Form number	**10.003A**	Version	**2.0**
Title	**Statistical Analysis Plan (SAP)**		

A phase IIa double blind, randomised placebo-controlled trial of Tocilizumab to investigate the effect on health-related quality of life in adults with Long COVID and persistent inflammation

PHOSP-I

STATISTICAL ANALYSIS PLAN (SAP)

Study Title:	A phase IIa double blind, randomised placebo-controlled trial of Tocilizumab to investigate the effect on health-related quality of life in adults with Long COVID and persistent inflammation		
Short Title:	PHOSP-I		
Sponsor Ref:	0862		
IRAS:	1006298		
EudraCT:	2022-003211-29		
ISRCTN:	46454974		
Sponsor:	University of Leicester		
Funded by:	Genentech/Roche		
Protocol Version:	v5.0	Date:	26/11/2025
SAP Version Number:	v1.0	Date:	09/06/2026

		Signature	Date
Prepared by:	Katie Dudman Biostatistician, Robertson Centre for Biostatistics	Signed by: KDudman 190CE24D0B20469...	09 June 2026
Approved by:	Alex McConnachie Director of Biostatistics, Robertson Centre for Biostatistics	DocuSigned by: 3463596312F04EA...	09 June 2026
	Rachael Evans Chief Investigator Professor of Respiratory Medicine / Honorary Consultant Respiratory Physician Centre for Exercise & Rehabilitation Science NIHR Leicester Biomedical Research Centre (Respiratory) Department of Respiratory Sciences, University of Leicester	DocuSigned by: Rachael Evans 44FFF00726424DE...	09 June 2026

CONTENTS

1. INTRODUCTION

1.1. STUDY BACKGROUND

Long Covid is a patient-derived description for persistent symptoms four weeks after SARS-CoV-2 infection. Those admitted to hospital with COVID-19 are at higher risk of persistent symptoms, alongside impairments across mental health, cognition, exercise capacity and health-related quality of life. There is a large population of adults with Long Covid. When PHOSP-I was initiated, there were no specific Long Covid pharmacological treatments.

A major challenge for trials dealing with Long Covid is the wide range of symptoms reported, with no single clear pathophysiological mechanism to target. There is an association between severity of impairments and persistently raised low grade C-reactive protein (CRP) level. This supports the concept of a precision medicine approach to test anti-IL6 (Tocilizumab) in adults with Long Covid and evidence of ongoing persistent systemic inflammation.

1.2. STUDY OBJECTIVES

The primary objective is to compare the effect of 12 weeks of subcutaneous Tocilizumab versus 12 weeks of subcutaneous placebo on health-related quality of life assessed by the Euroqol five-dimension five-level utility index (EQ5D-5L UI).

Secondary objectives are:

- To compare the effect of 12 weeks of subcutaneous (s/c) Tocilizumab versus 12 weeks of s/c placebo on symptoms, mental health, physical performance, daily physical activity, cognitive impairment, multi-organ function and systemic inflammation measured by blood biomarkers.

- To investigate the sustainability of any effect at 12 weeks after medication/placebo cessation on the primary and secondary outcome measures.

1.3. STUDY DESIGN

PHOSP-I is a double-blind, randomised placebo-controlled phase IIa trial of Tocilizumab administered subcutaneously (weekly where participants are ≥100 kg, and alternate weeks where participants are <100kg) for 12 weeks with a further 12 weeks of follow-up.

1.4. RANDOMISATION

Section 7.3 of the current protocol states:

> *Participants will be allocated to treatment groups using a mixed minimisation/randomisation procedure, designed to maintain balance with respect to:*
>
> - *Study site (1-15)*
> - *EQ5D-5L UI (<0.70 / ≥0.70)*
> - *CRP level (≤20mg/L / >20mg/L)*
> - *Participation in the MRI sub-study (Yes / No)*

- *Acute admission/Tocilizumab status (non-hospitalised / hospitalised, no Tocilizumab / hospitalised, with Tocilizumab)*

1.5. Sample Size and Power

Section 10.1 of the current protocol states:

> *The trial is powered based on the EQ5D-5L Utility Index score, a measure of general health-related quality of life. For context, a change of 0.05 units has been shown to be a minimum important difference for interventions for stable long-term conditions [31] and 0.08 units for cancer treatments [32]. In rheumatoid arthritis, a change in EQ5D-5L UI of 0.14 units has been reported for Tocilizumab [49]. Data from the PHOSP-COVID study show a mean (SD) score of 0.72 (0.25) at 5 months and 0.71 (0.25) at one year from hospital discharge, a deficit of 0.14 units compared to pre-COVID levels. The mean change between 5 and 12 months in PHOSP to date is 0.003 (n=585), with a standard deviation (SD) of 0.17 (unpublished). Given the aim of the trial is to identify an intervention with sizeable impact, we propose that an intervention effect of approximately one half of the SD of the change (0.085) between 5 and 12 months would represent a meaningful effect. A two-arm trial would require 64 participants per group with complete data to have 80% power to detect such an effect at a 5% significance level. Allowing for ~15% loss to follow-up, we anticipate that we will need to randomise approximately 76 participants per group.*

The final sample size achieved was 85 randomised participants with EQ-5D-5L available at 12 weeks. Under the same assumptions as above, this gives 80% power to detect an effect size of 0.62σ, where σ is the standard deviation of the change in the primary outcome.

1.6. Interim Analysis and Stopping Rules

No formal interim analyses were planned.

1.7. Statistical Analysis

1.7.1. Objectives

The objective of this SAP is to describe the analyses to be carried out for the PHOSP-I final analysis. This does not include the analysis of the exploratory outcomes (from stored blood and urine samples) or the exploratory outcomes for the sub-study. These will be detailed in a separate statistical analysis plan as needed.

The current version of the protocol at the time of writing is version 5.0 dated 26/11/2025. Future amendments to the protocol will be reviewed for their impact on this SAP, which will be updated only if necessary. This will be documented as part of the Robertson Centre Impact Assessment process.

1.7.2. General Principles

Analyses will follow intention to treat principles, in that participants will be compared according to randomised groups, regardless of compliance with their randomised treatment allocation. Complier Average Causal Effects analyses will be used as sensitivity analyses.

Continuous data will be summarised as the number of observations, number of missing values, mean, standard deviation, median, quartiles (25th and 75th percentiles), and range (minimum and maximum). Categorical data will be summarised as the number of observations, number of missing values, frequencies, and percentages. Data summaries will be presented for all participants and by randomised or treatment group, as relevant.

Efficacy outcomes will be analysed using baseline-adjusted regression models, adjusting for factors used to balance the randomisation. Interaction models will be used to explore subgroup effects. Missing data will not be imputed for main analyses, but multiple imputation will be used in sensitivity analyses.

Baseline and safety data will be reported descriptively, without formal statistical comparison, unless otherwise stated below.

1.7.3. Deviations To Those Specified In Study Protocol

The protocol notes efficacy outcomes will be analysed adjusting for factors used to balance the randomisation, however we will not include participation in the MRI sub-study as an adjustment factor in these analyses since this was stopped in most centres.

The data for some of the outcomes specified in the protocol (daily physical activity, Fried's frailty criteria) will not be available at the time of database lock and will be analysed using similar methods at a later date.

1.7.4. Additional Analyses To Those Specified In Study Protocol

The analysis plan in the protocol is brief, and additional detail is given in this SAP. However, no major additional analyses are planned.

1.7.5. Software

Analyses will be conducted using SAS for Windows v9.4, or higher, and R v4.5.3 or higher.

2. Analysis

2.1. Study Populations

The screened population will consist of all participants included in the screening log.

The randomised population will consist of the subset of the screened population that were randomised and will be analysed as per randomisation group.

The Full Analysis Set (FAS) will consist of the subset of the randomised population who have at least one follow-up visit for efficacy and will be analysed as per randomisation group.

The Per Protocol (PP) population will consist of the subset of the FAS who received at least 75% of expected doses of the randomised medication according to the protocol and completed the primary outcome within the correct time window, and who had no other major protocol deviations, and will be analysed as per randomisation group.

The safety population will consist of all randomised participants who received at least one dose of study drug and will be analysed as per treatment received.

2.2. STUDY STATUS

The following information, which will help to prepare the CONSORT diagram, will be summarised:

- Number of participants in the screening population
- Number of participants that consented
- Number of participants that satisfied the inclusion and exclusion criteria prior to randomisation
- Number and percentage of participants in the randomised population, overall and by randomisation group
- Number and percentage of participants in the FAS, overall and by randomisation group
- Number and percentage of participants in the PP population, overall and by randomisation group
- Number and percentage and percentage of participants in the safety population, overall and by treatment received
- Number and percentage of participants in the FAS and PP populations that received the alternative study treatment from that which they were randomised to, overall and by randomisation group
- Number and percentage of participants that withdrew from the study and the reasons for withdrawal (of those included in the FAS), overall and by randomisation group
- Number and percentage of participants that were lost to follow-up (of those included in the FAS), overall and by randomisation group
- Number and percentage of participants that died (of those included in the FAS), overall and by randomisation group

- Number and percentage of participants that completed the study (of those included in the FAS), overall and by randomisation group

- Number and percentage of participants that withdrew from study medication and the reasons for withdrawal from study medication (of those included in the FAS), overall and by randomisation group

Time to withdrawal from study will be summarised overall and by treatment group, and also displayed using cumulative incidence plots. Time to withdrawal from study medication will be summarised in the same way.

2.3. PROTOCOL DEVIATIONS

Protocol deviations reported by study sites or sponsor monitors will be reviewed by the Chief Investigator and classified as major or minor. These will be summarised overall and by randomised group. Any removal of participants from the analysis populations will be documented as part of database lock procedures.

2.4. BASELINE CHARACTERISTICS

For the FAS, baseline participant characteristics will be summarised overall and by randomised group. The following characteristics will be reported:

- Demographics
    - Age
    - Sex
    - Ethnicity (White/Mixed or Multiple ethnic backgrounds/Asian or Asian British/Black or African or Caribbean or Black British/Other ethnic group/Prefer not to say)
- Participation in MRI substudy
- Hospital admission
    - Acute admission status (Hospitalised/Non-hospitalised)
    - Highest level of respiratory support required after contracting COVID-19 that caused Long Covid, according to WHO clinical progression scale (Asymptomatic: viral RNA detected (Score 1)/Symptomatic: independent (Score 2)/Symptomatic: assistance needed (Score 3)/Hospitalised: no oxygen therapy (Score 4)/Hospitalised: oxygen by mask or nasal prongs (Score 5)/Hospitalised: oxygen by NIV or high flow (Score 6)/Intubation and mechanical ventilation (Score 7)/Mechanical ventilation (Score 8)/Mechanical ventilation and vasopressors, dialysis, or ECMO (Score 9))
    - Medication for acute COVID-19 (Bebtelovimab/Ritonavir-boosted nirmatrelvir (Paxlovid)/Molnupiravir/Sotrovimab/Remdesivir (Veklury)/Dexamethasone/Tocilizumab (RoActemra)/Sarilumab/Baricitinib (Olumiant)/Tofacitinib/Casirivimab and Imdevimab/Other)
    - If hospitalised, Anticoagulation for COVID-19 (Therapeutic/Intermediate/Prophylactic)

- Comorbidities
    - Number of comorbidities (0/1/≥2)
    - Cardiovascular (Previous myocardial Infarction/Ischaemic Heart Disease/Atrial fibrillation/Hypertension/Congestive heart failure/Congenital heart disease/Valvular Heart Disease/Pacemaker or Implantable defibrillator/Peripheral vascular disease/Hypercholesterolaemia or dyslipidaemia/Cerebrovascular accident (stroke) or transient ischemic attack/Other chronic cardiovascular disease)
    - Neurological and Psychiatric (Dementia/Any other chronic neurological disorder/Depression or Anxiety/Chronic fatigue syndrome, fibromyalgia or chronic pain/Any previous treatment with antidepressant medication/Any previous treatment by a mental health professional)
    - Respiratory (Chronic obstructive pulmonary disease/Asthma/Interstitial lung disease/Bronchiectasis/Obstructive sleep apnoea/Obesity hypoventilation syndrome/Pleural effusion/Other chronic respiratory disorder)
    - Rheumatological (Connective tissue disease/Rheumatoid arthritis/Osteoarthritis/Other chronic rheumatological disease)
    - Gastrointestinal (Peptic ulcer disease/Liver disease/Gastro-oesophageal reflux disease/Inflammatory bowel disease/Irritable bowel syndrome/Other chronic gastrointestinal disorder)
    - Metabolic/Endocrine/Renal (Diabetes/Hypothyroidism/Hyperthyroidism/Chronic kidney disease/Other chronic metabolic/endocrine disorder)
    - Previous Malignancy/Haematological (Solid Tumour Malignancy/Leukaemia/Lymphoma/Other chronic haematological disorder)
    - Infectious disease (AIDS/HIV/Chronic viral hepatitis/Mycobacterium Tuberculosis/Other chronic infectious disease)
    - Other
- Vaccination status
    - COVID-19 vaccination status prior to Long Covid
    - COVID-19 vaccination status prior to consent
- Clinical bloods
    - Full blood count (Platelets/Neutrophils/Red Blood Cell Count/Mean Cell Volume/Mean Cell Haemoglobin/Monocyte Count/Eosinophil Count/Basophil Count/Hb/White Cell Count/Lymphocytes/Haematocrit)
    - U&E (Sodium/Potassium/Urea/Creatinine/eGFR)
    - LFTs (Bilirubin/AST/ALT/Alk Phos/Albumin)
    - NT-proBNP
    - Other (C-Reactive Protein/Hepatitis B Core-Antibody/Hepatitis C Surface Antibody/HIV/IGRA)
- Pregnancy test done
    - Serum pregnancy test done
    - Urine pregnancy test done
- Measurements

    - Weight
    - Height
    - BMI
- Physical exam
    - Heart rate
    - SpO_2
    - Diastolic blood pressure
    - Systolic blood pressure
- PROMs (total scores and subscales if relevant)
    - EQ-5D-5L Utility Index (UI)
    - EQ-5D-5L Visual Analogue Scale (VAS)
    - Generalised Anxiety Disorder 7 questionnaire (GAD-7)
    - Patient Health Questionnaire-9 (PHQ-9)
    - FACIT-Fatigue
    - Dyspnoea-12
    - Nottingham Activities of Daily Living scale (NADL)
    - Saint George's Respiratory Questionnaire (SGRQ)
    - Modified Covid-19 Yorkshire Rehabilitation Scale (C19YRSm)
- Physical function
    - Incremental Shuttle Walk Test (ISWT)
    - Short Physical Performance Battery (SPPB)
    - Accelerometry
    - Handgrip strength
    - Fried's frailty criteria
- Organ function
    - Montreal Cognitive Assessment (MoCA)
    - Spirometry
- Availability of samples
    - Whole blood DNA (at any visit)
    - RNA PAXgene
    - Serum
    - Plasma
    - Plasma sodium citrate
    - Urine
    - Stool

2.5. EFFICACY OUTCOMES

All efficacy outcomes will be analysed using the FAS. The primary outcome and selected secondary outcomes will be additionally analysed using the PP population if different to the FAS.

2.5.1. PRIMARY OUTCOME

The primary outcome is the EQ-5D-5L UI (utility index) at 12 weeks. This will be calculated using the crosswalk function recommended by the NICE Decision Support Unit (age and sex conditional based mapping):

https://sheffield.ac.uk/nice-dsu/methods-development/mapping-eq-5d-5l-3l

Summary data will be presented for EQ-5D-5L UI at baseline, at 12 weeks, and for the change, overall and by randomised group. Similar summaries will be presented for the subgroup of the FAS with data available at both baseline and 12 weeks.

The adjusted mean difference in EQ-5D-5L health utility score at 12 weeks between randomised groups will be estimated using a mixed effects linear regression model. Randomised treatment will be included as a binary predictor. Study site will be included as a random effect. Sites with very small numbers of participants may be grouped as needed. Baseline CRP will be adjusted using the logarithm of the continuous measurement, rather than the binary variable used in the minimisation algorithm. Baseline EQ-5D-5L UI will be adjusted for as a continuous variable, rather than the binary variable used in the minimisation algorithm. Acute admission/Tocilizumab status will be adjusted as per the randomisation/minimisation procedure where possible. In the event of small numbers impacting the model, a binary variable (hospitalised or non-hospitalised) may be used instead. The regression coefficient for the treatment effect will be reported as an estimate of the adjusted mean difference between the Tocilizumab and placebo randomised groups, with a 95% confidence interval (CI) and p-value.

2.5.2. SECONDARY OUTCOMES

The secondary outcomes will be summarised and compared between randomised groups using similar methods, transformations of non-normally distributed variables will be considered prior to analysis as required. Specifically:

Outcome	Time Point	Analysis	Adjustment covariates	Population
EQ-5D-5L UI	4 weeks 8 weeks 24 weeks	Linear regression	Random: Site Fixed: other minimisation factors	FAS
EQ-5D-5L VAS	4 weeks 8 weeks 12 weeks 24 weeks	Linear regression	Random: Site Fixed: other minimisation factors Fixed: Baseline EQ-5D-5L VAS	FAS, PP
GAD-7	12 weeks 24 weeks	Linear regression	Random: Site Fixed: other minimisation factors Fixed: Baseline GAD 7	FAS, PP
PHQ-9	12 weeks 24 weeks	Linear regression	Random: Site Fixed: other minimisation factors Fixed: Baseline PHQ-9	FAS, PP
FACIT-Fatigue	12 weeks	Linear regression	Random: Site	FAS, PP

	24 weeks		Fixed: other minimisation factors Fixed: Baseline FACIT-Fatigue	
Dyspnoea-12	12 weeks 24 weeks	Linear regression	Random: Site Fixed: other minimisation factors Fixed: Baseline Dyspnoea-12	FAS, PP
NADL	12 weeks 24 weeks	Linear regression	Random: Site Fixed: other minimisation factors Fixed: Baseline NADL	FAS
SGRQ	12 weeks 24 weeks	Linear regression	Random: Site Fixed: other minimisation factors Fixed: Baseline SGRQ	FAS, PP
C19-YRSm	12 weeks 24 weeks	Linear regression	Random: Site Fixed: other minimisation factors Fixed: Baseline C19-YRSm	FAS, PP
ISWT	12 weeks 24 weeks	Linear regression	Random: Site Fixed: other minimisation factors Fixed: Baseline ISWT	FAS, PP
SPPB	12 weeks 24 weeks	Linear regression	Random: Site Fixed: other minimisation factors Fixed: Baseline SPPB	FAS, PP
Handgrip strength	12 weeks 24 weeks	Linear regression	Random: Site Fixed: other minimisation factors Fixed: Baseline handgrip strength	FAS, PP
MoCA	12 weeks 24 weeks	Linear regression	Random: Site Fixed: other minimisation factors Fixed: Baseline MoCA	FAS, PP
Spirometry	12 weeks 24 weeks	Linear regression	Random: Site Fixed: other minimisation factors Fixed: Baseline spirometry	FAS

Note, where relevant, continuous versions of minimisation variables will be used for model adjustments.

2.5.3. Exploratory Outcomes

The analysis of exploratory outcomes (from stored blood and urine samples) is not covered by this SAP.

2.5.4. Additional Analyses

No additional efficacy analyses are planned.

2.5.5. Sub-group Analyses

Subgroup analyses for the primary outcome will be performed by including the subgrouping variable as a predictor in the primary analysis regression model, and then fitting an interaction

between the subgrouping variable and randomised treatment group. The interaction p-value (from a likelihood ratio test) will be reported. Within-subgroup treatment effects will be reported with 95% CIs and p-values. A forest plot will be generated to illustrate the subgroup effects.

The following subgrouping variables will be assessed:

- baseline age (≥18 years and <40 years, ≥40 years and <60 years, ≥60 years)
- sex at birth (male, female)
- baseline EQ5D-5L UI (<0.70, ≥0.70)
- baseline CRP (≤20mg/L, >20mg/L)
- acute severity of acute illness described by the WHO clinical progression scale (1-6, 7-9)
- time from acute SARS-CoV-2 infection (≤1 year, >1 year)
- BMI (<30 kg/m^2, ≥30 kg/m^2)
- number of comorbidities (<2, ≥2)

Alternative cut-offs (e.g. above/below median) may be applied, instead of or in addition to the subgroups specified above (e.g. due to model convergence issues).

2.5.6. Sensitivity Analyses

Multiple imputation with chained equations will be applied to impute missing primary outcome data. All predictors in the primary analysis model (section 2.5.1), all secondary outcome measures at baseline (section 2.5.2), plus all factors considered as subgrouping variables (section 2.5.5) will be used in the imputation process.

2.5.7. Assumption Checking

Distributional assumptions for regression models will be checked visually, and alternative models (e.g. using data transformations, or alternative link and variance functions) may be used. As far as possible, this will be done prior to unblinding and documented within the SAP. Any changes to model specifications after unblinding will be documented in the final statistical outputs.

2.6. Safety Outcomes

All safety outcomes post randomisation will be reported for the safety population, overall and by treatment received.

2.6.1. Treatment Compliance

In the subset of the safety population that had baseline weight recorded <100kg, we will summarise the following, overall and by treatment group:

- The proportion of doses of IMP given, of those where a dose was planned
- The number of participants who have missed at least one dose and the number of doses missed

- The proportion of doses of IMP that were given out of window

In the subset of safety population that had baseline weight recorded ≥100kg, we will summarise the same information as noted above, and additionally the proportion of participants who had a change in regime.

2.6.2. ADVERSE EVENTS

All adverse event data will be presented separately for those events with an onset date before the date of randomisation, and those events with an onset date on or after the date of randomisation. No statistical testing of adverse event data will be performed.

The number of adverse events will be reported overall and by treatment received. The following characteristics of adverse events will be summarised:

- seriousness
- severity
- outcome
- relationship to study drug
- action taken regarding study drug
- duration (for those events that resolved during the study)
- time from randomisation (for events after randomisation)

Similar summaries will be provided for serious adverse events. The following characteristics of serious adverse events will be summarised:

- seriousness criteria
- severity
- outcome
- relationship to study drug
- action taken regarding study drug
- whether medication was given
- duration (for those events that resolved during the study)
- time from randomisation (for events after randomisation)
- expectedness (for events after randomisation possibly/probably/definitely related to the IMP)

Summaries of the number and percentage of participants experiencing at least one adverse event will be reported for all events, and for events classified by MedDRA system organ class and preferred term. Similar summaries will be provided for the number and percentage of participants experiencing at least one serious adverse event.

The above summaries will be repeated for events that are at least possibly related to the study drug.

All adverse event and serious adverse event data (including narratives) will be provided in data listings.

2.6.3. LABORATORY PARAMETERS

Clinical blood parameters will be summarised overall and by treatment received, at week 12 and week 24. Changes from baseline will also be summarised. Summaries will be presented for:

- CRP
- Platelets
- Neutrophils
- Red Blood Cell Count
- Mean Cell Volume
- Mean Cell Haemoglobin
- Monocyte Count
- Eosinophil Count
- Basophil Count
- Hb
- White Cell Count
- Lymphocytes
- Haematocrit
- Sodium
- Potassium
- Urea
- Creatinine
- eGFR
- Bilirubin
- AST
- ALT
- Alk Phos
- Albumin
- NT-proBNP

2.6.4. CONCOMITANT MEDICATIONS

The number and percentage of participants taking concomitant medications will be summarised overall and by study treatment received (Tocilizumab or placebo), for each of the following classes of medications:

- Analgesia (inc. opiods)
- NSAIDS
- Beta blockers
- Ca2 channel blockers
- ACE inhibitors
- Statin
- Anti-platelets
- Anticoagulants
- Diuretics
- Immunosuppressant
- Antibiotics

- Benzodiazepines
- Antidepressants
- Antipsychotics
- Anti-epilepsy
- Corticosteroids (oral)
- Inhalers containing ICS
- Bronchodilators (any form)
- Anti-fibrotics
- Vaccines
- Diabetes medications (inc. insulin)
- Other

Summaries will be provided for medications being used at baseline, week 12, and week 24. Separate summaries will be provided for medications taken at week 12 which were started after baseline, and for medications taken at week 24 which were started after week 12. Separate summaries will also be provided for any temporary medications taken during the treatment phase, i.e. started after baseline and stopped before week 12.

3. DATA CONVENTIONS

A separate assumptions document (PHOSP_I_Assumptions_vX_X.docx) detailing data derivation rules (e.g. partial dates, missing data imputation) will be created.

4. TABLES AND FIGURES

A dummy report, using fake treatment codes, will be provided for review and approval prior to the database lock.

5. REFERENCES

N/A

6. DOCUMENT HISTORY

This is the first version (v1.0) of the PHOSP-I final analysis SAP, dated 09/06/2026 (the initial creation).