDCap study database. The unique study identification number will be used to name the participants wearable device and to anonymously track their data throughout the study. The study team will provide each participant with a wearable device (Apple watch series 6 or above) and an iPhone series 8 or above if required. Training and on-boarding will take place at the baseline visit. In addition, MyCap, a companion App to RedCap research software, will be used to collected Patient reported outcome measure (PROM) data electronically, to enable comprehensive remote PROM collection over time[34]. The MyCap App will be downloaded to the participants phone and participants will be enrolled to the app by scanning a QR code they are presented with on their study App. Participants will receive a prompt “You have a MyCap activity due today” at 8am on the day a questionnaire is scheduled to be completed. In the case of a non-response to questionnaires participants will be sent a message to respond to the questionnaire. A MyCap on screen prompt will flash up to say, “You have a secure message waiting”. When participants open the MyCap app they can check their messages and complete the relevant questionnaire which is due. The questionnaire will not expire therefore follow up messages for outstanding questionnaires can be sent to the participant. If there is still no response, a phone call will be made to ensure participants are aware of when to engage with the study questionnaires. A further escalation if there is still no response to engage with study questionnaires will be to call out to the participant to troubleshoot issues,

they are having with the study devices.

Patient Reported Outcome Measures (PROMs)

Participants will be required to complete the following survey instruments at enrollment and at regular intervals, as per Table 1.

Table 1: Schedule of patient questionnaires

PROMS	Domain	Frequency
mMRC - Breathlessness	Dyspnoea	Baseline & Weekly
VAS - cough severity	Cough	Baseline & Weekly
Leicester Cough Questionnaire	Cough	Baseline & Monthly
FAS - Fatigue	Fatigue	Baseline & Monthly
K-BILD -Kings Brief Interstitial Lung Disease	Health Related Quality of Life	Baseline & 3 Monthly

1.*Breathlessness Questionnaire* – modified Medical Research Council dyspnoea scale *mMRC*[35,36], a self-rating tool used to assess the degree of baseline functional disability due to breathlessness on a scale from 0 to 4.

2.*Cough Severity Questionnaire – VAS*[37], A Visual analogue scale for cough severity where the patient indicates the severity of their cough over the last two weeks on a scale of 0 to 100. A ≥ 30-mm reduction in cough is considered a clinically meaningful change threshold for clinical trials in chronic cough[36].

3.*Cough Quality of Life Questionnaire – The Leicester Cough Questionnaire (LCQ)*[38], is a 19-item cough questionnaire comprising three health domains: physical, psychological, and social to assess the impact of cough in the previous two weeks which takes less than 5 minutes to complete.

4.*Fatigue Questionnaire – Fatigue Assessment Scale (FAS)*[39,40], FAS is a 10-item self-report scale (1-5) evaluating symptoms of physical and mental fatigue. A total FAS score < 22 indicates no fatigue, a score ≥ 22 indicates fatigue.

5.*King's Brief Interstitial Lung Disease – KBILD*[41,42], is a self-completed 15-item validated ILD-specific measure of health-related quality of life consists of three domains Physiological (3,5,6,8,12,14), Breathlessness and activities (1,4,11,13) and chest symptoms (2,7,9). The KBILD domain and total score ranges are 0-100; 100 represents best health status.

Wearable Devices

Study participants will be provided with an Apple watch (series 6 or later) wearable device and charging cable. The Apple Watch series 6 pairs with the iPhone series 8 or above and collected data can be viewed in the iPhone Health App. The watch will passively collect continuous real-time data on Heart Rate, blood oxygen levels, activity levels and 6-minute

walk distance (Table 1 supplementary material). All data collected by the smart watch automatically synchronizes to the Health App on the participants iPhone when participants have internet connection. The Health App has a function to export all health data collected. This anonymous file will be exported from the participant Health App while participants are attending clinic or during a home visit if participants do not attend a clinic appointment within 6 months of their baseline date. The export file will be transferred to a research MacBook using Apple Airdrop secure file transfer, from there it will be uploaded to the secure UCD research google drive where it will be stored for later analysis.

Participants will be required to wear the study wearable device for at least 20 hours per day and will be responsible to keep the device charged to allow this. Training materials will be provided in hard copy or can be emailed to the participant in electronic format. Participants will be shown how the watch must be worn snug to the wrist, to ensure correct readings are recorded while in use.

Study procedures

Follow up visit.

Participants follow up visits will coincide with clinic visits every four months. It is therefore anticipated that there will be three patient visits per year. If a participant is unable to attend a clinic appointment a home visit will be scheduled to collect the participants export file. Patients will be followed for 3 years, giving potentially a total of 9 visits per participant. Routine clinic visits include vital signs, spirometry and six-min walk test checking for desaturation. Additional periodic tests will include diffusion capacity of the lungs for carbon monoxide ($DLCO_2$), high resolution computed tomography (HRCT) and blood samples for biomarker studies.

All patient record data collected at clinic appointments will be captured in REDCap database for integration with MyCap PROMS and wearable datasets in our PRODIGY-ILD data warehouse (Figure 1).

Table 2: Study schedule of events and ILD Outpatient Clinic

Type of Visit	Screening	Baseline	Weekly	Monthly	Every 3 Months	ILD Clinic visits as per standard clinical care	Annually as per standard clinical care
Visit Number	1	2					
Week Number	-2	0					
Inclusion/Exclusion Criteria	✓						
Informed consent	✓	✓					
Medical history	✓						
Medications	✓					✓	
Acute exacerbations	✓					✓	
Spirometry	✓					✓	
DLCO[1]	✓					✓	
Six Minute Walk Test	✓					✓	
Serology	✓	✓					✓
HRCT[2]	✓	✓					✓
Dispensing of Apple watch and iPhone		✓					
Training on PROMS[3]		✓					
Breathlessness (PROMS mMRC))[4]		✓	✓				
Cough (PROMS VAS)[5]		✓	✓				
Cough (PROMS LCQ)[6]		✓		✓			
Fatigue (PROMS FAS)[7]				✓			
HRQoL (PROMS K-BILD)[8]					✓		
Apple Health Export Biometric and activity data						✓	

[1] DLCO: Diffusion capacity of the lungs for carbon monoxide, [2]HRCT: High Resolution Computed Topography, [3]PROMS: Patient reported outcome measures, [4]mMRC: Modified Medical Research Council dyspnoea scale, [5]VAS: Visual analogue scale for cough severity, [6]LCQ: Leicester Cough Questionnaire,[7]FAS: Fatigue Assessment Scale,[8]K-BILD: King's Brief Interstitial Lung Disease.

Clinical outcomes (disease progression, onset of acute exacerbations and mortality) will be assessed at clinic appointments, through changes in imaging (Computed Tomography (CT) of the chest), pulmonary function tests, 6-minute walk test distance and patient symptoms. Follow up data will be captured in the patient's research record in REDCap.

Data analysis

The overall objective of this study is to leverage the data collected to generate predictive models of outcome. Specifically, we will seek to develop individual patient level models, of the following categorical outcomes of

1. Acute exacerbation
2. Hospitalisation
3. Mortality

Models will be developed using a number of different approaches, with all models evaluated to determine best performance. These may include:

1) Logistic regression for mortality, acute exacerbation and hospitalization modelling the relationship between independent variables (for example biometric or PROM data) and each of the model outcomes of mortality, acute exacerbation and hospitalization.
2) Proportional Hazards to predict time to event (mortality) estimate the hazard ratio for input variables.
3) Linear regression, for continuous outcomes (exacerbations) where a linear relationship between variables (for example cough questionnaire) is anticipated
4) Survival Analysis: For time to event outcomes, where outcome (for example mortality) will be explored in the context of data collected from patients

To complement these approaches, we will also utilize machine learning approaches to explore more complex models, including

1) Random Forest, to classify outcomes based on biometric or PROM variable data, as both single and multiple predictors
2) Support Vector Machines- classifying patients as progressors or not based on multiple variables including PROM and biometric data, as well as clinical phenotype
3) K-Nearest neighbours - to make predictions about individual patients based on similarities in the dataset to other patients with known outcomes
4) Pattern discovery: Unsupervised machine learning methods will be utilized to uncover patterns in the data applying methods such as clustering to identify for example high symptom burden patients from PROMs data

All models will be evaluated and compared to determine performance. Key evaluation measures will include

- Accuracy: based on assessment of predicted vs observed outcomes, as participants continue in the cohort
- Sensitivity and Specificity: analysis of true predictors, based on comparison with actual participants outcome
- C-Statistics: to measure performance of models in discriminating between outcomes

Six-month patient data will be collected and analysed. Outcomes predicted will be compared with actual outcomes recorded at month twelve, allowing refinement, enhancement, and validation of the developed models. Data analysis will proceed as follows:

Step 1: Data quality control and validation:

Individual patient level data will be reviewed for completeness. Given the lack of previous studies integrating clinical, physiological, activity and patient reported data in this population, we will employ a conservative approach to data completeness with a requirement for at least 70% completeness for key variables. Missing data will be summarised and dealt with through

Case deletion, where the 70% threshold is not reached. Rational substitution will be employed where possible. For missing data at random, arithmetic imputation methods will be employed including worst case imputation for dropouts and interpolation/extrapolation where prior and after data is available.

Step 2: Descriptive analysis: Data from activity and physiological measurements will be summarised as per Table 3. Data summarisation will involve providing a concise overview of key characteristics of the dataset including central tendency values, variance to give insights into the spread of the data and frequency distributions, assessing for normality, while highlighting any notable patterns or peaks in the dataset.

Step 3: Feature selection, potential predictors of progression, exacerbation and mortality will be reviewed by specialist respiratory clinicians. The rationale behind the feature selection process will be clearly documented and the chosen predictors will be rigorously validated to ensure clinical relevance to the target outcomes. In parallel, activity and physiological measurements (individual and composite, single point and time-rolling average trends, raw values and individual-normalised) will be analysed to identify potential features, predictive of clinical outcome.

Step 4: Predicted outcomes, the prediction model will be simplified by the elimination of unnecessary variables. Methods such as correlation analysis will be performed to identify potential redundancies. Statistical tests will be used to rank features and their association with the target outcomes.

Step 5: Validation: we will evaluate the predictive model's effectiveness by comparing its performance to actual events per predictor both within individual and across individuals. By doing this we will gain insights into the practical relevance of individual predictors, and we will refine the model accordingly.

Table 3: Baseline characteristics from wearable

Variable	Min	Max	Mean (Standard Deviation)	Coefficient of variation	Weekly Δ max - min	Days with data	Hours worn	Number of recordings
Oxygen Saturation (%)								
Heart Rate (count/min)								
Heart Rate Variability (ms)								
Resting Heart Rate (count/min)								
Walking Heart Rate Average (count/min)								
Respiratory Rate (count/min)								
StepCount (count)								
Distance Walking Running (km)								
Six Minute Walk Test Distance (m)								
Sleep Analysis								
Wrist Temperature								

Patient and Public involvement

This study has been designed together with input from patients and patient advocates. Specifically, the Irish Lung Fibrosis patient and public involvement (PPI) Research Advisory Group and their volunteers have provided input into the relevance of wearables and research for this population, as well as recruitment and onboarding procedures to ensure these are designed with the end users in mind. In addition, we have co-developed the patient facing resources (including the website) to ensure their relevance, comprehensibility, and accessibility for this patient population.

Discussion

The primary objective of this study is to develop robust prediction models for clinical outcomes in ILD using a combination of clinical, physiological, activity and patient reported data fields. Through the comprehensive collection and analysis of these data sets, we aim to identify better predictors of disease progression, acute exacerbations, and mortality. Predictions will be compared to actual outcomes to validate prediction models. These predictions are expected to offer meaningful value for clinical management by providing clinicians with an improved tool for prognostication.

By developing a clinical prediction model of ILD outcomes, we anticipate that such predictors will be useful as patient level controls for clinical trials, where predicted outcomes for individual patients can be compared to actual recorded outcomes following an intervention. In this manner individual patients will act as their own control, enabling additional investigation of new therapeutic strategies.

A strength of this study lies in the ability of the wearable to passively collect a large quantity of biometric data without burdening the patient to record active measures. However, the patient reported outcome measures may represent a burden to patients.

Declarations

Acknowledgements

Health Research Board – Trials Methodology Research Network (HRB-TMRN), Galway, Ireland and the UCD Clinical Research Centre for their support of this study. Lindsay Brown Advanced Nurse Practitioner for assistance with PROMS selection. Dylan Keagan PIL/ICF. Irish Lung Fibrosis Association PPI Nicola Cassidy, Paula Jenkins, and Robert Hurley for their input to the design of the Website and Patrick McKay, Advantage Point Creative Consultancy

for Website design.

Author Contributions

Conception and design of the study: PD, MK, CMcC, ANF, SH, EG. Drafting the manuscript: SH, EG. Revising the manuscript for important intellectual content: PD, MK, CMcC, ANFSH, EG. Approval of the version of the manuscript to be published: PD, CMcC, MK. Peter Doran is the guarantor.

Funding

The grant support for this research study was provided by the Health Research Board – Trials Methodology Research Network (HRB-TMRN), The funding body had no role in the design of this study protocol, and will not be involved in the collection, analysis, and interpretation of data or manuscript preparation.

Availability of data and materials

Grouped data generated or analysed during this study will be published. No individual data will be made available to protect the recognition of individual participants.

Competing interests

All authors have no competing interest to declare.

Ethics and dissemination.

Full ethics approval was granted for this research study by the St. Vincent's University Hospital Research Ethics Committee, Dublin, Ireland; Ref no: RS23-023. Explicit consent will be sought from participants for the collection and processing of their data. Data Processing agreements will be in place to ensure that personal data will be processed as is necessary to achieve the objective of the health research and to ensure that data shall not be processed in such a way that might cause damage or distress to the data subject. St. Vincent's University Hospital and University College Dublin will be joint data controllers for the study. This study will comply with the General Data Protection Regulation.Study results will be presented at medical conferences and will be disseminated via peer-reviewed journals.

References

1. Kumar, Abbas, Fausto, Mitchell. Robbins Basic Pathology, Chapter 13 The Lung, Diffuse

Interstitial Lung Diseases, pg. 792 8th Edition Elsevier 2007.

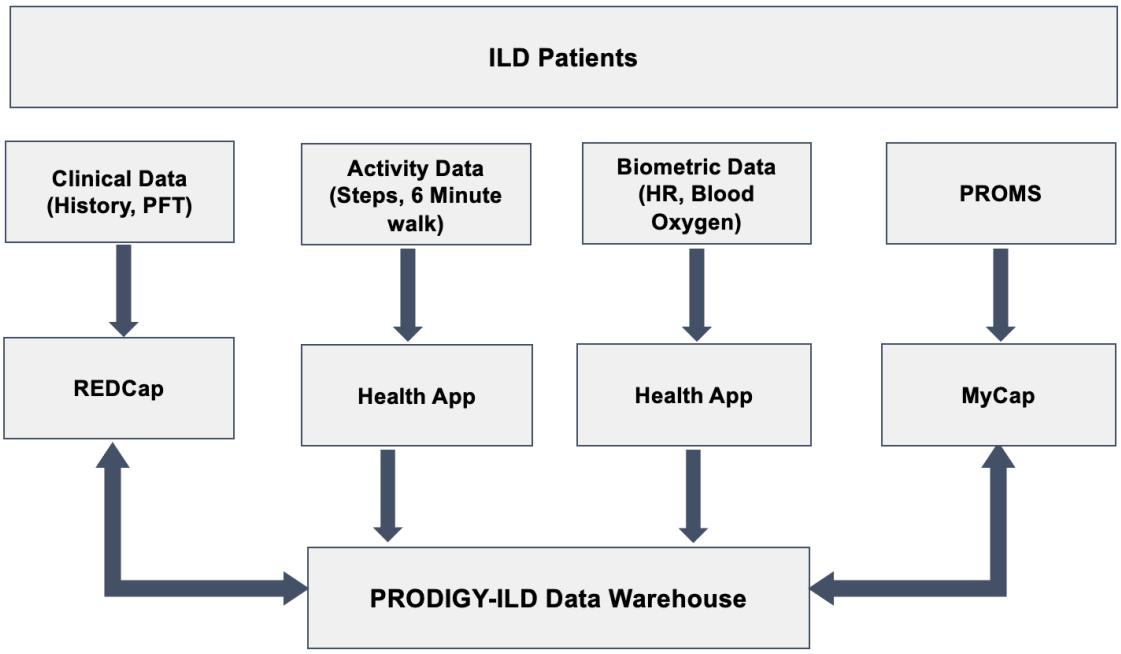
2. King T.E, Flaherty K.R and Dieffenbach P. Approach to the adult with interstitial lung disease: Clinical evaluation. UpToDate Wolters Kluwer Literature review current through Feb 2023. Accessed March 2023.
3. Raghu G et al. Idiopathic Pulmonary Fibrosis (an Update) and Progressive Pulmonary Fibrosis in Adults: An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. Am J Respir Crit Care Med. 2022 May 1; 205(9): e18-e47. Published online 2022 May1. Doi:10.1164/rccm.202202-0399ST
4. Salciccioli J.D, Marshall D.C, Goodall R, Crowley C et al. Interstitial lung disease incidence and mortality in the UK and the European Union: an observational study, 2001-2017.ERJ Open Research 2022 8: 00058-2022; DOI: 10.1183/23120541.00058-2022
5. Irish Thoracic Society Position statement on the management of Idiopathic Pulmonary Fibrosis. The Irish Thoracic Society Interstitial Lung Disease Group. August 2018 https://irishthoracicsociety.com/wp-content/uploads/2018/08/ITS-Pulmonary-Fibrosis-Position-Statement.pdf
6. Raghu G, Collard HR, Egan JJ, et al. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. Am J Respir Crit Care Med 2011; 183: 788–824.
7. American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. international consensus statement. American Thoracic Society (ATS), and the European Respiratory Society (ERS). *Am J Respir Crit Care Med.* 2000;161(2 Pt 1):646-664. doi: 10.1164/ajrccm.161.2.ats3-00.
8. Irish Lung Fibrosis Association. Available at https://ilfa.ie/ Accessed on the 8th of December 2022
9. Collard HR et al. Acute Exacerbation of Idiopathic Pulmonary Fibrosis. An International Working Group Report. *Am J Respir Crit Care Med.* 2016 Aug 1;194(3):265-75. doi: 10.1164/rccm.201604-0801CI
10. Kim JH., Perlman D. and Tomic R. Natural history of pulmonary fibrosis. Respiratory Medicine (2015) 109, 661-670. http://dx.doi.org/10.1016/j.rmed.2015.02.002
11. Lee SH, Kim SY, Kim DS et al. Predicting survival of patients with idiopathic pulmonary fibrosis using GAP score: a nationwide cohort study. Respiratory Research 2016; 17:131 doi: 10.1186/s12931-016-0454-0
12. King TE, Tooze JA, Schwarz MI, Brown KR, Cherniack RM. Predicting survival in idiopathic pulmonary fibrosis: Scoring system and survival model. Am J Respir Crit Care Med. 2001; 164:1171 – 81.
13. Wells AU, Desai SR, Rubens MB, Goh NSL, Cramer D, Nicholson AG, Colby TV, du Bois RM, Hansell DM. Idiopathic pulmonary fibrosis - A composite physiologic index derived from disease extent observed by computed tomography. Am J Respir Crit Care Med. 2003; 167:962–9.

14. du Bois RM, Weycker D, Albera C, Bradford WZ, Costabel U, Kartashov A, Lancaster L, Noble PW, Raghu G, Sahn SA, et al. Ascertainment of individual risk of mortality for patients with idiopathic pulmonary fibrosis. Am J Respir Crit Care Med. 2011; 184:459–66.
15. Richards TJ, Kaminski N, Baribaud F, Flavin S, Brodmerkel C, Horowitz D, Li K, Choi J, Vuga LJ, Lindell KO, et al. Peripheral blood proteins predict mortality in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med. 2012; 185:67–76.
16. Ley B, Ryerson CJ, Vittinghoff E, Ryu JH, Tomassetti S, Lee JS, Poletti V, Buccioli M, Elicker BM, Jones KD, et al. A Multidimensional Index and Staging System for Idiopathic Pulmonary Fibrosis. Ann Intern Med. 2012; 156:684–U658.
17. Fujii H, Hara Y, Saigusa Y, Tagami Y, et al.ILD-GAP Combined with the Charlson Comorbidity Index Score (ILD-GAPC) as a Prognostic Prediction Model in Patients with Interstitial Lung Disease. Can Respir J. 2023; 2023: 5088207.Published online 2023 Feb 8. Doi: 10.1155/2023/5088207
18. Lee J.H., Jang J.H., Jang H-J., Kim S.Y, et al. New prognostic scoring system for mortality in idiopathic pulmonary fibrosis by modifying the gender, age, and physiology model with desaturation during the six-minute walk test. Front Med (Lausanne). 2023 Jan 25:10:1052129. Doi:10.3389/fmed.2023.1052129
19. Althobiani M, et al. Variables used to monitor patients with Interstitial Lung Disease. BMJ Open Resp Res 2021; 8: e001088
20. Althobiani M., Alqahtani JS, Hurst JR, Russell AM, and Porter J. Telehealth for patients with interstitial lung diseases (ILD): results of an international survey of clinicians.
21. Althobiani M, Evans R, Alqahtani J, et al. Home monitoring of physiology and symptoms to detect interstitial lung disease exacerbations and progression: a systematic review. ERJ Open Res. 2021 Dec 20;7(4): 00441-2021. Doi:101183/23120541.00441-2021
22. Catharina C Moor, Rémy L M Mostard, Jan C Grutters, Paul Bresser, Joachim G J V Aerts, Niels H Chavannes, Marlies S Wijsenbeek. Home Monitoring in Patients with Idiopathic Pulmonary Fibrosis. A Randomized Controlled TrialAm J Respir Crit Care Med 2020 Aug 1;202(3):393-401. Doi: 10.1164/rccm.202002-0328OC.
23. Imre Noth, Vincent Cokn, Nazia Chaudhuri, Tamera J Corte, Kerri A Johannson, Marlies Wijsenbeek, Stephane Jouneau, Andreas Michael, Manuel Quaresma, Klaus B Rohr, Anne-Marie Russell, Susanne Stowasser, Toby M Maher, INMARK trial investigators. Home spirometry in patients with idiopathic pulmonary fibrosis: data from the INMARK trial Eur Respir J. 2021 Jul 8;58(1): 2001518. Doi:10.1183/13993003.01518-2020. Print 2021 Jul
24. Marlies S Wijsenbeek, Elisabeth Bendstrup, Claudia Valenzuela, Michael T Henry, Catharina C Moor, Stéphane Jouneau, Alessandro G Fois, Onofre Moran- Mendoza, Syed Anees, Mirela Mirt, Monica Bengus, Frank Gilberg, Klaus-Uwe Kirchgaessler, Carlo Vancheri. Disease Behaviour During the Peri-Diagnostic Period in Patients with Suspected Interstital Lung Disease: The STARLINER Study Adv Ther. 2021 Jul;38(7):4040-4056. doi: 10.1007/s12325-021-01790-y. Epub 2021 Jun 11.

25. S. Walsh, T. Cahill, C. Edwards, E. Costello, J. Walsh, A.-M. Russell, A.W. O'Regan. PaGent-Reported Monitoring of Symptoms and Spirometry Via the patientMpower Plaform in Idiopathic Pulmonary Fibrosis. American Journal of Respiratory and CriGcal Care Medicine 2018; 197: A4933
26. Wim A. Wuyts, * Marlies Wijsenbeek, Benjamin Bondue, Demosthenes Bouros, Paul Bresser, Carlos Robalo Cordeiro, Ole Hilberg, Jesper Magnusson, Effrosyni D. Manali, António Morais, Spyridon Papiris, Saher Shaker, Marcel Veltkamp, and Elisabeth Bendstrupm. Idiopathic Pulmonary Fibrosis: Best Practice in Monitoring and Managing a Relentless Fibrotic Disease. Respiration. 2020 Jan; 99(1): 73–82. Published online 2019 Dec12. Doi: 10.1159/000504763
27. Toby M Maher 1 2, Courtney Schiffman 3, Michael Kreuter 4 5, Catharina C Moor 6, Steven D Nathan 7, Judit Axmann 8, Paula Belloni 3, Monica Bengus 8, Frank Gilberg 8, Klaus-Uwe Kirchgaessler 8, Marlies S Wijsenbeek 6. A review of the challenges, learnings, and future directions of home handheld spirometry in interstitial lung disease. Respir Res 2022 Nov 11;23(1):307. Doi: 10.1186/ s12931-022-02221-4.
28. Marlies S Wijsenbeek, Catharina C Moor, Kerri A Johannson, Peter D Jackson, Yet H Khor, Yasuhiro Kondoh, Sujeet K Rajan, Gabriela C Tabaj, Brenda E Varela, Pieter van der Wal, Richard N van Zyl-Smit, Michael Kreuter, Toby M Maher. Home monitoring in interstitial lung diseases. Lancet Respir Med. 2022 Oct 4; Doi:10.1016/ S2213-2600(22)00228-4
29. Apple Heart Study Protocol: Evaluation of the use of Apple Watch faeatures for identification of cardiac arrhythmias. NCT03769207 Available at https://Clinicaltrials.gov Accessed March 2023.
30. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. J Clin Epidemiol1995;48:1503-10.
31. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol1996;49:1373-9
32. PA Harris, R Taylor, R Thielke, J Payne, N Gonzalez, JG. Conde, Research electronic data capture (REDCap) – A metadata-driven methodology and workflow process for providing translational research informatics support, J Biomed Inform. 2009 Apr;42(2):377-81.
33. PA Harris, R Taylor, BL Minor, V Elliott, M Fernandez, L O'Neal, L McLeod, G Delacqua, F Delacqua, J Kirby, SN Duda, REDCap Consortium, The REDCap consortium: Building an international community of software partners, J Biomed Inform. 2019 May 9 [doi: 10.1016/j.jbi.2019.103208]
34. PA Harris, J Swafford, E S Serdoz, J Eidenmuller, G Delacqua, V Jagtap, R J Taylor, A Gelbard, A C Cheng, S N Duda, "MyCap: a flexible and configurable platform for mobilizing the participant voice", JAMIA Open, Volume 5, Issue 2, July 2022, ooac047. https://doi.org/10.1093/jamiaopen/ooac047
35. Mahler DA, Wells CK. Evaluation of clinical methods for rating dyspnea. Chest 1988; 93: 580–586. [PubMed] [Google Scholar]

36. Bestall JC, Paul EA, Garrod R, et al. Usefulness of the medical research council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. Thorax 1999; 54: 581–586. [PMC free article] [PubMed] [Google Scholar]
37. Martin Nguyen A, Bacci ED, Vernon M, et al. Validation of a visual analog scale for assessing cough severity in patients with chronic cough. Ther Adv Respir Dis 2021 Oct 26. doi: 10.1177/17534666211049743.
38. Birring S, Prudon B, Carr A, Singh S, Morgan M, Pavord I. Development of a symptom specifc health status measure for patients with chronic cough: Leicester Cough Questionnaire (LCQ). Thorax 2003; 58:339–43
39. FAS: Fatigue Assessment Scale and how to use FAS. Accessed June 2023, Available at: https://www.ildcare.nl/index.php/condition-use-the-fas-fatigue-assessment-scale/ https://www.ildcare.nl/index.php/how-to-use-the-fas-fatigue-assessment-scale/
40. Hendriks C, Drent M, Elfferich M, De Vries J. The Fatigue Assessment Scale (FAS): quality and availability in sarcoidosis and other diseases. Curr Opin Pulm Med 2018; 24 (5): 495-503. https://pubmed.ncbi.nlm.nih.gov/29889115
41. Patel AS, Siegert RJ, Brignall K, et al. The development and validation of the King's Brief Interstitial Lung Disease (K-BILD) health status questionnaire. BMJ Thorax 2011 doi: 10.1136/thoraxjnl-2012-201581
42. Claire M. Nolan, Surinder S. Birring, Matthew Maddocks, Toby M. Maher, Suhani Patel, Ruth E. Barker, Sarah E. Jones, Jessica A. Walsh, Stephanie C. Wynne, Peter M. George, William D-C. Man. King's Brief Interstitial Lung Disease questionnaire: responsiveness and minimum clinically important difference. European Respiratory Journal 2019, 54:1900281; DOI: 10.1183/13993003.00281-201

Figure 1: Data Warehouse of data collected from wearable, app, and patient record.



DEPARTMENT OF RESPIRATORY MEDICINE, ST. VINCENT'S UNIVERSITY HOSPITAL, DUBLIN

PARTICIPANT INFORMATION AND CONSENT FORM

STUDY TITLE: IPF/PPF Apple Study

NAME OF PRINCIPAL INVESTIGATOR: Dr Cormac McCarthy

You are being invited to participate in a research study. Thank you for taking time to read this.

WHAT IS THE PURPOSE OF THIS STUDY?

The aim of the study is to monitor participants with a diagnosis of Idiopathic Pulmonary Fibrosis or Progressive Pulmonary Fibrosis. Participants will be given an Apple smartwatch series 6 or above, and an iPhone series 8 (if they do not have one) with an App to answer questions about their symptoms this information will be updated to a research database. The results of this study will be used to better understand the symptoms and trajectory of the disease, as well as develop prediction models for clinical outcomes.

WHY HAVE I BEEN CHOSEN?

You have been chosen for this study as you have a diagnosis of Idiopathic Pulmonary Fibrosis or Progressive Pulmonary Fibrosis and the specialist team providing your care have deemed you to be a suitable candidate for the study.

WHAT WILL HAPPEN IF I VOLUNTEER?

Your participation is entirely voluntary. If you initially decide to take part you can subsequently change your mind without any adverse consequences. This will not affect your future treatment in any way. Furthermore, your doctor may decide to withdraw you from this study if they feel it is in your best interest.
If you agree to participate, you will be requested to wear an Apple watch for at least 20 hours per day, and to complete electronic questionnaires on a mobile application on an iPhone about your symptoms. Weekly questionnaires consist of a breathlessness and cough rating which take less than 1 minute to complete. Monthly fatigue and cough questionnaires take approximately 15 minutes in total to complete. Every three months a health-related quality of life questionnaire will take approximately 10 minutes to complete.

You will also come to your usual outpatient clinic appointments where you will have routine lung function tests, radiology tests and give blood samples where they are necessary for your care.

Study visits, apart from the baseline visit, will happen on clinic days. The baseline visit will require additional time spent with the study research team member for initial training on the use of the Apple watch and iPhone App. Once participants are on-board the study

subsequent study visits will be routine as per usual outpatient clinic visits and standard care with the specialist rare lung disease team. Before you leave your clinic visit Apple watch data will be exported from your iPhone Health App to our research system. It is important to understand that an Apple Watch is not a medical device, the data gathered from the Apple watch will not be monitored by hospital staff real-time but will be used to update our research system every four months when you come to clinic. You should follow your doctor's advice on the use of approved medical devices for monitoring your blood oxygen levels to adjust your supplemental oxygen requirements.

ARE THERE ANY BENEFITS FROM MY PARTICIPATION?

It is hoped the study information we will obtain may provide further knowledge of this condition, and can be used to inform treatment in future.

ARE THERE ANY RISKS INVOLVED IN PARTICIPATING?

There are no increased risks associated with this study above standard care at outpatient clinics, which will include minimal risks associated with taking blood samples and radiology images. It is an observational study to collect real-time participant data.

WHAT HAPPENS IF I DO NOT AGREE TO PARTICIPATE?

If you decide not to participate in this study your treatment will not be affected in any way.

CONFIDENTIALITY

Your identity will remain confidential. A study number will identify you. Your name will not be published or disclosed to anyone. Anonymised results will be published in medical journals and presented at peer groups.

COMPENSATION

Your doctors are adequately insured by virtue of their participation in the Clinical Indemnity Scheme.

WHO IS ORGANISING AND FUNDING THIS RESEARCH?

This study is organised and funded by HRB-Trials Methodology Research Network, UCD Clinical Research Centre, Prof Michael Keane & Prof Peter Doran Research funds.

Will, I be paid for taking part in this study? No

Will my expenses be covered for taking part in this study? Travel costs related to attending the baseline visit will be reimbursed.

HAS THIS STUDY BEEN REVIEWED BY AN ETHICS COMMITTEE?

The St. Vincent's Healthcare Group, Ethics and Medical Research Committee have reviewed and approved this study.

CONTACT DETAILS

Principal Investigator: Dr Cormac McCarthy (01) 221 3702

RESEARCH PARTICIPANT'S RIGHTS

If you have any questions about your rights as a research participant, then you may contact the Hospital's Quality & Patient Safety Department 01 2214013

CONFIDENTIALITY & DATA PROTECTION

1. INTRODUCTION

1.1 This Participant Information and Consent Form provides guidance and information to IPF/PPF Apple study research participants regarding the processing of the research participants' personal data. St. Vincent's University Hospital is committed to protecting and respecting your privacy. This Participant Information and Consent Form together sets out the basis on which any personal data we collect from you or that you provide to us will be processed by us an independent data controller. Please read this Participant Information and Consent Form carefully to understand our treatment and use of your personal data.

1.2 The processing of your personal data will be in compliance with the Data Protection Acts 1988 to 2018 (as amended) and the General Data Protection Regulation (the "Data Protection Legislation").

1.3 Please note that agreeing to participate in a research program with St. Vincent's University Hospital, you acknowledge that you have read, understood and agree to this Participant Information and Consent Form.

2. IDENTITY OF THE CONTROLLER OF PERSONAL INFORMATION

For the purposes of the Data Protection Legislation, St. Vincent's University Hospital is an independent data controller in the following circumstances:

1.	**SVUH – Facility Provider for patients and clinical care** **Company Type**: PLC **Company Registration number** : 338585 Having its registered office at: SVUH, Elm Park, Dublin 4 **Dr Cormac McCarthy – Principal Investigator and clinical care** **Company Type:** SVUH Employee, **Company registration number -** NA Having its registered office at: SVUH, Elm Park, Dublin 4

3. CONTACT DETAILS OF THE DATA PROTECTION OFFICER*

3.1

The data protection officer for SVUH and PI is: Sean Gibney, SVUH, Elm Park, Dublin 4, Contact 01 221 3591 Email: dataprotection@svuh.ie

4. PROCESSING OF YOUR PERSONAL DATA

St Vincent's University Hospital will process your personal data for the following purposes on the basis of your consent:

Personal data	Purpose of processing[1]
1. **Identification e.g. name, address, DOB (please note this and subsequent information will be anonymised/coded once leaving SVUH)**	1. **(a)Originally captured as part of medical care (b) used for purpose of carrying out research**
2. **Test results**	2. **Research outcome analysis**
3. **Clinical History**	3. **PMH relevant to study research outcome analysis**
4. **Real-time biometric data collected by wearable device Apple watch**	4. **Statistical analysis , to build prediction models for clinical outcomes**
5. **Questionnaire (PROMS) data collected by iPhone App**	5. **PROMS to measure disease progression also for statistical analysis , to build prediction models for clinical outcomes**

.

4.1 Where does St. Vincent's University Hospital obtain my personal data from?

Most of the personal data we process is obtained from you directly but we also obtain personal data about you from your

- medical notes,
- lab test results,
- radiology results,
- lung function tests
- wearable Apple device,
- iPhone App questionnaire responses,

5. SHARING OF PERSONAL DATA

Your personal data will in particular be shared with:

*NOTE: These parties will either be acting as Processors of your information as part of this research study e.g. CROs, non-SVUH employees supporting research process or Controllers in their own right.

Person/Company/institute	Requirement for sharing
Research Study team Sinead Holden (UCD) Emer Gunne (UCD)	Data analysis and Prediction modelling - UCD Clinical Research Centre

5.1 Service Providers

We use third party service providers who provide services including financial services, occupational health and IT services. In providing the services, your personal data will, where applicable, be processed by the service provider on our behalf.

We will check any third party that we use to ensure that they can provide sufficient guarantees regarding the confidentiality and security of your personal data. We will have written contracts with them which provide assurances regarding the protections that they will give to your personal data and their compliance with our data security standards and international transfer restrictions.

5.2 Disclosures to Third Parties

In certain circumstances, we share and/or are obliged to share your personal data with third parties outside St. Vincent's Hospital, for the purposes described above and in accordance with Data Protection Legislation.

These third parties include but are not limited to:

- the Health Products Regulatory Authority;
- the Health Service Executive;
- the Joint Commission International;
- relevant industry bodies;
- external professional advisors; and
- others, where it is permitted by law, or where we have your consent.

6. TRANSFERS OUTSIDE THE EUROPEAN ECONOMIC AREA[2]

6.1a We will not transfer, store or process your personal data outside the European Economic Area.

6.1 If you would like to see a copy of any relevant provisions, please contact your data protection officer (see "Contact Us" section below).

7. HOW IS MY PERSONAL DATA SECURED

7.1 St. Vincent's University Hospital operates and uses appropriate technical and physical security measures to protect your personal data.

7.2 We have in particular taken appropriate security measures to protect your personal data from accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access, in connection with this research study. Access is only granted on a need-to-know basis to those people whose roles require them to process your personal data. In addition, our service providers are also selected carefully and required to use appropriate protective measures.

8. STORAGE OF PERSONAL DATA

We will keep your personal data for 10 years. This may mean that some Information is held for longer than other information.

9. YOUR RIGHTS

9.1 You may have various rights under Data Protection Legislation. However, in certain circumstances, these rights may be restricted[3]. In particular, your rights may be restricted where this is necessary: (i) for the prevention, detection, investigation and prosecution of criminal offences; (ii) in contemplation of or for the establishment, exercise or defence of a legal claim or legal proceedings (whether before a court, tribunal, statutory body or an administrative or out-of-court procedure); and/or (iii) for the performance of a task carried out in the public interest or in the exercise of official authority vested in St. Vincent's University Hospital.

These rights may include (as relevant):

(i) **The right of access** enables you to check what type of personal data we hold about you and what we do with that personal data and to receive a copy of this personal data;

(ii) **The right to object** to processing of your personal data where that processing is carried out on the basis of our legitimate interests. We will stop using your personal data unless we can demonstrate an overriding legitimate ground for the continued processing of this personal data;

(iii) **The right to rectification** enables you to correct any inaccurate or incomplete personal data that we hold about you;

(iv) **The right to erasure** enables you to request that we erase personal data held about you in certain circumstances;

(v) **The right to restrict processing** of your personal data by us in certain cases, including if you believe that the personal data held about you is inaccurate or our use of the personal data is unlawful; and

(vi) **The right to data portability** enables you to receive your personal data in a structured, commonly used and machine readable format and to have that personal data transmitted to another data controller

10. YOUR RIGHT TO LODGE A COMPLAINT WITH A SUPERVISORY AUTHORITY

10.1 Without prejudice to any other administrative or judicial remedy you might have, you may have the right under data protection legislation in your country (where applicable) to lodge a complaint with the relevant data protection supervisory authority in your country (i.e. the Office of the Data Protection Commissioner in Ireland) if you consider that we have infringed applicable data protection legislation when processing your personal data. This means the country where you are habitually resident, where you work or where the alleged infringement took place.

11. CHANGES TO THIS INFORMATION

11.1 We may decide to make changes to this Participant Information and Consent Form. If a change is made, we will notify you in person of such changes. An updated Participant Information and Consent Form will be provided to you in advance of any change actually taking effect.

12. CONTACT US

12.1 For further information or if you have any questions or queries about this Participant Information and Consent Form, please contact:

By letter:	**Sean Gibney,** Data Protection Officer, St Vincent's University Hospital Elm Park, Dublin 4
By email:	dataprotection@st-vincents.ie
By telephone:	(01) 221 3591

CONSENT FORM

PLEASE TICK YOUR RESPONSE IN THE APPROPRIATE BOX

- I have read and understood the Participant Information and Consent Form YES ☐ NO ☐
- I have had the opportunity to ask questions and discuss the study YES ☐ NO ☐
- I have received satisfactory answers to all my questions YES ☐ NO ☐
- I have received enough information about this study YES ☐ NO ☐
- I understand that I am free to withdraw from the study at any time without giving a reason and without this affecting my future medical care YES ☐ NO ☐
- I am aware of the potential risks, benefits and alternatives of this research study. YES ☐ NO ☐
- I consent to take part in this research study having been fully informed of the risks, benefits and alternatives. YES ☐ NO ☐
- I give informed consent to have my data processed as part of this research study. YES ☐ NO ☐

STORAGE & FUTURE USE OF INFORMATION:

- I give permission for material/data to be stored for possible future research related to the current study only if consent is obtained at the time of the future research but only if the research is approved by a Research Ethics Committee. YES ☐ NO ☐
- I give permission for material/data to be stored for possible future research related to the current study without further consent being required but only if the research is approved by a Research Ethics Committee. YES ☐ NO ☐
- I give permission for material/data to be stored for possible future research unrelated to the current study only if consent is obtained at the time of the future research but only if the research is approved by a Research Ethics Committee. YES ☐ NO ☐
- I give permission for material/data to be stored for possible future research unrelated to the current study without further consent being required but only if the research is approved by a Research Ethics Committee. YES ☐ NO ☐

- I agree that some future research projects may be carried out by researchers working for commercial/pharmaceutical companies. YES ☐ NO ☐

- I understand I will not be entitled to a share of any profits that may arise from the future use of my material/data or products derived from it.
YES ☐ NO ☐

Participant's Signature: _____
Date: _____
Participant's Name (block capitals): _____

Research Team Member's Signature: _____
Date: _____
Research Team Member's Name (block capitals): _____

Translator's Signature: _____
Date: _____
Translator's Name (block capitals): _____

Legal Rep./Guardian's Signature: _____
Date: _____
Legal Rep./Guardian's Name (block capitals): _____

Table 1: Apple Watch Continuous Monitoring Data

Variable	Data recorded
Heart Rate (HR) (BPM)	Heart Rate (Beats Per Minute): Apple watch records HR approximately every 5 minutes. Resting Heart Rate: The average heart beats per minute measured when the wearer has been inactive or relaxed for several minutes. Heart Rate Variability: A measure of the variation in the time interval between beats. Apple watch calculates HRV by using standard deviation of beat-to-beat measurements which are captured by the heart rate sensor. Average HRV per day is recorded. Walking Heart Rate Average: the average heart beats per minute measured by Apple Watch during walks at a steady pace throughout the day.
Respiratory Rate (breaths/min)	Apple watch records Respiratory Rate while the wearer is sleeping
Blood Oxygen (SpO2)	Apple watch attempts to record oxygen saturation every 30 minutes during sleep. It will also opportunistically record oxygen saturation during the day while arm is at rest. Also, a participant can opt to do a recording themselves at their convenience.
Six-Minute Walk estimate	A weekly estimate of how far the wearer can walk on flat ground in six minutes based on recent motion and workout data. The Apple Watch can give a predicted six-minute walk distance up to 500 meters.
Step Count (count)	Step count is the number of steps you take throughout the day.
Distance Walking Running (km)	Calculates an average distance you have walked and run over the last 7 days
Exercise minutes	Measure of how many minutes of brisk activity you do
Sleep analysis	Time In Bed Time Asleep (Core, Deep, REM) Time Awake
Wrist Temperature	Wrist temperature is a measurement related to your body's temperature taken by Apple Watch while you are sleeping. Each value is an average of several measurements taken during sleep.