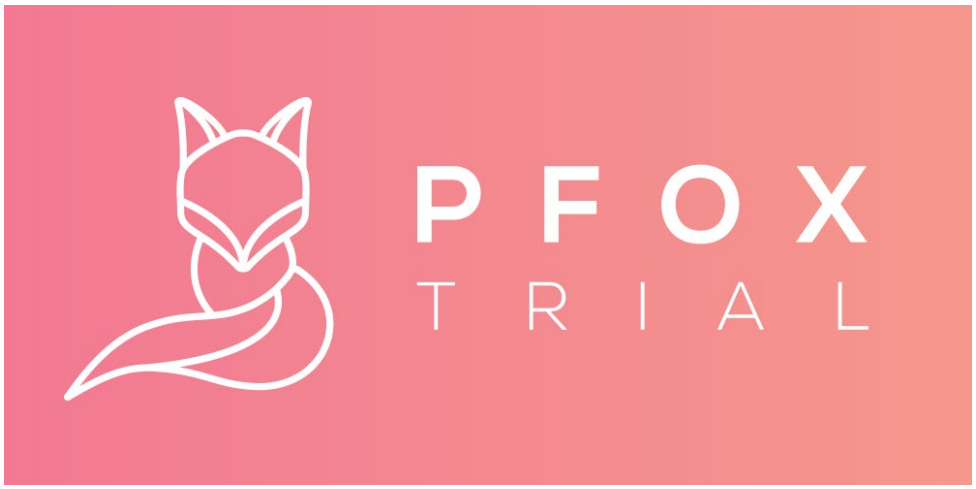

Ambulatory oxygen for treatment of exertional hypoxaemia in pulmonary fibrosis (PFOX): a randomised controlled trial.

Statistical Analysis Plan

Version 3.0

28th June 2024

Anne Holland, Graham Hepworth, Christine McDonald, Tamera Corte, Daniel Chambers, Andrew Palmer, Magnus Ekström, Ian Glaspole, Nicole Goh.

Table of Contents

LIST OF ABBREVIATIONS

6MWT	6-minute walk test
CI	confidence interval
CT	computed tomography
DSMB	Data Safety and Monitoring Board
FEV_1	forced expiratory volume in one second
FVC	forced vital capacity
GLMM	generalised linear mixed model
HRQL	health-related quality of life
ILD	interstitial lung disease
IPF	idiopathic pulmonary fibrosis
ITT	intention to treat
KBILD	King's Brief Interstitial Lung Disease questionnaire
LTOT	long term oxygen therapy
MAR	missing at random
NHMRC	National Health and Medical Research Council
PFOX	Pulmonary Fibrosis ambulatory OXygen trial
POC	portable oxygen concentrator
SpO_2	oxyhaemoglobin saturation
TEAE	treatment emergent adverse events
TLCO	diffusing capacity for carbon monoxide

1 ADMINISTRATIVE INFORMATION

1.1 STUDY IDENTIFIERS

- Protocol: HREC/18/Alfred/42, version 5, dated 29th November 2023
- ClinicalTrials.gov register Identifier: NCT03737409
- Published protocol: Holland AE, Corte T, Chambers DC, et al. Ambulatory oxygen for treatment of exertional hypoxaemia in pulmonary fibrosis (PFOX trial): a randomized controlled trial. BMJ Open 2020;10:e040798. doi:10.1136/bmjopen-2020-040798 (1)
- Funders: National Health and Medical Research Council (Australia) GNT 1139953, Swedish Society of Medicine (SLS-786791)

1.2 REVISION HISTORY

Version	Date	Changes made to document	Authors
1.0 (draft)	26th April 2024	Initial draft	Anne Holland
2.0 (draft)	18th June 2024	Incorporated edits from Graham Hepworth	Anne Holland, Graham Hepworth
3.0 (final)	28th June 2024	Incorporated edits from investigators	Anne Holland, Graham Hepworth

1.2.1 APPROVALS

The undersigned have reviewed this plan and approve it as final. They find it to be consistent with the requirements of the protocol as it applies to their respective areas. They also find it to be compliant with ICH-E9 principles and, in particular, confirm that this analysis plan was developed in a completely blinded manner (i.e. without knowledge of the effect of the intervention(s) being assessed).

Anne Holland, Principal Investigator

28/6/24

Graham Hepworth, Study Statistician

18/7/24

2 STUDY SYNOPSIS

The PFOX trial is a prospective, parallel group, randomized, multi-centre international clinical trial examining the effects of ambulatory oxygen vs ambulatory air in people with fibrotic interstitial lung disease (ILD) and exertional desaturation. Written, informed consent is obtained for each participant. Participants are randomised 1:1 into two groups: (i) Oxygen Group: ambulatory oxygen therapy using a portable oxygen concentrator (POC) and (ii) Air Group: sham therapy using an identical POC. Randomisation is conducted via a secure web-based interface. Participants, health professionals and trial staff are blinded to the POC being used. Participants are followed up at 3 and 6 months.

2.1 STUDY OBJECTIVES

2.1.1 PRIMARY OBJECTIVE

The primary objective of PFOX is to determine whether ambulatory oxygen therapy delivered using a POC improves physical activity in daily life at 3 months when compared to ambulatory air in people with fibrotic ILD.

Primary hypothesis: Ambulatory oxygen therapy (compared to ambulatory air) improves physical activity in daily life, measured by the number of steps per day, at 3 months.

2.1.2 SECONDARY OBJECTIVES

Secondary objectives of PFOX are to determine the effect of ambulatory oxygen therapy delivered using a POC on functional exercise capacity, health-related quality of life (HRQL), breathlessness, fatigue, anxiety, depression, time spent in moderate to vigorous physical activity, sedentary time and oxygen saturation in daily life at 3 and 6 months in people with fibrotic ILD.

Secondary hypotheses:
Ambulatory oxygen therapy (compared to ambulatory air):

1. Improves functional exercise capacity at 3 and 6 months
2. Results in better HRQL at 3 and 6 months
3. Reduces symptoms of breathlessness and fatigue at 3 and 6 months
4. Reduces anxiety and depression at 3 and 6 months
5. Improves time spent in moderate to vigorous physical activity and reduces sedentary time at 3 and 6 months
6. Increases steps per day at 6 months

2.2 PATIENT POPULATION

2.2.1 INCLUSION CRITERIA

- aged 18 years and over
- have a physician-confirmed diagnosis of fibrotic ILD, such as IPF, connective tissue disease-associated ILD, fibrotic hypersensitivity pneumonitis, idiopathic non-specific interstitial pneumonia, unclassifiable idiopathic interstitial pneumonia, environmental/occupational lung disease or sarcoidosis, with features of diffuse fILD of >10% extent on high-resolution CT, with ILD being the predominant pathological process (2)
- have had stable pharmacotherapy over the previous 3 months
- exhibit exertional desaturation, defined as oxyhaemoglobin saturation (SpO_2) ≤88% for at least 10 consecutive seconds during a 6-minute walk test (6MWT) performed on room air.

2.2.2 EXCLUSION CRITERIA

- Currently using or eligible for long-term oxygen therapy (LTOT), with eligibility defined as arterial oxygen pressure ≤55 mmHg at rest on room air, or 56–59 mm Hg with evidence of right heart failure
- current smokers, due to the risk of oxygen use near flames
- have predominantly obstructive lung disease, with forced expiratory ratio less than the lower limit of normal;
- pregnancy
- cognitively unable to consent
- currently participating in pulmonary rehabilitation
- non-ambulant
- have been admitted to an acute care hospital within the last 30 days
- death or transplant are anticipated within the study period.

OUTCOMES

2.2.3 PRIMARY OUTCOME

- Change from baseline in physical activity at 3 months, measured by the number of steps per day. Steps per day is measured using the StepWatch activity monitor (SAM) (Modus Health, Washington DC, USA). The SAM is worn on the ankle continuously for 7 days (except for bathing) at baseline, 3 months and 6 months. Days with < 8 hours recording or < 200 steps will be excluded. The mean number of steps per day will be calculated using the mean for all included days.

2.2.4 SECONDARY OUTCOMES

- Change from baseline in functional exercise capacity assessed by 6-minute walk distance at 3 and 6 months
- Change from baseline in health-related quality of life evaluated using the St George's Respiratory questionnaire at 3 and 6 months. Change in total score and domain scores will be reported.
- Change from baseline in health-related quality of life evaluated using the King's Brief Interstitial Lung Disease questionnaire at 3 and 6 months. Change in total score and domain scores will be reported.
- Change from baseline in dyspnea measured using the Dyspnea-12 questionnaire at 3 and 6 months. Change in total score and domain scores will be reported.
- Change from baseline in fatigue evaluated by the Fatigue Severity Scale at 3 and 6 months.
- Change from baseline in anxiety and depression measured by the Hospital Anxiety and Depression Scale at 3 and 6 months. Scores for anxiety and depression can vary from 0 to 21 and are stratified as follows: 0-7 (indicates absence of anxiety/depression symptoms); 8-10 (presence of symptoms of anxiety and depression in moderate degree - borderline); 11 or more (significant number of anxiety/depression symptoms - confirmed cases). Scores for anxiety and depression, as well as number of confirmed cases of anxiety and depression, will be reported.
- Change from baseline in time spent in moderate to vigorous physical activity, measured using the GeneActiv (GENEActiv, Cambridgeshire, a wrist-worn, triaxial accelerometer, at 3 and 6 months
- Change from baseline in sedentary time measured using the GeneActiv (GENEActiv, Cambridgeshire, a wrist-worn, triaxial accelerometer, at 3 and 6 months

2.2.5 SAFETY OUTCOMES

Adverse events are defined according to the Good Clinical Practice (GCP) guidelines including treatment emergent adverse events (TEAE's), which are undesirable events not present prior to medical treatment, or that worsens either in intensity or frequency following treatment. AEs of specific interest are:

- worsening of fILD (worsening of lung function, development of resting hypoxaemia)
- exacerbation of fILD

- burns (from smoking whilst using a POC, using the POC around an open flame or equipment that sparks)
- nosebleed or dry nose
- musculoskeletal injury from tripping on a POC
- bruising or infection at blood draw site
- fainting related to blood draw
- fainting, dizziness, chest pain, ataxic gait, lower limb pain or mental confusion related to 6-minute walk test
- hospitalization
- death.

2.2.6 TERTIARY OUTCOMES

- Use of oxygen therapy, downloaded from POC flash memory, measured in hours at 3 and 6 months
- Reason for cessation of therapy where relevant (patient request, commencement of LTOT, other).
- Oxygen saturation in daily life, measured using a Nonin 3150 Wrist Oximeter during waking hours for 2 consecutive days at 3 and 6 months.

2.3 INTERVENTION

For the oxygen group, ambulatory oxygen therapy will be delivered using the Inogen One G3 HF POC. Ambulatory air will be delivered using the Inogen One G3 HF POC that has been modified by the manufacturer to deliver only air. The POCs are identical in appearance.

All participants are informed that the aim of using a POC is to assist them to be more active, with fewer symptoms. They are encouraged to use the POC at all times when they are moving about, including walking at home or in the community, during exercise or during other activities. The POC should not be used when sitting still or sleeping. Written and verbal education are provided. Participants are encouraged to use their allocated POC during physical activity for the 6-month study period. The Inogen One G3 HF POC is used at its maximum flow setting of 5 for both groups.

2.4 RANDOMISATION AND BLINDING

Following baseline assessment participants are randomly allocated (1:1) to receive AOT or sham AOT (air) delivered by the Inogen One G3 HF POC. A computer generated, permuted block randomisation is used with stratification for (i) desaturation during 6MWT (<80% vs ≥80%) and (ii) site of recruitment. A remote, web-based, computer-generated randomisation procedure is used (provided by NHMRC Clinical Trials Centre).

Participants, clinicians and researchers will be blinded to group allocation. The Inogen One G3 HF POCs for AOT and air groups will be identical in appearance, display, weight and operation, with the only difference being the gas delivered. POCs were coded by the distributor, BOC, who have no involvement in trial conduct. All participants are advised against measuring oxygen saturation at home during the duration of the trial, as this does not represent usual clinical practice in any of the centres and may unblind the participants.

At the conclusion of the trial, participants are asked two questions to evaluate the success of blinding:

- Which treatment do you think you were receiving, oxygen or air?
- Did you have a pulse oximeter at home over the last 6 months? If yes, how did you use it?

2.5 SAMPLE SIZE

A total of 220 participants (110 per group) would provide 80% power to detect, at the two-sided 5% level, a clinically important difference between groups in the primary outcome of 599 steps per day (3). This assumes an SD of 1582 steps, based on physical activity data previously collected at our centre in 52 patients with fILD

(4). Previous experience suggests that 5% of participants could start LTOT (and cease POC use) over 6 months. We therefore intended to randomise 260 participants to ensure that 220 participants completed the study.

3 TRIAL REPORTING

The PFOX trial underwent important modifications because of the COVID-19 pandemic. We will report trial results according to the CONSERVE (CONSORT and SPIRIT Extension for RCTs Revised in Extenuating Circumstance) statement (5) as a framework to report the impact of the pandemic on PFOX and describe our mitigation approaches.

4 STATISTICAL ANALYSIS

4.1 GENERAL PRINCIPLES

- All outcomes and analyses are prospectively categorized as primary, secondary, or tertiary
- Summaries of continuous variables which are reasonably symmetrical will be presented as means and standard deviations, or as medians and inter-quartile ranges for skewed data. Categorical variables will be presented as frequencies and percentages.
- Differences in all endpoints between the two arms of the study will be tested independently at the two-tailed 0.05 level of significance. All estimates of treatment effects will be presented with 95% confidence intervals (CIs).
- Primary and secondary efficacy and safety outcomes analyses will be conducted on an intention-to-treat (ITT) basis so that all patients will be analysed in the group to which they were randomized irrespective of whether or not they received the allocated treatment.
- The ITT analysis strategy for PFOX is defined as follows:
    - Is based on an ITT design that aims to collect all outcome data on all randomized subjects;
    - Includes a main analysis that keeps subjects in their randomized groups, analyses all available outcome data, and is valid under a named plausible assumption about the missing data;
    - Includes sensitivity analyses that consider a range of plausible alternative assumptions that differ from the main assumption about the missing data;
    - All individuals are included in sensitivity analyses.
- For primary and secondary analyses, the treatment effect on the primary efficacy outcome will be adjusted for the baseline measurement of the outcome.
- Subgroup analyses will be carried out irrespective of whether there is a significant treatment effect on the primary outcome. Their purpose is to supplement evidence from the primary analysis to help to fully characterise the treatment effect. Results from subgroup analyses will be interpreted in this context.
- Analyses will be conducted primarily using R statistical software.

4.2 INTERIM ANALYSES

There are no formal interim analyses specified for this trial. The DSMB periodically review group data for the primary outcome measure and the major subsidiary analyses, including the safety analyses.

4.3 PARTICIPANT DISPOSITION

Flow of patients through the study will be displayed in a standard CONSORT diagram. The report will include the number of screened patients who met the inclusion criteria, the number included, and the major reasons for exclusion of eligible patients. At follow up, the number of patients included, withdrawn, lost to follow up and the number who died within that period will also be reported.

4.4 MULTIPLICITY ADJUSTMENT

No formal adjustments will be undertaken to constrain the overall type I error associated with the secondary, tertiary, and exploratory analyses. Their purpose is to supplement evidence from the primary analysis to more fully characterise the treatment effect. Results from the secondary and tertiary analyses will be interpreted in this context.

4.5 BLIND REVIEW

The persons responsible for developing this statistical analysis plan will be kept blind until after the plan has been signed off, trial close-out is complete, and data lock is done. The results will not be unblinded to the rest of the study team until the final statistical report has been completed.

4.6 DATA SETS TO BE ANALYSED

All data will be analysed by intention to treat, including all randomised participants in the groups to which they were allocated, regardless of adherence.

4.7 PATIENT CHARACTERISTICS AND BASELINE COMPARISONS

In order to assess balance, description of the specified baseline characteristics will be presented for the Oxygen and Air groups (Table 1). Discrete variables will be summarized as frequencies and percentages. Unless otherwise indicated in the tables percentages will be calculated according to the number of patients for whom data are available. Continuous variables will be summarized by use of either mean and standard deviation or median and interquartile range. Durations and time intervals will be summarized by medians, IQRs, and ranges.

Baseline measures for patients will be presented by group for the following variables:
age, gender, body mass index, type of fibrosing ILD, respiratory function (FEV1, FVC and TLCO, absolute and %predicted), 6-minute walk distance (metres), resting SpO_2, nadir SpO_2 on 6MWT (%), % of waking hours with $SpO_2 \leq 88\%$ during daily life, physical activity (average steps/day), Dyspnoea-12, KBILD, Hospital Anxiety and Depression Scale.

4.8 CONCOMITANT THERAPIES

The number of participants in each group who commence LTOT will be reported and compared across randomised arms. The number of participants in each group who commence new pharmacotherapies and their nature will be reported. Outcome data for participants who commence LTOT or new pharmacotherapies will be analysed according to allocated treatment group, as per intention to treat principles.

4.9 ANALYSIS OF THE PRIMARY OUTCOME

The primary outcome is change from baseline in physical activity at 3 and 6 months, measured as steps/day using the Stepwatch Activity Monitor.

4.9.1 MAIN ANALYSIS

The primary analysis will use linear mixed models, with site as a random effect. We will compare change from baseline at 3 and 6 months between groups, with the baseline measurement included in the model as a covariate. The treatment effect will be presented as the mean difference between groups with the corresponding 95% confidence interval. Desaturation during 6MWT, although a stratifying factor in the design, will be modelled as a fixed effect.

The assumption of homogeneity of variance will be examined for each outcome by plotting residuals from the model against fitted values. Where right skewness results in heterogeneity of variance, a log transformation will be applied.

Missing data are anticipated for the primary outcome. The primary outcome is assumed to be missing-at-random (MAR), i.e. it is assumed that the probability of 3 month steps/day data being missing may depend on the values of the observed data, but not on the values of the missing data. Important explanatory and auxiliary variables such as age, type of ILD, respiratory function and extent of exertional desaturation are being collected and will be examined to assess the plausibility of the MAR assumption.

4.9.2 SUBGROUP ANALYSES

No correction for multiple testing in subgroup analysis will be undertaken and multiplicity-unadjusted p-values will be reported, although the number of declared subgroup analyses will be specified in all publications.

The following subgroup analyses will be carried out for the primary outcome:

- Type of fibrosing ILD (IPF vs other)
- Severity of ILD (FVC<50% vs ≥50%predicted)

The effect of this factor will be assessed by adding an interaction term to the main model.

4.9.3 OTHER SENSITIVITY ANALYSES

Not planned.

4.10 ANALYSIS OF SECONDARY OUTCOMES

Changes in secondary outcomes listed in 2.2.4 will be compared between groups using linear mixed models, using the same random and fixed factors as for the primary outcome, including the baseline measure of the outcome as a covariate.

4.10.1 Change in functional exercise capacity assessed by 6-minute walk distance
4.10.2 Change in health-related quality of life evaluated using the St George's Respiratory Questionnaire
4.10.3 Change in health-related quality of life evaluated using the K-BILD Questionnaire
4.10.4 Change in dyspnea measured using the Dyspnea-12 questionnaire
4.10.5 Changes in fatigue evaluated by the Fatigue Severity Scale
4.10.6 Change in anxiety and depression measured by the Hospital Anxiety and Depression Scale
4.10.7 Change in time spent in moderate to vigorous physical activity
4.10.8 Change in sedentary time

4.11 ANALYSIS OF SAFETY OUTCOMES

Adverse events are defined according to the Good Clinical Practice guidelines including treatment emergent adverse events (TEAE's), which are undesirable events not present prior to medical treatment, or that worsen either in intensity or frequency following treatment. AEs of specific interest are defined according to the criteria used in the LOTT trial (6):

- worsening of lung function (absolute FVC decline >5% predicted over the 6-month trial) (7)
- development of resting hypoxaemia (arterial oxygen pressure ≤55 mmHg at rest on room air, or 56–59 mm Hg with evidence of right heart failure)
- exacerbation of fILD
- hospitalization
- death
- bruising or infection at blood draw site
- fainting related to blood draw
- fainting, dizziness, chest pain, ataxic gait, lower limb pain or mental confusion related to 6-minute walk test
- Adverse events related to POC use:
    - burns (from smoking whilst using a POC, using the POC around an open flame or equipment that sparks)
    - nosebleed or dry nose
    - musculoskeletal injury from tripping on a POC

For each category, we will report both the total number of events and the proportion of participants with at least one event. The proportion of participants with at least one event will be compared between randomised groups using generalised linear mixed models (GLMMs) with a binary outcome and the same random effects as for the main analysis. The number of exacerbations and hospitalisations will be compared between groups using GLMMs with a Poisson on Negative Binomial model for the counts. The number of adverse events related to POC use will be compared between groups using similar GLMMs.

4.12 ANALYSIS OF TERTIARY OUTCOMES

4.12.1 Portable oxygen concentrator usage, reported as total hours at 3 and 6 months, and mean hours per day. Differences between randomized groups will be compared.

4.12.2 Reason for cessation of therapy where relevant (patient request, commencement of LTOT, other) will be reported as number and percent of randomized group.

4.12.3 Oxygen saturation in daily life at 3 and 6 months, measured using a Nonin 3150 Wrist Oximeter during waking hours for 2 consecutive days. Reported as % of monitored time with $SpO_2 \leq 88\%$ during daily life and compared between randomised groups.

These outcomes will be analysed using similar models to those used for the primary and secondary outcomes, depending on whether the data are continuous, binary or counts.

4.12.4 We will conduct an exploratory analysis to investigate the impact of portable concentrator usage (mean hours per day) on the primary outcome of mean steps/day using linear mixed models, with POC usage modelled as a random effect.

5 REFERENCES

1. Holland AE, Corte T, Chambers DC, Palmer AJ, Ekstrom MP, Glaspole I, et al. Ambulatory oxygen for treatment of exertional hypoxaemia in pulmonary fibrosis (PFOX trial): a randomised controlled trial. BMJ Open. 2020;10(12):e040798.
2. Flaherty KR, Brown KK, Wells AU, Clerisme-Beaty E, Collard HR, Cottin V, et al. Design of the PF-ILD trial: a double-blind, randomised, placebo-controlled phase III trial of nintedanib in patients with progressive fibrosing interstitial lung disease. BMJ Open Respiratory Research. 2017;4(1):e000212.
3. Demeyer H, Burtin C, Hornikx M, Camillo CA, Van Remoortel H, Langer D, et al. The Minimal Important Difference in Physical Activity in Patients with COPD. PLoS One. 2016;11(4):e0154587.
4. Prasad JD, Paul E, Holland AE, Glaspole IN, Westall GP. Physical activity decline is disproportionate to decline in pulmonary physiology in IPF. Respirology. 2021;26(12):1152-9.
5. Orkin AM, Gill PJ, Ghersi D, Campbell L, Sugarman J, Emsley R, et al. Guidelines for Reporting Trial Protocols and Completed Trials Modified Due to the COVID-19 Pandemic and Other Extenuating Circumstances: The CONSERVE 2021 Statement. Jama. 2021;326(3):257-65.
6. Long-Term Oxygen Treatment Trial Research G, Albert RK, Au DH, Blackford AL, Casaburi R, Cooper JA, Jr., et al. A Randomized Trial of Long-Term Oxygen for COPD with Moderate Desaturation. N Engl J Med. 2016;375(17):1617-27.
7. Maher TM, Corte TJ, Fischer A, Kreuter M, Lederer DJ, Molina-Molina M, et al. Pirfenidone in patients with unclassifiable progressive fibrosing interstitial lung disease: a double-blind, randomised, placebo-controlled, phase 2 trial. The Lancet Respiratory Medicine. 2020;8(2):147-57.

APPENDIX 1: PROPOSED TABLES AND FIGURES

Figure 1: Consort flowchart

Table 1: CONSERVE-CONSORT report of trial modifications due to the COVID-19 pandemic

Table 2: Baseline characteristics by randomized group

Table 3: Primary and key secondary outcomes

Figure 2: Longitudinal mean plot of primary outcome, steps per day

Table 4: Adverse events