Therapies for Long COVID in non-hospitalised individuals: From symptoms, patient reported outcomes and immunology to targeted therapies (The TLC Study)

Short Study Title	Therapies for Long COVID in non-hospitalised individuals
IRAS Number	296374
Sponsor reference number	1535867
REC reference number	
Protocol version and date	2.0– 08/11/2021

Chief Investigator – Dr Shamil Haroon

Institute of Applied Health Research

College of Medical and Dental Sciences

The University of Birmingham

Edgbaston

Birmingham

B15 2TT

Tel: 0121 414 6890

s.haroon@bham.ac.uk

Deputy-Chief Investigator, Professor Melanie Calvert

Professor of Outcomes Methodology, NIHR Senior Investigator

Director Centre for Patient Reported Outcomes Research

Director Birmingham Health Partners Centre for Regulatory Science and Innovation

Institute of Applied Health Research

The University of Birmingham Edgbaston Birmingham

0121 414 8595

M.Calvert@bham.ac.uk

Sponsor- The University of Birmingham, Birmingham, United Kingdom.

Sponsor Representative – Dr Birgit Whitman

Head of Research Governance and Integrity

Research Support Group

Room 117 Aston Webb Building

The University of Birmingham

Tel: 0121 405 8011

Email: researchgovernance@contacts.bham.ac.uk

Co-investigators

Dr Krishnarajah Nirantharakumar

The University of Birmingham

K.Nirantharan@bham.ac.uk

Dr Joht Singh Chandan

The University of Birmingham

J.S.Chandan.1@bham.ac.uk

Dr Elin Haf Davies

Aparito Limited

elin@aparito.com

Professor Janet Lord

The University of Birmingham

J.M.LORD@bham.ac.uk

Dr Olalekan Lee Aiyegbusi

The University of Birmingham

O.L.Aiyegbusi@bham.ac.uk

Professor Alastair Denniston

The University Hospitals Birmingham NHS Foundation Trust

A.Denniston@bham.ac.uk

Professor David Wraith

The University of Birmingham

D.Wraith@bham.ac.uk

Dr Louise Jackson

The University of Birmingham

L.Jackson.1@bham.ac.uk

Dr Grace Turner

The University of Birmingham

G.Turner.1@bham.ac.uk

Ms Anuradhaa Subramanian

The University of Birmingham

A.Subramanian@bham.ac.uk

Professor Georgios Gkoutos

The University of Birmingham

GkoutosG@adf.bham.ac.uk

Professor Elizabeth Sapey

The University of Birmingham

E.Sapey@bham.ac.uk

Professor Tom Marshall

The University of Birmingham

T.P.Marshall@bham.ac.uk

Mr Gary Price

PPI representative
gary.price3m@gmail.com

Dr Puja Myles
Clinical Practice Research Datalink (CPRD)
Puja.Myles@mhra.gov.uk

Dr Tim Williams
Clinical Practice Research Datalink (CPRD)
Tim.williams@mhra.gov.uk

Mrs Sarah Hughes
The University of Birmingham
s.e.hughes@bham.ac.uk

Dr Kirsty McGee
The University of Birmingham
K.C.McGee@bham.ac.uk

Dr Christel McMullan
The University of Birmingham
c.mcmullan@bham.ac.uk

Professor Steven Marwaha
The University of Birmingham
S.Marwaha@bham.ac.uk

Nikita Simms-Williams
The University of Birmingham
N.Simms-Williams@bham.ac.uk

Study Administrator
Anita Walker
Institute of Applied Health Research
The Learning Centre

The University of Birmingham

Edgbaston

Birmingham

B15 2TT

Mobile: +44 (0) 7912 091 473

Email: a.a.walker@bham.ac.uk

Funder(s)

National Institute for Health Research (NIHR) and UK Research and Innovation (UKRI)

Financial and non-financial support

Funders – Grant number: COV-LT-0013

CI Signature Page

The undersigned confirm that the following protocol has been agreed and accepted and that the Chief Investigator agrees to conduct the study as per GCP guidelines, the Sponsor's SOPs, and other regulatory requirements as applicable.

I agree to ensure that the confidential information contained in this document will not be used for any other purpose other than the evaluation or conduct of the clinical investigation without the prior written consent of the Sponsor.

I also confirm that I will make the findings of the study publicly available through publication or other dissemination tools without any unnecessary delay and that an honest accurate and transparent account of the study will be given; and that any discrepancies from the study as planned in this protocol will be explained.

This protocol has been approved by:

Trial Name:	
Protocol Version Number:	Version: 1.2
Protocol Version Date:	05/07/2021

CI Name:	
Trial Role:	Chief Investigator
Signature and date:	__ __ / ____/ __ __ __ __

Sponsor statement:

Where the University of Birmingham takes on the sponsor role for protocol development oversight, the signing of the IRAS form by the sponsor will serve as confirmation of approval of this protocol.

Version control log

Version Number	Date	Summary of changes
1.2	05/07/2021	

Protocol contributors

Name	Affiliations	Contribution to protocol
Dr Shamil Haroon	The University of Birmingham	Chief Investigator. Responsible for drafting and reviewing the protocol and accompanying materials.
Professor Melanie Calvert	The University of Birmingham	Co- chief Investigator. Responsible for drafting and reviewing protocol, patient-reported outcome strategy, development of the Symptom Burden Questionnaire for Long COVID. Input to design of digital platform. Final review/protocol sign-off
Dr Elin Haf Davies	Aparito	Co-Investigator. Design of the digital platform and use of wearables
Dr Olalekan Lee Aiyegbusi	The University of Birmingham	Co-Investigator. Drafting of participant information sheets, co-ordination of patient input, and usability testing of the Study App.
Dr Grace Turner	The University of Birmingham	Co-investigator, drafted patient-facing documents for TLC main study, TLC sub-study and co-production workshops. Responsible for co-production work package.
Dr Christel McMullan	The University of Birmingham	Co-Investigator. Description of PROs and review of protocol
Mrs Sarah Hughes	The University of Birmingham	Co-Investigator. Validation of the Symptom Burden Questionnaire for Long Covid (SBQ-LC™)

Ms Anuradhaa Subramanian	The University of Birmingham	Co-Investigator. CPRD data acquisition, handling and analysis.
Dr Puja Myles	Clinical Practice Research Datalink	Co-Investigator. Oversight of data linkage and input into observational study designs, approvals for linked data
Dr Tim Williams	Clinical Practice Research Datalink	Co-Investigator. Oversight of recruitment of patients with respect to CPRD operations

Contents

List of abbreviations and definitions

Abbreviation	Term
TLC	Therapies for Long COVID
PROs	Patient Reported Outcomes
CPRD	Clinical Practice Research Datalink
WP	Work Package
PPIE	Patient and Public Involvement and Engagement
HES	Linked Hospital Episode Statistics
UK-CIC	UKRI funded Coronavirus Immunology Consortium
NIHR	National Institute for Health Research
IRSP	Intervention Research Service Platform
CHESS	PHE COVID-19 Hospitalisations in England Surveillance System
PID	Patient Identifiable Data
GCP	Good Clinical Practice
PPE	Personal Protective Equipment
EQ-5D 5L	EuroQol-5D 5 level
FACIT-Fatigue	The Functional Assessment of Chronic Illness Therapy – Fatigue Scale
GAD-2	Generalised Anxiety Disorder -2
PCL-2	Abbreviated Post Traumatic Stress Disorder Checklist – Civilian version
PHQ-2	Patient Health Questionnaire -2
GP	General Practice
SBQ-LC™	Symptom Burden Questionnaire for Long COVID
SARS CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2

STUDY SUMMARY

Study Title	Therapies for Long COVID in non-hospitalised individuals: From symptoms, patient reported outcomes and immunology to targeted therapies (The TLC Study)
Short Title	Therapies for Long COVID in non-hospitalised individuals (TLC Study)
Study Design	Population based cohort
Study Participants	Non-hospitalised individuals with a diagnosis of COVID-19 and symptoms lasting at least 12 weeks ("Long COVID"). Matched control patients: i) individuals without a diagnosis of COVID-19; ii) individuals with a diagnosis of COVID-19 but without long-term symptoms. Co-production workshops: Patients and their family members, healthcare professionals, experts in relevant symptoms and corresponding interventions, regulators, and policymakers (NICE).
Planned Size of Sample	Minimum of 2000 individuals with a diagnosis of COVID-19. Minimum of 500 individuals with Long COVID identified from the main cohort of individuals with COVID-19. Minimum of 500 matched control patients without a diagnosis of COVID-19. Minimum of 500 matched control patients with a diagnosis of COVID-19 but without long-term symptoms (i.e., without Long COVID). Up to 200 individuals from the Long COVID cohort, 50 matched control patients without COVID-19, and 50 individuals with COVID-19 but without long-term symptoms, selected for bio-sampling (sample clusters) and wearables. Co-production workshops: Approximately 40 stakeholders including patients and their family members, healthcare professionals, experts in relevant symptoms and corresponding interventions, regulators, and policymakers (NICE).
Follow-up duration	Monthly for 12 months
Planned Study Period	16 months
Research Question/Aim(s)	To evaluate the symptom burden and underlying pathophysiology of Long COVID syndromes in non-hospitalised individuals. To co-produce an intervention that can be delivered remotely for supporting individuals with Long COVID syndromes. To gain consensus on therapies to recommend, therapies to evaluate and therapies to recommend against using for the clinical management of Long COVID.
Design and methods	**Work package 1** will identify people that had COVID-19 from GP records. Study participants will be invited to use an app or website to report symptoms, quality of life and work capability. **Work package 2** will assess Long COVID symptom burden and define symptom clusters. A sub-cohort will be invited to provide blood and saliva samples to investigate the immune basis for Long COVID and help identify potential therapeutic targets. Wearable devices will be provided to understand the effects of Long COVID on heart rate, oxygen levels, physical activity and sleep.

	Work package 3 will draw upon existing evidence for treatments for Long COVID including drugs and supportive interventions (e.g., for supporting mental health or fatigue). Working with patients, doctors and other experts, consensus statements will be developed on: i) therapies for recommended use in the clinical management of Long COVID; ii) therapies to evaluate and iii) therapies not recommended for use. A virtual supportive intervention will be co-produced with patients and health service providers. Data from work packages 1-3 will also be used to inform the development of a clinical trial that will evaluate the remote-delivered intervention produced during the project, and to plan future evaluation of pharmacological and supportive interventions for Long COVID. This will be done under a separate, linked protocol.

KEY WORDS: Long COVID, COVID 19, Symptoms, Quality of life, Patient Reported Outcomes, Therapies

STUDY PROTOCOL

Therapies for Long COVID in non-hospitalised individuals

1.0 Background

Coronavirus Disease 2019 (COVID-19), caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS CoV-2), is the most significant pandemic since the Spanish Influenza Pandemic of 1918. There have been over 140 million cases worldwide, including over 4.4 million cases in the UK[1]. In the UK, this has resulted in over 460,000 hospital admissions and in excess of 150,000 deaths[1,2].

SARS CoV-2 enters human cells using the angiotensin II receptor, which is widely expressed throughout the body. COVID-19 is consequently a multisystem disorder affecting the lungs, heart, gut, kidneys, brain and skin. A wide range of symptoms have been associated with COVID-19 including breathlessness, fatigue, fever, myalgia, “brain fog” and anosmia.

These symptoms resolve within 12 weeks in most affected individuals. However, in up to 10% of individuals, symptoms of COVID-19 persist, sometimes in a relapsing remitting manner[3]. This can have significant impacts on physical and neurocognitive functioning, quality of life and work capability. Those having symptoms lasting beyond 12 weeks are referred to as experiencing post-COVID-19 syndrome (NICE 2019) or more widely known as Long COVID.

The cause of Long COVID is poorly understood. However, there is evidence from other post-viral syndromes that may be applicable to Long COVID. Furthermore, evidence from existing studies on COVID-19 suggest a significant burden of pathophysiological insults and sequelae such as lung scarring, kidney injury, myocarditis and systemic inflammatory states that may promulgate longer term symptoms. There is also a possibility that autoimmune pathways may be triggered by COVID-19, leading to multi-system inflammatory damage[4].

2.0 Rationale

Approximately 1 in 10 individuals with COVID-19 experience symptoms and impaired quality of life and frequently report experiencing heterogeneous physical and psychological symptoms beyond 12 weeks (“Long COVID”) that can be debilitating. People living with Long COVID have indicated that they are suffering with a

range of symptoms, feel 'abandoned' and 'dismissed' by healthcare providers and receive limited or conflicting advice. The aetiology, pathophysiology and treatments for Long COVID are not well understood, creating an unmet need for the growing number of affected individuals. Although efforts are being made to study Long COVID in hospitalised patients, there is a large unmet need among non-hospitalised individuals. Long COVID may comprise several distinct syndromes yet to be fully characterised.

Existing evidence on Long COVID is based on hospitalised cohorts and non-selected populations that are unlikely to be generalisable to the wider UK population of non-hospitalised individuals. A representative population-based cohort, ideally derived from primary care, is needed to understand the burden of Long COVID, associated disability and impact on work capability. The literature suggests the physical and psychological symptoms of Long COVID are highly diverse and may represent several distinct syndromes [5,6]. The characterisation of these syndromes using real-world data in combination with patient reported outcomes would aid health services to deliver appropriate interventions to meet these health needs.

The aetiology and risk factors for Long COVID also need further investigation as literature suggests a disproportionately higher prevalence of Long COVID among women, older adults and individuals with specific symptom clusters[7]. One possible explanation for this may relate to differing immunological profiles of individuals who go on to develop Long COVID[8]. For example, older adults with COVID-19 show higher levels of senescent T cells[9]. As these are pro-inflammatory, more likely to be autoimmune and tissue-damaging, their persistence could mediate distinct symptoms in Long COVID among these individuals. Understanding the immunological basis of Long COVID would better enable clinicians to offer targeted therapies[10].

There is an urgent need for evidence-based, accessible interventions that are co-produced with patients with lived experience of COVID-19 to better support the growing number of non-hospitalised individuals with Long COVID. We propose to derive treatment recommendations drawing upon existing approaches to the management of post-viral syndromes in addition to a detailed understanding of Long COVID syndromes. There is a pressing need to develop a trial platform that can link symptom, primary care and hospital data with immunological profiling to evaluate interventions for Long COVID in non-hospitalised patients. The Recovery trial and PHOSP-COVID study have demonstrated the strength of platforms where several interventions can be rapidly evaluated in secondary care[11,12]. However, such an infrastructure is still urgently needed for affected individuals in the wider community.

The proposed research will directly address the patient needs highlighted in the NIHR Evidence Report, by the LongCovidSOS campaign, our patient and public involvement group and patient partners, who have co-designed this study[13,14]. Our research will characterise the symptoms, health impacts and underlying pathophysiology of Long COVID and define its component syndromes in non-hospitalised individuals. In addition, we will provide tailored resources to support symptom management and nurse-led support for those with the severest symptoms. We will use our findings to co-produce with patients a targeted intervention for Long COVID, tailored to individual patient need.

There is currently a lack of published data relating to the pathogenesis of Long COVID. The University of Birmingham is internationally renowned for leading the exploration of acute COVID-19 immune profiles. The proposed research will build upon this expertise and provide the first immune profile for patients with Long COVID syndromes.

3.0 Research aims and objectives

3.1 Research aims

To evaluate the symptom burden and underlying pathophysiology of Long COVID syndromes in non-hospitalised individuals and to co-produce a remotely delivered intervention for supporting individuals with Long COVID syndromes. Data will also be used to inform the development of a feasibility study to evaluate the intervention produced during this project, and to plan future evaluations of supportive and pharmacological interventions for Long COVID symptoms. This will be done under a separate, linked protocol.

3.2 Research Objectives

1. To establish a representative population-based cohort of non-hospitalised individuals with COVID-19 and prolonged symptoms ("Long COVID").

2. To develop and validate a new patient-reported outcome (PRO) measure of symptom burden in individuals with Long COVID.

3. To collect validated PROs including symptoms, quality of life and work capability data longitudinally from the representative cohort of individuals with Long COVID using the Aparito Atom5™ platform.

4. To identify distinct symptom clusters ("syndromes") in non-hospitalised individuals with Long COVID and compare these Long COVID syndromes to known post-viral syndromes.

5. To measure immunological, proteomic and miRNA biomarkers to determine the underlying pathogenesis associated with Long COVID syndromes. To measure physical parameters via a wearable device to determine the physical exertions (heart rate intensity and movement) and sleep patterns observed in Long COVID syndromes.

6. To make recommendations on pharmacological and supportive therapies that: i) should be implemented without need for further evaluation; ii) should be evaluated in a clinical trial, and iii) should not be recommended for treating long COVID based on the available evidence.

7. Co-produce a supportive, remotely delivered intervention for individuals with long COVID, to address the symptoms associated with each Long COVID syndrome.

8. To evaluate patient and public involvement in the study.

4.0 Research Outputs and Deliverables

1. To establish a population-based representative cohort of 4000 individuals (minimum 2000) with COVID-19, with a minimum of 500 individuals with Long COVID, 500 matched control patients without COVID-19 (control group 1) and 500 matched control patients with COVID-19 but without Long COVID (control group 2). We will capture detailed data on symptoms, quality of life, work capability and objective health measures for individuals with Long COVID and patients in the matched control groups. This will include the development and validation of a new disease-specific PRO of symptom severity and interference for individuals with Long COVID.

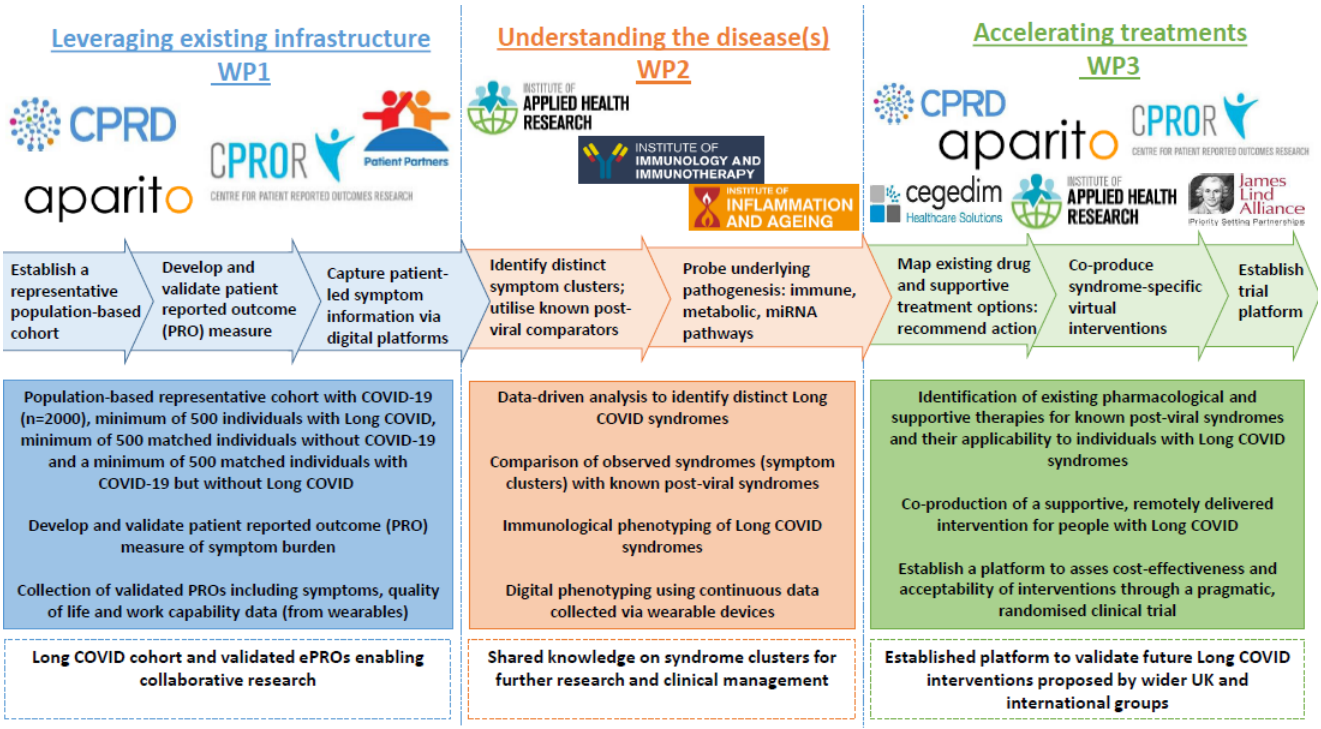
2. To develop a new, validated PRO for the measurement of symptom burden in individuals with Long COVID: the Symptom Burden Questionnaire™: Long COVID.

3. To describe distinct Long COVID syndromes and their epidemiology and characterise their immune, inflammatory and proteomic profile, and physical parameters (heart rate intensity, sleep and movement) providing mechanistic understanding of each syndrome and identifying potential therapeutic targets for future evaluation. This will accelerate the identification of potential therapies for Long COVID that are targeted to each syndrome. Based on symptom clusters we will identify up to 200 individuals from the Long COVID cohort, 50 matched control patients without COVID-19, and 50 matched control patients with COVID-19 but without Long COVID to receive bio-sampling and be provided with wearable devices (Garmin Vivosmart4) to measure heart rate, oxygen saturation, step count and sleep quality (BioWear Sub-Study). This will be used in WP2 to deep phenotype Long COVID syndromes.

4. To co-produce a targeted intervention for Long COVID syndromes that can be delivered remotely on a digital platform. This intervention will have patients and carers at the heart of its design and will be based on the best available evidence on Long COVID. It will be designed to be effective, targeted to address symptoms associated with each specific syndrome, acceptable to patients, and scalable.

We will fulfil these objectives through four work packages (WPs). Work package 4 will be undertaken following further ethical review and approval of a substantial amendment to this study. We will hold regular patient and public involvement and engagement (PPIE) meetings throughout the project to ensure that we obtain and incorporate input from patients and members of the public and will actively engage with members of the LongCovidSOS campaign.

Figure 1: Overview of the TLC Study (including future planned work)



5.0 Study design

WP1: Establish a representative population-based cohort of individuals with Long COVID and collect patient-reported outcomes (PROs) (Objectives 1, 2 and 3)

Overview

We will establish a population-based representative cohort of individuals with COVID-19 and matched controls. We will capture detailed data on symptoms, quality of life, work capability and objective health measures for individuals with Long COVID. This will include the development and validation of a new disease-specific PRO measure that will be used to assess the symptom severity and interference for individuals with Long COVID.

Establish a representative population-based cohort of individuals with Long COVID (Objective 1)

A representative population-based cohort of non-hospitalised individuals with Long COVID will be established using primary care records in the Clinical Practice Research Datalink (CPRD) and recruitment through the Interventional Research Services Platform (IRSP). We will use linked data from the Second-Generation Surveillance PCR-confirmed COVID-19 registry to identify all confirmed COVID-19 patients in CPRD. Using linked Hospital Episode Statistics (HES) data we will identify and exclude hospitalised patients. Further details on each of the data sources is provided (Box 1).

All potential cases will be reported to the Clinical Practice Research Datalink's (CPRD) Interventional Research Services Platform (IRSP). This is a web-based research platform linked to CPRD's anonymised primary care database of coded clinical electronic health care data from approximately 1700 practices across the UK. The platform enables primary care staff to review pre-selected patient lists, confirm eligibility and invite patients to join a study. The selection criteria are applied within the CPRD IRSP, based on the predefined patient selection strategy. A delegated health professional at each practice can subsequently access the pseudonymised patient list, re-identify the patients, and review the clinical record to a set of screening questions. For patients that meet selection criteria, the primary care team will send out an invitation letter informing them about the study and containing key links and information required to enable enrolment onto the study.

Via IRSP, all potential cases will be presented to participating general practices to review their suitability for study recruitment. Those assessed as suitable for study recruitment will be sent an invitation letter to participate in the study. We aim to recruit a minimum of 2000 individuals with COVID-19. Individuals from underrepresented sociodemographic groups (e.g., Black and Minority Ethnic (BAME) and those from areas of high socioeconomic deprivation) will be oversampled. These individuals with COVID-19 (minimum n=2000) will be invited to report symptoms, quality of life and work capability through the Aparito Atom5™ digital platform. Data collected from this cohort of patients with COVID-19 will be used to identify a sub-cohort of patients with long-term symptoms of Long COVID (minimum n=500, aiming for n≥750).

Patients with COVID-19 will be propensity score matched to patients without a COVID-19 diagnosis (minimum n=500), who will similarly be invited to participate in the study following eligibility checking by their general practice. This will constitute control group 1. Propensity scores will be estimated for the likelihood of COVID-19 using logistic regression, including the following data obtained from CPRD Aurum: age, sex, socio-economic status, body mass index (BMI), smoking status, comorbidities and risk factors for COVID-19 identified by Public Health England. A second control group of at least 500 participants will be derived from those with COVID-19 who do not report long-term symptoms (i.e., do not have Long COVID).

Box 1. Data sources to derive the Long COVID cohort and assess patient reported outcomes

Clinical Practice Research Datalink (CPRD)

Clinical Practice Research Datalink (CPRD) is a real-world research service supporting retrospective and prospective public health and clinical studies. CPRD is jointly sponsored by the Medicines and Healthcare products Regulatory Agency and the National Institute for Health Research (NIHR), as part of the Department of Health and Social Care. CPRD collects anonymised patient data from a network of GP practices across the UK. Primary care data are linked to a range of other health related data to provide a longitudinal, representative UK population health dataset. The data encompass 60 million patients, including 16 million currently registered patients.

CPRD primary care data is linked to Hospital Episode Statistics, (HES) COVID-19 Hospitalisation in England Surveillance System (CHESS) and Second-Generation Surveillance (SGSS) data via a unique patient pseudo-identifier which allows this data to be linked together to create a combined dataset for analysis.

Second-Generation Surveillance System (SGSS)

SGSS is the national laboratory reporting system used in England to capture routine laboratory data on infectious diseases and antimicrobial resistance. The SARS-CoV-2 testing started in UK laboratories on 24/02/2020, with the SGSS data reflecting testing (swab samples, PCR test method) offered to those in hospital and NHS key workers (i.e., Pillar 1). The linked SGSS data only contains positive tests results. The primary care data itself includes both positive and negative Pillar 2 PCR and other RNA detection COVID-19 test results and it is estimated that over 80% of all such tests undertaken in England are being matched to GP records.

Hospital Episode Statistics (HES)

HES Admitted Patient Care (HES APC) data contains details of all admissions to, or attendances at English NHS healthcare providers. It includes private patients treated in NHS hospitals, patients who are resident outside of England and care delivered by treatment centres (including those in the independent sector) funded by the NHS. All NHS healthcare providers in England, including acute hospital trusts, primary care trusts and mental health trusts provide data. HES APC data includes the complete set of hospital episode information (admission and discharge dates, diagnoses (identifying primary diagnosis), specialists seen under and procedures undertaken) for each linked patient with a hospitalisation record. In addition, Augmented care data (intensive and/or high dependency levels of care) and Maternity data are available.

COVID-19 Hospitalisation in England Surveillance System (CHESS)

CHESS was established by Public Health England (PHE) across all NHS Trusts in England on 15/03/2020 to collect epidemiological data on COVID-19 infection in persons requiring hospitalisation and ICU/HDU admission. It has more details over HES including laboratory test results for COVID-19 and other respiratory viruses, mechanical ventilation, complications like pneumonia, COVID-19 specific risk factors and treatment.

Development and validation of a new patient reported outcome (PRO) measure of symptom burden in individuals with Long COVID (Objective 2)

We will follow international best practice guidance in the development of PRO measures to construct and evaluate the psychometric measurement properties of a new PRO measure that will assess symptom severity and interference in the population of individuals with Long COVID, the Symptom Burden Questionnaire™ for Long Covid (SBQ-LC™)[15–20].

Ethical approval for SBQ-LC™ development has been granted by the University of Birmingham based on a separate linked protocol with participants recruited via non-NHS routes (Reference ERN_21-0191). Evaluation of the SBQ-LC™'s psychometric measurement properties will be undertaken utilising study data collected via the Aparito Atom5™ digital platform collected in WP1.

Collection of validated patient-reported outcomes (PROs) including symptoms, quality of life and work capability data longitudinally from the representative cohort of individuals with Long COVID using Aparito Atom5™ (Objective 3)

Aparito Atom5™ is a digital platform with a patient facing mobile application (iOS and Android) and online portal. It has a number of capabilities that make it ideally suited to study Long COVID. These include the capability to: track quality of life and work capability with the use of electronic PROs, integrate wearable devices, allow direct communication via healthcare professionals, deliver tailored notifications and alerts with links to relevant support services, and convey data in real-time to clinicians and researchers.

The Aparito Atom5™ app is available for use on Android and iOS operating. Atom5 ™ is a highly configurable, disease agnostic and multilingual platform designed to support research and clinical trials. It has five components: patient facing desktop portal, patient facing app (iOS and Android), clinician/study staff-facing desktop portal, hardware/wearable integration, and data analytics dashboard. Each deployment is GxP compliant. Aparito is accredited for Quality Management System (QMS), ISO 13485 and Information Security Management System (ISMS) ISO 27001. The online portal can be accessed through an internet browser.

Participants will be e-consented to use the Atom5™ platform and asked about the presence of comorbidities, established Long COVID symptoms, quality of life and work/education capability using validated tools via a study specific smartphone app. They will also be asked if they have previously been hospitalised for COVID-19 and if they have received a vaccine for SARS CoV-2.

Outputs and deliverables: A population-based representative cohort of individuals with Long COVID and two groups of propensity score matched control participants [(i) without COVID-19, (ii) with COVID-19 but without Long COVID]. This will include detailed data on symptoms, quality of life, work capability and objective health measures for individuals with Long COVID and matched controls. We will also have validated of a new disease-specific PRO of symptom severity and interference for individuals with Long COVID.

WP2: Analysis and confirmation of distinct Long COVID syndromes (Objectives 4 and 5)

Overview: Long COVID may be a heterogeneous group of post-viral syndromes affecting multiple body systems and pathophysiological pathways. It is important to describe the heterogeneity of symptoms and identify whether there are distinct clinical phenotypes. This will enable identification of individuals with distinct syndromes for further detailed study, including understanding differences in immunological response, and to develop targeted approaches.

Data-driven analysis to identify distinct Long COVID syndromes (Objective 4, part A)

We will define distinct symptom clusters (syndromes) using linked data from CPRD and Atom5™ platforms. We will employ clustering algorithms that have been shown to work well in medical settings, with binary and mixed variables: (1) complete linkage hierarchical clustering; (2) probabilistic c-means cluster analysis; (3) density peaked clustering analysis. We will evaluate algorithm performance and the optimal number of clusters (e.g. by internal validation against a holdout dataset using metrics such as Gap statistic and silhouette index).

We will describe the epidemiology of patients within each Long COVID syndrome as well as health service use (primary care consultations, prescriptions, A&E attendances and hospital admissions) and mortality risk and compare these to both sets of matched control patients. We will identify whether Long COVID syndromes are

associated with particular population groups such as BAME groups, patients with long-term conditions, and other vulnerable population groups identified by Public Health England.

Comparison of observed syndromes (symptom clusters) with known post-viral syndromes (Objective 4, part B)

We will undertake a rapid review describing post-viral syndromes. The data will be synthesised to describe the post-viral syndromes that resemble the Long COVID syndromes identified in our data-driven cluster analysis. We will build on a rapid scoping review that we have undertaken to inform this proposal, on how post-viral syndromes from previous pandemic coronaviruses map onto the currently described complications experienced by those with COVID-19.

Immunological phenotyping of Long COVID syndromes (Objective 5, part A)

Long COVID is associated with both systemic and organ-specific sequelae, including changes in the central nervous system. These are reminiscent of autoimmune phenomena arising following infections. Often these are transient; however, in some individuals they persist and generate post-infectious autoimmune sequelae. To study this, we will invite patients to obtain bio samples (saliva and blood) for a minimum of 250 patients with Long COVID and 50 patients in each of the control groups. We will measure common autoantibody profiles and use both antibody and T cell screens to identify specific self-antigens associated with long COVID, which may allow the identification of potential immunotherapies.

The extended hyperinflammation and immune activation seen with COVID-19 is also likely to accelerate the ageing of the immune system, including the generation of pro-inflammatory, tissue damaging senescent T cells. We will assess this in a sample of each of the symptoms clusters in the Long COVID cohort, and both sets of matched control patients, to determine if there is a preponderance amongst one or more syndromes. Their presence will suggest the utility of interventions to reverse immune ageing such as use of rapamycin analogs.

We will also look for the presence of potential diagnostic biomarkers in saliva and blood, focussing on small non-coding RNAs in saliva (sncRNAs) and over 7000 serum proteins using a proteomics platform (SomaScan®)[21] to define dysfunction of a range of body systems. The main sncRNAs of interest are micro RNAs (miRNAs). These are RNA molecules (20-25 nucleotides) that regulate gene expression post-transcriptionally by either repressing translation or inducing degradation of messenger RNA. They are conserved across species, and are particularly stable in an extracellular environment, making them excellent candidates for use as biologic indicators of different physiologic states. The human genome encodes over 2000 miRNAs[22]. They are produced by a wide variety of cells in different organs and secreted into the blood and other bodily fluids where they can exert biological effects. They are believed to control the activity of 60% of all of our genes[22]. MiRNAs have been shown to play a central role in many biological processes including: the cell cycle, cell metabolism, apoptosis, immune responses and cancer progression[23]. Amongst other roles, sncRNAs have been demonstrated to be very accurate biomarkers of acute and chronic brain injury[24].

Digital phenotyping (Objective 5, Part B)

Continuous data collected on physical parameters such as exertions (heart rate, intensity and movement) and sleep patterns via the wearable devices integrated to the digital phenotyping app (Atom5™), will be used to describe patients with distinct Long COVID syndromes and create their physiological profile providing an understanding of the physical activities patterns linked to sleep, heart rate and intensity of physical movement.

Outputs and deliverables: We will describe distinct Long COVID syndromes and how they map onto the post-viral syndromes from previous pandemic coronaviruses. We will describe the epidemiology of Long COVID syndromes and characterise their immune, inflammatory, proteomic and physiological profile, providing mechanistic understanding of each syndrome and identifying potential therapeutic targets for future evaluation.

WP3: Review of potential therapies for Long COVID syndromes and co-production of a remotely delivered supportive intervention for individuals with Long COVID (Objectives 6 and 7)

Overview: We will identify potential pharmacological and supportive (non-pharmacological) therapies for managing individuals with Long COVID based on evidence from other post-viral syndromes, in combination with the evidence generated from WP2. We will work with patient partners and clinicians to co-produce a supportive intervention providing targeted support for managing long COVID syndromes that can be delivered remotely via the Atom5™ platform.

Identification of existing pharmacological and supportive therapies for known post-viral syndromes and their applicability to individuals with Long COVID syndromes (Objective 6)

We will conduct a rapid review evaluating the management of post-viral syndromes. We will identify current pharmacological and supportive interventions targeted at ameliorating common symptoms associated with impaired quality of life in similar post-viral syndromes (see Annex) in addition to treatments currently being considered for the management of Long COVID. We will map therapies to each of the Long COVID syndromes identified in WP2 as a foundation for discussion with experts and stakeholders in a series of participatory workshops (see next section).

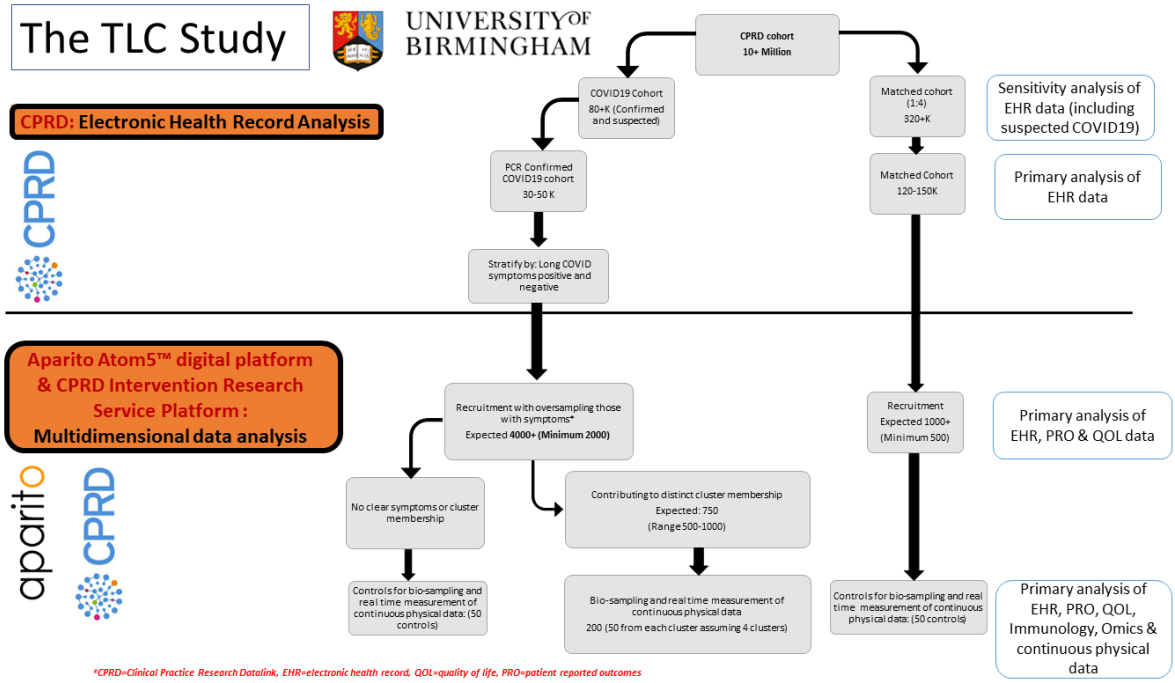
Co-production of a supportive, remotely delivered intervention for people with Long COVID (Objective 7)

We will apply person-centred design principles to co-produce a supportive, remotely delivered intervention with key stakeholders:patients and their family members, healthcare professionals, experts in relevant symptoms and corresponding interventions, regulators, and policymakers (NICE).

Co-production will be guided by the "building blocks of co-production" framework[25]. A series of expert consensus workshops will be convened to make recommendations on: 1) supportive/pharmacological therapies to be used in each Long COVID syndrome; 2) therapies to be evaluated in a clinical trial; 3) therapies that should not be recommended, based on the cumulated evidence from the rapid systematic review and primary findings from WP2. The consensus workshop will follow the James Lind Alliance final priority setting method[26], based on Nominal Group Technique. In addition, decision making will be guided by the APEASE criteria[27]. Particular consideration will be given to non-pharmacological therapies (e.g., physiotherapy, breathing techniques, sleep hygiene) that can be delivered digitally and at scale.

Intervention co-production workshops will then be held to determine which of the agreed interventions are suitable for or could be adapted for remote delivery.

Outputs and deliverables: A co-produced targeted intervention for Long COVID syndromes that can be delivered remotely on a digital platform.



6.0. Participant selection and Eligibility

6.1 Population-based representative cohort of individuals with COVID-19 (minimum 2000 participants)

6.1.1 Eligibility Criteria

Non-hospitalised individuals with a positive RT-PCR test for SARS-CoV-2 in SGSS data linked to CPRD Aurum. Those from that cohort reporting symptoms associated with COVID-19 at least 12 weeks from the diagnosis will be defined as having Long COVID.

6.1.2 Inclusion criteria

Non-hospitalised individuals aged 18 years and over with a SARS CoV-2 RT-PCR result in SGSS data linked to CPRD Aurum, and willing to provide informed consent and to undertake the protocol activities.

6.1.3 Exclusion criteria

Individuals without a positive SARS CoV-2 RT-PCR test result in CPRD Aurum.

Individuals with a hospital admission within 28 days of a positive SARS CoV-2 RT-PCR result in CPRD AURUM.

Individuals aged under 18 years.

Individuals unable/unwilling to provide informed consent.

Individuals unable/unwilling to undertake the protocol activities.

Individuals deemed appropriate for exclusion by their GP (e.g., on a palliative care register).

6.2 Matched controls - Group 1 (minimum of 500 participants)

6.2.1 Eligibility Criteria

Propensity score-matched individuals without a diagnosis of COVID-19 (suspected or confirmed) in CPRD Aurum and without a positive SARS CoV-2 RT-PCR result in linked SGSS data and without a record of hospital admission within 28 days of the index date will comprise Control Group 1. For propensity score matching we will use covariates such as sociodemographic characteristics, lifestyle and metabolic profile measures and presence of comorbidities, especially those identified as risk factors for COVID-19.

6.2.2 Inclusion criteria

Individuals aged 18 years or over without a diagnosis of COVID-19 (suspected or confirmed) in CPRD Aurum and without a positive SARS CoV-2 RT-PCR result in linked SGSS data and able to provide informed consent and undertake the protocol activities.

6.2.3 Exclusion criteria

Individuals with a diagnosis of COVID-19 in CPRD Aurum or positive SARS CoV-2 RT-PCR result in SGSS.

Individuals coded in their primary care record as having suspected COVID-19.

Individuals aged under 18 years.

Individuals unable or unwilling to provide informed consent.

Individuals unable or unwilling to undertake the protocol activities.

Individuals deemed appropriate for exclusion by their GP (e.g., on a palliative care register).

6.3 Matched controls - Group 2 (50 participants)

6.3.1 Eligibility criteria

Participants taking part in the study who have had a positive SARS CoV-2 RT-PCR test result in SGSS data linked to CPRD but not reporting symptoms on ATOM5™ beyond 12 weeks from the positive RT-PCR test result will comprise Control Group 1. These control participants will be propensity score-matched to participating Long COVID patients.

6.3.2 Inclusion criteria

Participants taking part in the study who have had a positive SARS CoV-2 RT-PCR test result in SGSS data linked to CPRD but not developed Long COVID symptoms based on the data reported in ATOM5 and able to provide informed consent and to undertake the protocol activities.

6.3.3 Exclusion criteria

Participants unable/unwilling to provide informed consent.

Participants unable/unwilling to undertake the protocol activities.

Participants deemed appropriate for exclusion by their GP (e.g., on palliative care register).

6.4 Bio-sampling and wearables (BioWear) cohort (200 Long COVID participants, 50 participants from Control Group 1 and 50 participants from Control Group 2)

6.4.1 Eligibility Criteria

Participants in the study who have self-reported symptoms associated with COVID-19 more than 12 weeks since a positive SARS CoV-2 RT-PCR test or COVID-19 diagnosis. This will include 50 randomly sampled participants from each Long COVID symptom cluster, up to a maximum of 200 Long COVID participants. 50 participants randomly sampled from Control Group 1 and 50 participants randomly sampled from Control Group 2 will also be eligible.

6.4.2 Inclusion criteria

Participants able to provide informed consent who have consented to being contacted for the biosampling/wearable study and able to undertake the protocol activities.

6.4.3 Exclusion criteria

Participants who have not previously consented to be contacted for bio-sampling/use of wearables.

Participants unable/unwilling to provide informed consent.

Participants unable/unwilling to undertake the protocol activities.

6.5 Co-production workshops

6.5.1 Eligibility Criteria

Patients and their family members, healthcare professionals, experts in relevant symptoms and corresponding interventions, regulators, and policymakers (NICE).

6.5.2 Inclusion criteria

Individuals able to provide informed consent who have an interest in treatments for Long COVID, including people who have experienced Long COVID and their family members, friends or carers; healthcare professionals; experts in symptom management and potential treatments relevant to Long COVID; and academic researchers.

6.5.3 Exclusion criteria

Individuals unable/unwilling to provide informed consent.
Individuals unable/unwilling to attend a co-production workshop.

6.5 Participant identification and recruitment

6.5.1 Population-based representative cohort of individuals with Long COVID

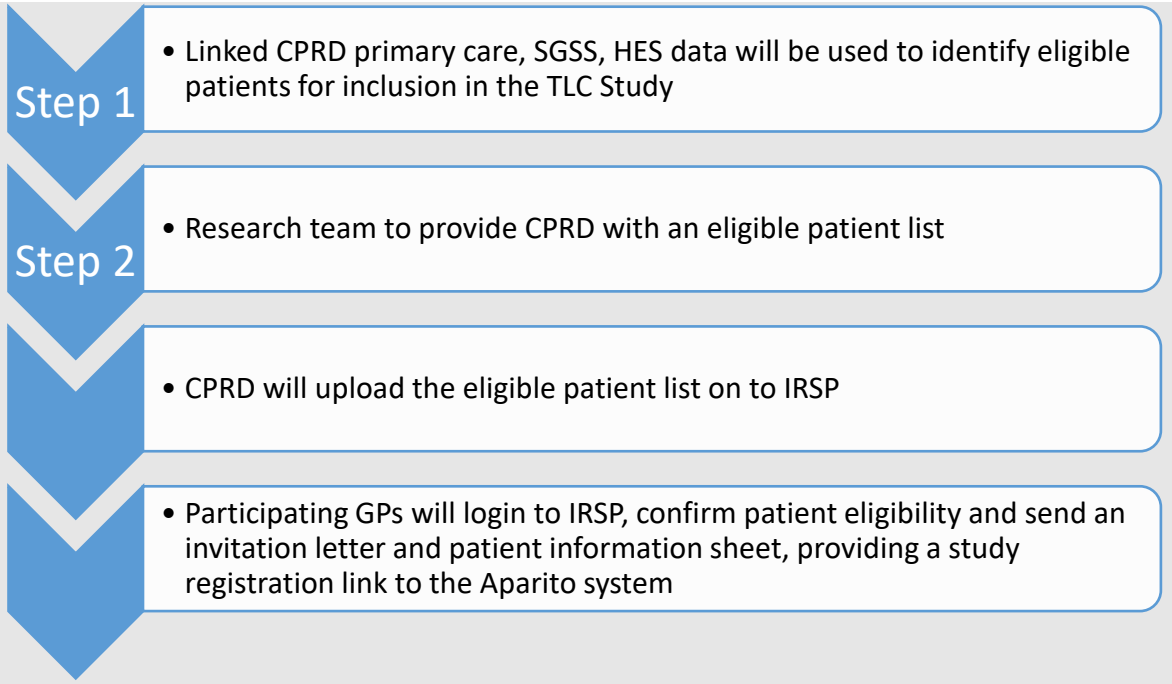
Researchers from The University of Birmingham will use linked data from the Second-Generation Surveillance (SGSS) data that contain SARS CoV-2 positive RT-PCR results to identify patients with confirmed COVID-19 in CPRD Aurum. Using linked Hospital Episode Statistics (HES), the research team will identify and exclude patients who have been hospitalised within 28 days of a positive SARS CoV-2 RT-PCR result or COVID-19 diagnosis).

The list of CPRD Aurum IDs of patients identified as potentially eligible (as above) by the research team will be provided to CPRD. The list of suitable patients will then be uploaded onto the CPRD Interventional Research Services Platform (IRSP). IRSP will generate a unique "username" (Global ID) for each patient on that list, which will be shared with Aparito to assist with verifying participants registering on the Atom5™ app. CPRD will inform participating GP practices that the patient list is available for them to review on by logging onto IRSP. GPs will identify the patients using the corresponding EMIS ID and screen them for their suitability for this study. Researchers at the University of Birmingham will not have access to Patient Identifiable Data until they have consented to participate on Atom5TM.

GPs will then invite eligible patients to take part via invitation letters, text message or email which will contain the patient information sheet (PIS) and a username and unique one time password for the Aparito Atom5 ™ platform. Patients will then log on Atom5 ™ and provide e-consent (which includes confirming their eligibility) and log data on their symptoms and patient-reported outcomes (PROs). From this data we will identify individuals with prolonged symptoms (>12 weeks) associated with COVID-19, which will form the long COVID cohort.

Participants will be provided a £10 Amazon voucher code via Atom5™ upon completing the baseline questionnaires and a further £10 Amazon voucher upon completing the final study questionnaires at 12 months of follow-up.

Figure 2.0 Participant recruitment



6.5.2 Matched controls

Sampling from the CPRD Aurum database, we will generate a propensity score matched control population of 500 individuals without a positive SARS CoV-2 RT-PCR result (Control Group 1) and 500 individuals with COVID-19 without Long COVID (Control Group 2), to assess the additional burden of Long COVID symptoms beyond what is observed in the infected population. Propensity scores will be generated using data on sociodemographic factors, lifestyle and metabolic profile measures and comorbidities. Matching will be done using the nearest-neighbour algorithm without replacement, considering callipers of width equal to 0.1. The index date for matching will be 1st January 2020. The CPRD Aurum ID for the matched control patients within Control Group 1 will be provided to the CPRD IRSP. CPRD will inform participating GP practices of these patients by providing the corresponding EMIS codes and will ask them to check their suitability for recruitment. The list of suitable patients will then be recorded on IRSP with the generation of a unique Global ID. Researchers at the University of Birmingham will not have access to Patient Identifiable Data until they have consented to participate on Atom5TM. Participating GP's will invite eligible patients to take part via invitation letters, text message or email which will contain a copy of the PIS and a unique on-boarding code. Patients willing to take part will be instructed to use the on-boarding code to register on the Atom5™ platform.

6.5.3 Bio-sampling and wearables (BioWear) sub-cohort

Researchers will identify distinct symptom clusters (syndromes) using linked data from CPRD and data collected via the Atom5™ platform. A random sample of up to 300 participants (200 participants with Long COVID, with 50 sampled from each of the four most prevalent Long COVID syndromes, and 50 participants from each of the two control groups) who have previously consented to be contacted for bio-sampling will be identified and invited to take part in the BioWear sub-study. Selected participants will be sent an invitation to participate and the BioWear sub-study PIS via the Atom5TM app, SMS message or by post for those who have expressed having difficulty accessing digital platforms. Those individuals showing an interest in taking part will be contacted by a member of the research team to arrange a date and time for blood and saliva sample collection and to be provided with a wearable device (Garmin Vivosmart4).

6.5.4 Co-production cohort

We will invite individuals who have an interest in treatments for Long COVID to take part in a series of expert consensus and co-production workshops, including:

- People who have experienced Long COVID and their family members, friends, or carers.
- Healthcare professionals.
- Experts in symptoms and potential treatments relevant to Long COVID.
- Regulators.
- Policy makers.

Participants will be recruited from advertisements (distributed in newsletters and social media), word of mouth/known contacts (snowballing), patient support groups (including LongCOVIDSOS), research networks, and Long COVID clinics. Those expressing an interest in participating in the workshops will be given or sent the PIS

electronically or by post prior to the event. In addition, the research team will explain the purpose of the research, the format of the workshop and invite participants to ask questions at the beginning of each workshop.

7.0 Consent

7.1 Main study cohort and matched controls

Participants identified as eligible by their GP and who have shown an interest in taking part in the TLC study will be e-consented via the Aparito Atom5 ™ app or website, or by telephone. E-consent will comprise completion of the participant consent form on the Atom5 ™ app or website and using DocuSign to provide an electronic signature. When this is done by telephone, the researcher taking consent will talk through the consent form with the participant and complete the participant's consent form on Atom5™. Participants will have the option to consent to be contacted for future bio-sampling and wearable monitoring devices, if eligible, and other future related studies. Consent forms will be stored within the Aparito Atom5™ platform. For participants who opt to consent by telephone, a copy of the completed consent form will be sent by post by the research nurse. Both online and telephone support for participants will be provided during the consent process.

7.2 Bio-sampling and wearable device

Participants selected for bio-sampling and provision of the Garmin wearable device (TLC BioWear Sub-Study) will be asked to read the PIS and provide e-consent via Atom5™. They will be able to contact the research team to clarify any aspects of the PIS and consent form using the contact details provided in the PIS. Consent will be verified and co-signed by a member of the research team at the bio-sampling appointment. A copy of the consent form will be stored within the Aparito Atom5™ platform. For participants who opt to consent by telephone, a copy of the completed consent form will be provided at the appointment or sent by post. If for any reason it is not possible to take e-consent due to device failure, consent will be taken on paper and the completion of the consent process will be recorded on Atom5™ as soon as possible.

7.3 Co-production

Potential participants will be given or sent the PIS prior to the workshop. Individuals who register to take part in a workshop will be required to provide e-consent via the Aparito Atom5 ™ platform. E-consent will comprise completion of the participant consent form on the Atom5 ™ platform website and using DocuSign to provide an electronic signature. If an individual is unable to provide consent electronically, consent will be taken by telephone. When this is done by telephone, the researcher taking consent will talk through the consent form with the participant and complete the participant's consent form on Atom5™.

Consent will include consent to the discussions being audio recorded and the meeting being photographed. Photographs will be taken by a researcher who will be provided with a list of participants who have not consented to their photographs being taken. During in-person workshops, all participants will be wearing name badges which will facilitate identification of participants. Furthermore, the researcher will announce when photographs are being taken.

8.0 Study procedures

8.1 Population-based representative cohort of individuals with long COVID and matched controls

Eligible individuals who wish to participate will download the Aparito Atom5™ app or visit the website and register using their unique one-time use on-boarding code. They will then be asked to provide electronic consent on the app or website, which includes confirming that participants with a positive SARS CoV-2 RT-PCR result have not been hospitalised within 28 days of being diagnosed with COVID-19. Control participants will be asked if they have ever had COVID-19 to check eligibility for inclusion as a control subject. Once the consent form has been completed, participants will be asked to complete a brief demographic and clinical questionnaire. The questionnaire will ask for the following details: name, sex, date of birth, contact number, postal address, ethnic group, GP address, height, weight, smoking status, pregnancy status, SARS CoV-2 vaccination status, confinement (shielding) status, care home residence, occupation, occupational status (in work, furlough, etc.), and over-the-counter medications used to manage symptoms. Full details of these questions are provided in the appendix (Baseline data collection in Atom5™) and in Table 1.0.

Participants will then be invited to report the presence of established Long COVID symptoms and quality of life using validated tools at monthly intervals for a period of 12 months. At each time point, participants will be asked to confirm their willingness to continue to participate in the study. Those wishing to withdraw will be asked to complete a withdrawal form and advised that any data collected up to that point will need to be retained as it will have been used in aggregated data analyses.

As requested by our PPI group, we will ensure that people who are unable to use the app or website (due to no access to a smartphone or internet connection, patient preference, language, or other requirements) can participate through nurse interview-led appointments (telephone or video conference) with translators as required. Information and contact numbers will be provided on the PIS. Alternative language versions of each PRO will be provided where validated translations exist (see table below) The translated questionnaires are currently being programmed on the app.

	English	Bengali	Polish	Welsh	Punjabi	Gujarati	Urdu
EQ5D5L	X	X	X	X	X	X	X
FACIT F	X	X	X		X	X	X
PHQ-2	X	X	X		X	X	
GAD-2	X		X		X	X	X
PCL-2	X						
modified MRC Dyspnoea Scale	X						
COVID-19 core outcome measure for recovery	X						
SBQ-LC	X						

All participants will be advised where to seek additional support should completion of the PROs raise any concerns regarding their well-being. Proxy-completion will be permitted via the app or website, or via telephone if the participant is unable to complete the PRO and will be documented on the app or website, or via the clinical portal.

Participants will be sent push notification reminders each month to complete the questionnaires on the day completion is due, at 2 days and again at 6 days. Both research personnel and patients will be provided with training resources and telephone support to help minimise missing data. Aparito will provide up to a maximum of two training sessions to the site staff online. A training video will be made available for those that need training after those two sessions. Finally, the site staff will be made aware of the IT support procedures and how to contact Aparito as needed. Aparito will be responsible for providing first line support only to the clinical team via UoB during the duration of the project. Training leaflets outlining how to download and use the app will be provided to staff and patients.

The health service use (primary care consultations, prescriptions, A&E attendances and hospital admissions) and mortality of patients from each Long COVID syndrome will also be described using data from CPRD Aurum and linked data from Hospital Episode Statistics and Office for National Statistics mortality statistics (via CPRD). We aim to use the most up-to-date healthcare data available from CPRD available up to 12 months of follow-up from initial registration on Atom5™.

8. 2 Clinical support

The majority of symptoms of long COVID, while potentially debilitating, are not associated with acute medical emergencies. However, some symptoms of long COVID can be distressing, and also can mirror symptoms of potentially dangerous medical conditions, such as a blood clot or stroke. The TLC study cannot provide overarching medical care for participants, and their care remains with their usual care providers.

To provide support and guidance, access to nurse-led clinical support (during usual working hours of 8am - 5pm, Monday-Friday (excluding public holidays)) with oversight by a consultant respiratory physician (Professor Elizabeth Sapey) will be provided for participants reporting severe symptoms on Atom5™. Prof Sapey will be mentoring and supporting the nurses, providing medical guidance when needed by answering the nurses queries and ensuring signposting to appropriate medical services, as needed. This might include reviewing questionnaire answers from participants, speaking to participants should any severe symptoms be flagged or should the participant wish to speak to a doctor involved in the TLC study, or speaking to the participant's usual medical practitioner to discuss responses to the questionnaires. Prof Sapey can only provide signposting advice (including asking the participant to present to hospital for an emergency review if needed), to ensure appropriate safety netting for the participant, and cannot provide medical treatments or further investigation, as the participant will remain under the care of their usual medical practitioner.

Should participants be concerned about their symptoms or experience a sudden deterioration in their symptoms, they will be encouraged to discuss this with their GP, call 111 or 999, depending on the severity of their symptoms, as per usual standard of care practices and as advised by the medically qualified support team. This will be explained on the PIS. The nurse-led support team can signpost to medical assistance, but will not be able to provide direct medical care or aid.

GPs or the participants' usual care team will not be routinely informed about the results of the research or the data collected. However, Atom5™ will be configured to report the recording of severe symptoms to the research nurse via the Atom5™ clinical portal. The research nurse will then contact the participant to assess the level of support needed and signpost to relevant services and information or escalated to the consultant physician for further advice if needed. The advice given may include a recommendation for the participant to contact their GP or NHS 111 to discuss their symptoms or seek more urgent help by contacting 999. The participant's GP will only be contacted if the participant cannot be reached after three attempts over the

course of 24 hours and the participant has provided their GP details. We will seek specific consent for these actions. Any verbal contact with the participant or GP will be followed by a written letter.

Participants will be advised in the PIS and in the Atom5™ app and website that the data will not routinely be seen by their GP or usual care team. If they are concerned about symptoms or well-being they should dial 111. In an emergency they should dial 999. They will also be provided with the link to the NHS Your COVID Recovery website (https://www.yourcovidrecovery.nhs.uk). Participants reporting symptoms of urgent clinical concern will receive the following notification on Atom5™: "You have reported symptoms that may require urgent follow-up. We recommend you contact 111 to seek further advice. In an emergency dial 999."

8. 3 Distinct Symptom clusters

Participants selected for bio-sampling and who have agreed to be contacted to provide blood and saliva samples and use wearable devices (BioWear Sub-Study) will be contacted by a member of the research team to arrange a date and time for blood collection. Visits for blood and saliva sampling and collection of the wearable devices will take place at the NIHR/Wellcome Trust Clinical Research Facility (CRF) based at the Queen Elizabeth Hospital Birmingham at a time suitable for the participant. If for whatever reason participants are not able or willing to attend the CRF, home visits will be offered by a research nurse. Sample tubes will be pre-labelled with a unique anonymised global ID number, with no patient identifiable information and transported back to the CRF by the research nurse collecting the sample, in appropriate containment. Research staff conducting these visits will be members of the central research team and have consent, Good Clinical Practice (GCP), and phlebotomy training.

Participants selected for wearable monitoring will be provided with wearable devices (Garmin Vivosmart4) at the same visit that the bio-sampling is undertaken. The wearables will measure heart rate, oxygen saturation, step count and sleep quality. Via Atom5™ participants will be asked to undertake a 40-step test (they will be asked to walk 40 steps and measure their oxygen saturation pre and post-test, as well as logging activity on their app) once a week for each week that they use the wearable device. They will be asked to wear the device continuously (except for when the device requires charging) for 28 consecutive days from the time the device is first received, and then again for 28 consecutive days at six and 12 months. Participants will be provided with instructions and both online and telephone support in the use of the wearable devices. They will also be clearly informed that the devices will not be used to monitor their health and that their data will not be used to inform their care.

The data from the biosampling and wearables will be used in combination with symptom data from Atom5™ and clinical data from CPRD Aurum to deep phenotype Long COVID syndromes. The health service use (primary care consultations, prescriptions, A&E attendances and hospital admissions) and mortality of patients from each Long COVID syndrome will also be described using data from CPRD Aurum and linked data from Hospital Episode Statistics and Office for National Statistics mortality statistics (via CPRD).

8. 4 Co-production workshops

Participants will be invited to attend one or more of a series of co-production workshops that will be held either remotely or face-to-face. The aim of the workshops is to reach a consensus on which interventions should be recommended for use to treat Long COVID, which are suitable for evaluation in a clinical trial, and which should not be recommended, based on the available evidence. They will also be used to determine which of the agreed interventions are suitable for or could be adapted for remote delivery. Potential

participants showing an interest will be provided with the relevant PIS, information on how to provide informed consent and event details. Participants will be asked to register prior to the event.

9.0 Data collection

9.1 Questionnaire data collection

Table 1.0-Summary of data collection

	Study time point												
	Baseline (BL)	1 M	2 M	3 M	4 M	5 M	6 M	7 M	8M	9 M	10 M	11 M	12 M
Consent	✓												
Confirmation of COVID-19 infection†	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Confirmation of hospital admission within 28 days of COVID-19‡	✓												
Name	✓												
Date of birth	✓												
Telephone number	✓												
e-mail address	✓												
Postal address	✓												
Sex at birth	✓												
Ethnic group	✓												
GP details	✓												
Height	✓												
Weight	✓												
Smoking status	✓												
Pregnancy status¥	✓												
SARS CoV-2 vaccination status	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Shielding status	✓												
Care home residence	✓												
Occupation	✓												
Occupational status	✓						✓						✓
Change in duration of working/education	✓												
Reason for change in working or education	✓												
Over-the-counter medication use	✓												

	Study time point												
	Baseline (BL)	1 M	2 M	3 M	4 M	5 M	6 M	7 M	8M	9 M	10 M	11 M	12 M
SBQ-LC™	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
FACIT-Fatigue	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Modified MRC Dyspnoea Scale	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
COVID-19 core outcome measure for recovery	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
PHQ-2	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
GAD-2	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
PCL-2	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
EQ-5D	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Bio-Sampling*					✓								
Garmin Wearable Device*					✓						✓		

† Only for controls

‡ Only for cases

¥ Only for females

*****Participants selected for bio-sampling and wearable Garmin devices. Data will be recorded from the Garmin devices at baseline (when the devices are provided) for 4 weeks, then 6 months and 12 months on.

Questionnaires will contain the unique Global ID and participant identifiable data (e.g., name, date of birth, address) will be stored separately from research data within the Atom5™ platform and will be accessible only to designated research staff for the purpose of contacting participants for research visits, registering participants at University Hospitals Birmingham for bio-sampling appointments, and where participants report severe Long COVID symptoms and may need advice or support.

Participants will complete the questionnaires monthly within a timeframe window of -/+ 7 days. Reminders/notifications will be sent by Atom5™ at Baseline/+2 days/+6 days for the baseline questionnaires and at –7 days/0/+2 days for subsequent ones.

At registration, all participants will be asked to provide the onboarding code provided in their invitation letter. Once participants have registered and provided informed consent, they will then be invited to provide baseline demographic data and report the presence of a wide range of symptoms, quality of life and work/education capability using validated tools via the Atom5™ platform.

Patient-reported outcomes will be assessed using symptom specific measures (Table 1): Symptom Burden Questionnaire-Long COVID (SBQ-LC™), FACIT-Fatigue, modified MRC Dyspnoea Scale, COVID-19 core outcome measure for recovery, Patient Health Questionnaire-2 (PHQ-2), Generalized Anxiety Disorder-2 (GAD-2), Abbreviated PTSD Checklist (PCL-2), and quality of life using the EuroQol-5D (EQ-5D).

The SBQ-LC™ will undergo further validation during the study. The development and initial validation of SBQ-LC™ has been applied for through a separate ethics application to the University of Birmingham Research Ethics Committee since participants will not be recruited via the NHS. The SBQ is a disease-specific, multi-domain instrument that will be used to assess symptom burden and functional status (interference). Respondents will be asked to indicate the presence of a symptom over the last 7 days and then provide an indication of symptom burden by providing a response on a 5-point ordinal scale measuring either symptom severity, symptom frequency, or interference. Higher scores will be indicative of greater symptom burden and interference.

The Functional Assessment of Chronic Illness Therapy –Fatigue Scale (FACIT-Fatigue) is a 13-item measure that assesses self-reported fatigue and its impact upon daily activities and function[28]. It has been validated for use across a range of populations. The FACIT-Fatigue has a 7-day recall period and is scored on a 5-point Likert Scale from "0-Not at all"- to "4-Very much". Individual items scores are summed (two items of which are reverse scored), multiplied by 13 and then divided by the number of items answered, with a higher score indicating less fatigue and better quality of life. It takes less than 5 min to complete.

The modified MRC Dyspnoea Scale is a modification to the widely used MRC Dyspnoea scale[29]. The item has a 24-hour recall period and is scored on a 5-point Likert scale from "0 - I only get breathless with strenuous exercise" to "4 - I was breathless when dressing, talking or at rest". It has a high degree of conceptual overlap with its parent clinical measure, the MRC Dyspnoea scale, which is in wide-scale clinical practice.

The COVID-19 core outcome measure for recovery is a single item intending to measure a return to the pre-illness state. The item has a same day recall period and is scored on a 5-point Likert scale from "0 - Completely recovered" to "5 - Not recovered at all"[30].

The Euroqol EQ-5D-5L comprises five items plus one visual analogue scale. It has been widely validated and used to assess health outcomes from a wide variety of interventions on a common scale for purposes of evaluation, allocation and monitoring. It is used by the National Institute for Health and Care Excellence (NICE) in health technology assessment. EQ-5D-5L takes only a few minutes to complete and has a same day recall period.

The GAD-2 is a screening tool for generalised anxiety disorder derived from the GAD-7. It comprises the first 2 items of the GAD-7, which are considered as the core anxiety symptoms ("feeling nervous, anxious or on edge"/Not being able to stop or control worrying")[31]. The GAD-2 performs well as a screening tool for three other common anxiety disorders (panic disorder, social anxiety disorder, and PTSD). It has a recall period of two weeks. The GAD-2 has a global score (0-6, no weighting). A higher score indicates an increased likelihood of underlying anxiety disorder. The recommended cut-off score for further investigation is ≥ 3. The GAD-2 has been validated in many studies and has retained the same psychometrics properties of the GAD-7 (86% sensitivity/83% specificity)[32]

The PHQ-2 is a screening tool for depression derived from the PHQ-9. It comprises the first 2 items of the PHQ-9 (depressed mood and anhedonia). The PHQ-2 has a recall period of 2 weeks. It has a global score (0-6, no weighting). A higher score indicates an increased likelihood of an underlying depressive disorder. The recommended cut-off score for further investigation is ≥ 3. The PHQ-2 has been validated in many studies and has shown a sensitivity of 83% and specificity of 92%[33].

The PCL-2 is an abbreviated version of the PTSD Checklist – Civilian version (PCL-C) and is used to screen for post-traumatic stress disorder (PTSD). It comprises two items (intrusive memories/distress associated with reminders of the traumatic event). It has a recall period of one month. An individual is considered to have

screened positive if the sum of these two items is ≥ 4 (Lang et al, 2005). Previous studies have shown that the PCL-2 has good psychometric properties and have shown sensitivity of .97 and specificity of .58 [34] .

The selection of appropriate patient- reported outcome measures has been informed by: i) outcomes that matter to patients identified through relevant literature, ii) input from the TLC lived experience advisory group, and iii) to align with outcomes being collected in HEAL-COVID. The team has included PRO measures that have demonstrated good psychometric properties in other clinical populations. The SBQ-LC™ will be validated in people with Long COVID in a separate study application.

Questionnaires will take approximately 15 minutes in total to complete. Completion will be permitted in all settings (home, care home, other). The order of administration will be standardised. Participants will be sent up to two reminders to help minimise missing data. For participants who are unable to use the Atom5™ app or website (due to patient preference, language, or other requirements) the research team will offer interview-led completion by telephone or video conference, with translators as required. Alternative language versions of each PRO will be provided when validated translations exist. All PROs will be used in accordance with user manuals when available. Where possible all data should be provided directly by the research participant. However, proxy completion by a carer, family member or friend will be permissible if the individual is too unwell or has technical difficulties. Proxy completion will be recorded on Atom5™.

Questionnaires will contain the unique Global ID and participant identifiable data (e.g., name, date of birth, address) will be stored separately from research data within the Atom5™ platform and will be accessible only to designated research staff for the purpose of contacting participants for research visits, registering participants at University Hospitals Birmingham for bio-sampling appointments, and where participants report severe Long COVID symptoms and may need advice or support.

9.2 Garmin wearable devices

Eligible participants will be provided with wearable devices (Garmin Vivosmart4) to measure heart rate, oxygen saturation, step count and sleep quality. Aparito has a partnership with Garmin which allows the hardware sensor data to be communicated directly to the Atom5™ platform and does not therefore need to interact with the Garmin software (Garmin Connect). Participants will be asked to wear the devices continuously for four consecutive weeks (except when recharging the devices), when they first receive the device, at 6 months and then at 12 months. During each week of use, participants will be prompted to undertake a 40-step test, which will involve participants being asked to walk 40 steps as an exertional test to assess for changes in heart rate and oxygen saturation before and after the test.

9.3 Laboratory Data

Once analysed, blood sample and saliva data will be inputted directly into the Atom5™ platform by researchers within the laboratory. Laboratory staff will have no access to patient identifiable data and will input results using the participant's unique Global ID.

9.4 Co-production

Co-production workshops will be audio recorded using an encrypted digital recorder. Participants will also complete a short demographic questionnaire and field notes will be taken, as necessary.

Recordings will be transcribed verbatim by a professional transcription service who is approved by the University of Birmingham and will sign a confidentiality agreement. The transcription company will delete the recording once the researcher is happy with the accuracy of the transcript. The transcripts will be anonymised and assigned a unique identification number.

9.5 Data storage

9.5.1 Aparito Atom5™

Data collected from questionnaires, wearable devices, saliva and blood sample analysis will be stored on the Atom5™ platform to ensure one single-sign on portal for all study staff and participants. The type and level of data access granted will be in accordance with the staff member's role (view only vs data entry and editing rights). Access to Atom5™ will be via entry of a username (Global ID) and password. Access to the system by a user will be logged with a date and time stamp for audit purposes. Data will be stored on an Azure Microsoft Cloud in the UK (Cardiff). Data export will be conducted in CSV files at several time points during the study to the research team to enable data checking and interim as well as final study analyses. Anonymised data will be stored on a secure server within UoB for a period of 10 years. For those participants who have consented to be contacted for future ethically approved research their data will be stored in a secure database for 10 years after study completion. All Participant Identifiable Data (PID) stored on the Atom5™ platform will be deleted at the time of decommissioning and all data transferred to the study sponsor.

9.5.2 Co-Production

Audio files will be transferred from the audio recorder into an encrypted folder on to the secure university server as soon as possible after the interviews. The recordings on the digital audio recorder will be deleted as soon as successful transfers to the encrypted laptop have been confirmed. The audio files will be securely shared with an approved transcription service authorised by the University of Birmingham. Only anonymised quotes from the transcript will be used in any arising publications or reports.

10.0 Data Analysis Plan

10.1 Long COVID symptom clusters

Unsupervised exploratory clustering techniques will be employed to identify distinct Long COVID symptom clusters (objective 4).

We will first perform dimension reduction using one of: multi-dimensional scaling (MDS), t-stochastic neighbour embedding (t-SNE) and uniform manifold approximation and projection (UMAP). This first step will reduce dimensionality of the global data set while retaining local distances between individuals.

Following the dimensionality reduction step, we will employ clustering algorithms that have been shown to work well in medical settings:

(1) k-means clustering, which locates the stable set of centroids of a predetermined number of k clusters.

(2) Fuzzy c-means cluster analysis, a version of k-means clustering which allows data points to have partial cluster membership.

(3) Hierarchical agglomerative clustering, which joins observations in a recursive fashion satisfying some linkage criterion until k clusters have been generated. We will use the Ward linkage, which joins observations that minimise the variance of merged clusters.

(4) Gaussian finite mixture models, a model-based approach to clustering which associate each data point with a multivariate (ellipsoidal) Gaussian distribution.

(5) DBSCAN/HDBSCAN, algorithms, which identify clusters of dense observations among unassigned lower-density observations.

The optimal number of clusters for all methods (except DBSCAN) will be determined using model selection methods and internal cross validation. Cluster membership goodness of fit will be evaluated by: (1) the Jaccard similarities (intersection size divided by union size) of the original clusters to the most similar clusters in the resampled data, and (2) the average silhouette score, where the silhouette score varies between +1 for an observation aligned with its cluster centre and -1 for an observation aligned with the centre of another cluster. We will determine mean bootstrapped Jaccard and silhouette indices to assess cluster stability. The plot of within-groups sum of squares vs K will be checked graphically for optimal K (elbow method) and a minimum BIC.

10.2 Validation of the SBQ-LC™

Anonymised PRO data collected via the Aparito Atom5™ digital platform during WP1 will be downloaded from the Aparito cloud at multiple time points as a csv file, cleaned and prepared for statistical analyses. A range of statistical analysis software packages will be used to undertake Rasch analysis, report descriptive item/scale characteristics (e.g., response frequency, means, skewness, kurtosis), and to undertake psychometric analyses using traditional methods including evaluation of internal consistency, reliability, test-retest reliability, construct/criterion validity, measurement error, and responsiveness. Analysis will be led by one researcher and discussed and verified by a second researcher with expertise in PRO development and psychometrics. The analysis plan is described below:

<u>Sample and item characteristics</u>

Demographic characteristics of participants will be reported using descriptive statistics (i.e., frequency counts). Per item, distribution of responses across response categories, score means, skewness, kurtosis, and presence of floor/ceiling effects) will be reported.

<u>Rasch analysis</u>

To confirm fit of the SBQ-LC™'s items to the Rasch model for interval-level measurement, a range of statistical and graphical analyses will be performed for each scale/subscale:

Assessment of rating scale functioning - As greater symptom burden should equate to higher scores on the SBQ-LC™, it is expected that the higher the level of severity, frequency or interference, the higher the response category that will be endorsed. Items will be inspected to ensure that all items orient with the latent variable. Response category thresholds will be inspected to confirm that categories advance monotonically across the measurement continuum. The item category probability curves (CPC) will be inspected for threshold ordering. We will check that each response category is endorsed by a minimum of ten persons per response category[35] . Outfit MNSQ values < 2.0 logits will be considered evidence of acceptable category fit[36].

Assessment of item and person fit to the model - Fit statistics show the extent that item performance and person ability differ from the Rasch modelled expectations. Item fit will be confirmed by inspecting the item characteristic curves (ICCs) for each item. The empirical and model ICCs should present a close alignment of the observed and predicted scores for an item. Person fit will be assessed in the same manner as item fit. Infit and outfit statistics with mean square values outside the 0.5 – 1.5 logit range will be considered evidence of misfit. Misfitting persons will be removed from the dataset and the Rasch analysis repeated to ascertain the impact of misfitting persons to the analysis.

Assessment of unidimensionality and local independence - Unidimensionality (i.e., the set of items under assessment represents a single construct) and local independence of items are core requirements of the Rasch model. Unidimensionality will be assessed using a principal component analysis of the residuals (PCAR). The following criteria will be applied as evidence of unidimensionality:

1. an eigenvalue less than 3.0 on the first residual contrast[39,40]
2. The percentage variance explained by the first contrast of 5% or less[41]
3. Disattenuated correlations between the person measures for the item clusters on the first residual contrast greater than 0.70[40]

Local independence will be appraised by examining item fit statistics for overfit and correlations between items of the standardised residuals. Overfit and a substantial correlation of the standardised residuals for two items will be considered evidence of local dependence. A range of critical values for residual correlations have been reported in the literature. For this study, COSMIN-recommended residual correlation values of $r < 0.2$ will be considered evidence of low dependency.

Assessment of differential item functioning (DIF) – Measurement invariance is a criterion of the Rasch model and is established through assessment of DIF. The items of the SBQ-LC™ will be evaluated for DIF for the sample characteristics of gender and age. If possible, and depending on the sample, exploration of cross-cultural DIF and DIF for groups classified according to comorbidity will also be undertaken. DIF contrast values (i.e., effect size) for establishing statistically significant DIF ($p \leq 0.05$) will be based on the size of the reference and focal groups.

Targeting of the scale - Targeting assesses the ability of an instrument's items to measure the full range of persons in a sample. Items should collectively form a hierarchy of symptom burden. The item hierarchy should range from the item representing the least burden to the item measuring maximum burden measured using log-odds units or logits. Targeting of the SBQ-LC™'s scales will be assessed by inspecting: 1) the person-item distribution map 2) the summary statistics of item and person measures; and 3) the location range for items and persons on the Rasch scale.

Assessment of reliability - Reliability will be evaluated in Winsteps using separation indices. Separation indices indicate the number of distinct levels of functioning that can be distinguished for both items and persons. A person separation index of 1.50 may be considered to represent an acceptable level of separation, whereas an index of 2.00 to represent a good level of separation and index of 3.00 represent to an excellent level of separation [42]. A second statistic, the person reliability index, indicates the replicability of person ordering if the same sample were given another set of parallel items measuring the same construct. Higher values are indicative

of better reliability with values exceeding thresholds of 0.7, 0.8, and 0.9 indicating "acceptable", "good", and "excellent" replicability [43]. Item separation will be reported in the same manner as person separation.

Internal Consistency Reliability

Internal consistency reliability, a measure of how closely related a set of items are as a group, will be evaluated by computing Cronbach's alpha (α) for each unidimensional scale of the SBQ-LC™. Alpha values greater than 0.70 in addition to confirmed fit to the Rasch model will be considered evidence of good internal consistency reliability.

Test-retest reliability

Scale-level intra class correlations (ICCs) using a two-way mixed effects model with absolute agreement and its 95% confidence interval will be specified[17]. A mixed effects model was selected over a random effects model as the test-retest time points will be fixed (± 7 days). An ICC coefficient value ≥ 0.7 will be considered evidence of adequate reliability[44]. Respondents whose symptoms have remained stable and report that that they have not started any new treatment for their Long COVID symptoms within the last four weeks will be included in analyses of test-retest reliability.

Measurement error

The standard error of measurement (SEM) is considered to be a parameter for the amount of measurement error present in an instrument. Using a test-retest design, standard error of measurement (SEM) will be calculated. SEM will be calculated by first creating a variable for the different between the score obtained during the first and second administration (test score – retest score = *Difference*). The standard deviation of *Difference* in the sample ($SD_{difference}$) will be calculated. $SD_{difference}$ will be divided by the square root of 2 to calculate the SEM [45]. Smallest detectable change (SDC) provides a value for the minimum change that needs to be observed in order to be confident that the observed change is real and not due to measurement error. SDC will be calculated for individuals (SDC_{ind}) and group-level scores (SDC_{grp}).

Construct validity

A priori hypotheses will be generated to assess the SBQ-LC™'s construct validity. These hypotheses will be refined once the SBQ-LC™ scale structure, as a new PRO, is confirmed. Indicatively:

1. FACIT-Fatigue/– Scores on the FACIT-Fatigue will be positively correlated with SBQ-LC™ scores.
2. Modified MRC Dyspnoea Scale – Scores on the Modified MRC Dyspnoea Scale will be positively correlated with scores on the SBQ-LC™ .
3. Covid-19 core outcome measure for recovery – The Covid-19 core outcome measure for recovery scores will be negatively correlated with SBQ-LC™'s scale scores.
4. PHQ-2 – Scores on the PHQ-2 will be positively correlated with scores on the SBQ-LC™.
5. GAD-2 – Scores on the GAD-2 will be positively correlated with scores on the SBQ-LC™.
6. PCL-2 – Scores on the PLC-2 will be positively correlated with scores on the SBQ-LC™.
7. EQ5D-5L – Scores on the EQ5D VAS will be negatively correlated with scores on the SBQ-LC™.

For continuous scores, Pearson correlations will be used to examine associations. Interpretation of Pearson's r-values will be based on the guidelines by Cohen (1988): small (r = 0.10 to 0.29); medium (r = 0.30 to 0.49); and large (r = 0.50 to 1.0).

Known-groups validity

Known-groups validity will be established by comparing SBQ-LC™ scores from individuals with Long COVID with matched controls. A hypothesis testing approach and parametric tests will be used to compare scores on the SBQ-LC™ between groups. A further sub-group analysis will be undertaken whereby Individuals with Long COVID will be grouped using their score rating on the Covid-19 recovery measure as an anchor question. Mean SBQ-LC™ scores will be compared across groups using parametric statistical tests.

Criterion validity and establishing cut-off values

Using SBQ-LC™ data from matched controls and participants with Long COVID, we will conduct Receiver Operating Characteristic (ROC) curves analyses. Area Under the Curve (AUC) will be calculated to evaluate the SBQ-LC™'s ability to identify individuals with Long COVID from matched controls. COSMIN guidance recommends an AUC ≥ 0.70 as sufficient evidence of criterion validity. The ROC curve will be used to establish cut-off values to support the identification of individuals with Long COVID from the general population.

Responsiveness

ROC analysis and hypotheses testing (construct approach) using PRO change scores (I.e., change in PRO scores between two timepoints) will be used to evaluate responsiveness (I.e., see construct validity above). Responsiveness testing, as with all other aspects of SBQ-LC™ validation, will comply with COSMIN recommendations on methodological quality for studies of PRO validation.

10.3 Co-Production

QSR NVivo12 will be used to manage, sort, code and organise the transcribed data. All transcribed data will be anonymised. Interview field notes, audio files and transcripts will be analysed using a rapid assessment procedure (RAP) sheet[46] . The RAP sheet will be a working document summarising the main findings after each workshop.

11.0 Bio-samples

11.1 Patient selection

A sample of approximately 300 participants with Long COVID from each identified syndrome (symptom cluster) will be invited for bio-sampling (50 individuals per syndrome [total of four syndromes] plus 50 from each of the two control groups) 3 months following study enrolment.

11.2 Sample collection

For venous blood draws, 1 x 10ml serum, 6 x 6ml Lithium Heparin and 8.5ml Paxgene DNA samples will be collected. SncRNAs will be extracted from a small sample of serum collected in a standard red-top tube (2ml aliquot). A maximum of 60 ml blood will be drawn from each participant.

The salivary sample collection will be performed by using a commercial buccal swab kit (DNA Genotek P-157 saliva collection kit). The buccal swab will be applied by the research team or the patient to the inside of the cheek near the floor of the mouth for a few seconds and causes no discomfort to the patient. Buccal swabs will not be collected from patients with bleeding or open sores in the mouth.

Once collected, blood and saliva samples will be labelled with the Global ID before being transported to the University of Birmingham Research Laboratory based at the Queen Elizabeth Hospital Birmingham. Research staff within the laboratory will be unable to identify research participants.

11.3 Sample processing

Processing will follow a standard laboratory research protocol. Processed samples will be stored and batch tested later. Samples may also be stored for repeat or further testing. Most samples will be analysed in Birmingham. However, serum/plasma analysed for proteins will be carried out using the Somalogic platform based in the USA. A material transfer agreement will be in place with other centres where testing will take place outside of UoB. Only pseudo anonymised samples will be transferred between sites and this will be recorded centrally so there is full traceability of samples.

The saliva kit includes a collection tube containing stabilisers that preserve the miRNAs at room temperature for up to eight weeks. RNA species will be extracted from the salivary swabs and serum using standard kits in accordance with the manufacturer's instructions. The sncRNA will then be analysed using Next Generation Sequencing analysis in the first instance. This produces a list of candidate biomarkers that includes all human sncRNAs (<200 basis). The candidate biomarkers that survive a set False Discovery Rate (typically 0.1) are taken forward for the next step, which includes amplification and measurement by standard qPCR. The results will be analysed and interpreted to detect unique biomarker profiles.

At the end of this study samples will be fully anonymised and stored for 5 years for use in other ethically approved research, either in the UK, or overseas or stored in a licensed tissue bank. There will be no mechanism to identify participants once samples have been anonymised.

Participants will be asked if they agree to consent that any leftover samples can be stored within the Clinical Immunology Service Laboratory/UoB Research Laboratories at the University of Birmingham for future ethically approved research. Only anonymised samples will be stored – the laboratory will hold no personal data.

All freezers used in this study will be temperature controlled and monitored 24 hours a day, with a fitted Tutela monitoring alarm system,in accordance with CIS protocols which have been United Kingdom Accreditation Service (UKAS) accredited.

11.4 Laboratory investigations

Participants will be screened with a comprehensive panel of autoantibodies related to systemic autoimmune conditions, a wide range of organ-specific autoantibodies, and a complete set of neuro-immunological markers for central and peripheral nervous system disorders. We will study any frequently recurring self-antigens to provide evidence of molecular mimicry between SARS CoV-2 antigens and the specific self-antigen. We will align these data with that from a separate study of hospitalised COVID-19 patients through our involvement in the UKRI funded Coronavirus Immunology Consortium (UK-CIC).

Participants will also have their serum analysed for proteins using the Somalogic platform. This will determine levels of 7000 proteins associated with pathogenesis of a range of long-term conditions and includes pro- and anti-inflammatory markers (IL1β, IL6, IL8, IL10, IL1ra, TNFα) and markers of tissue damage and breakdown such as 3-methylhistidine for muscle catabolism. This will delineate mechanistic pathways for the different symptom clusters and will allow us to compare data with the PHOSPCOVID platform, which we contribute to in both the immunology and sarcopenia working groups. This work will complement our involvement in the UKRI-funded COVID-CNS consortium assessing mental health and cognitive decline in hospitalised COVID-19 survivors.

With consent, patient DNA/RNA will be studied/stored for genetic analysis.

These analyses will reveal the potential biological processes driving Long-COVID, some of which may be common to one or more syndromes and others that may be specific to one cluster. The identification of immune pathways associated with specific Long COVID syndromes will enable the identification of existing candidate drugs that target those pathways, for example B cell targeting therapies such as Rituximab if autoantibodies are raised in one or more clusters, or therapies such as Fisetin to remove senescent cells. This will ensure efficient targeting of therapies for acceleration of patient benefit.

Details of laboratory tests to be undertaken in this study are shown in Table 2.0.

Table 2

Sample Collected	**Vacutainer Required**	**Total volume collected (mls)**	**Location**	**Downstream Application**
Plasma	Green Lithium Heparin Vacutainer	6 x 6mls	IIA UoB Research Labs/Clinical Immunology Service Laboratory (CIS)	*Analysed for proteins using the Somalogic platform (USA)
Peripheral Blood Mononuclear Cells (PBMCs)	Green Lithium Heparin Vacutainer		IIA UoB Research Labs/Clinical Immunology Service Laboratory (CIS)	Immunophenotyping (flow cytometry), DNA/RNA isolation
Serum	Red Serum Vacutainer	1 x 10mls	IIA UoB Research Labs/Clinical Immunology Service Laboratory (CIS)	Analysed for autoantibodies (CIS)
Saliva	Commercial buccal swab kit (DNA Genotek P-157 Saliva collection kit)	Swab	IIA UoB Research Labs	SncRNA Analysis
RNA/DNA	PaxGene DNA tube	8.5ml	IIA UoB Research Labs/Clinical Immunology	Downstream genetic analysis

			Service Laboratory (CIS)	
Total volume:		Max 60mls whole blood		

Samples to be batch sent to Somalogic for downstream analysis will be stored in temperature monitored -80 freezers prior to shipping.

12.0 Study Withdrawals

12.1 Participant withdrawal

Participants will be made aware when they consent to the study that they will be free to withdraw at any time. Participants will be encouraged to complete the study measures, but they will not need to offer a reason for withdrawal. If participants are lost to follow-up their samples and data will be included. If a participant withdraws, they will be asked whether any of their collected samples to date can still be used. Alternatively, participants may opt to have their samples and data discarded. If data has already been analysed as part of a locked dataset then participants will be advised that it will not be possible to withdraw their data; however, samples will be destroyed.

12.2 Participant withdrawal - Co-production workshops

Due to the nature of data collection at the co-production workshops, it will not be possible to disaggregate individual participants' contributions from the discussions. Therefore, their data will be retained despite withdrawal. All data will be treated as confidential and anonymised both during analysis and in any subsequent reports.

As detailed in the PIS, for participants who later withdraw consent for photography, software will be used to remove the participant's image from the photograph.

13.0 Safety reporting

No risks are expected to arise from taking part in this observational study and no Serious Adverse Events are anticipated as a unique consequence of participation.

13.1 Expected Events

1. COVID-19 related Death
2. Flare-up of existing Long COVID symptoms and related hospitalisation.
3. Adverse Events are commonly encountered in patients having blood tests. However, blood testing is usually well tolerated. The most common adverse event is an individual feeling faint, during phlebotomy. Facilities, including a bed on which to lie, will be available to mitigate this if required. Sometimes there is bruising at the puncture site. This is a common adverse event and so will not be recorded or reported.

14.0 Sampling

14.1 Sample Size

To define Long COVID symptom clusters:

Reliable detection of clustered data via well-chosen combinations of these methods has been shown to require at minimum 20-30 observations per subgroup provided good cluster separation exists (effect size Δ=4 or over). We expect that with 500 patients in the long Covid cohort we will have power above 80% to detect well-separated clusters comprising at least 5% of the cohort for K=4 to K=6 clusters.

14.2 Sampling technique

We will use linked data from the Second-Generation Surveillance PCR-confirmed COVID-19 registry to identify all confirmed COVID-19 patients in CPRD Aurum. Using linked Hospital Episode Statistics (HES) data we will identify and exclude hospitalised patients. Participating general practices will be asked to review cases and exclude those not suitable for study invitation. We aim to derive a cohort of 4000 (minimum 2000) individuals with COVID-19. Individuals from underrepresented sociodemographic groups (e.g., Black and Minority Ethnic (BAME)) will be oversampled.

A sample of participants with Long COVID from each identified syndrome (symptom cluster) will be invited for bio-sampling (50 individuals per syndrome plus a control group). Power achieved for Kruskal-Wallis one-way analysis of variance ability to detect a difference (rejecting at $\alpha < 0.05$) between two samples with n1=50 and n2=50, given some location difference (d), with samples drawn from the normal distribution and the t5 distribution is shown in Table 3.

TABLE 3 Power achieved for Kruskal-Wallis one-way analysis of variance ability to detect difference (rejecting at $\alpha < 0.05$) between two samples with n1=50 and n2=50, given some location difference (d), with samples drawn from the normal distribution and the t5 distribution.

d	power (normal)	power (t5 distribution)
0.1	7.7%	6.8%
0.2	16.3%	13.6%
0.3	30.5%	24.3%
0.4	48.4%	39.1%
0.5	66.3%	56.2%
0.6	80.8%	71.7%
0.7	90.5%	84.5%
0.8	95.8%	92.5%
0.9	98.4%	96.5%
1	99.4%	98.6%

15.0 ETHICAL AND REGULATORY CONSIDERATIONS

15.1 Assessment and management of risk

Risks and Mitigation

There is a risk is of deriving a non-representative population of individuals with Long COVID. We are addressing this through the use of CPRD for recruitment. University of Birmingham (UoB) has a multi-study license to use the CPRD and have already obtained large parts of the necessary data as part of a number of COVID-19 studies. CPRD and UoB are already working on the largest NIHR-funded data-driven clinical trial assessing the effectiveness of novel anticoagulants in preventing stroke among patients with atrial fibrillation. These existing arrangements will be leveraged to ensure secure data transfer, recruitment, e-consenting, and data linkage.

A further risk is patients reporting severe symptoms on the Aparito Atom5™ platform. We are mitigating this by appointing a nurse at University Hospitals Birmingham who will be line-managed by Professor Elizabeth Sapey who can provide consultant support and rapid referral to the Long COVID clinic if required- please refer to section 8.2.

15.2 Risks and mitigation strategies associated with the Aparito Atom5™ platform

Risk	Mitigation	Probability (H/M/L)	Impact (H/M/L)
Time delay live deployment of data capture platform (WP1).	Aparito will leverage their current Covid specific deployment and rapid configuration	L	M
Aparito has no way of knowing in advance if all patients will have access to smartphone device.	Aparito will supply software with compatible operating systems on both iOS and Android. Telephone completion will be offered.	M	M
Low recruitment	If the number of eligible patients for study recruitment with access to smartphone is low Aparito can provide a study handset on loan	M	M
Low retention / High drop out	Aparito have deployed their app and website with patient input, and a focus on good user experience and ease of use. The inclusion of alerts and notifications support adherence as well as a dashboard for the clinical team to follow up with patients that are showing low engagement to check for reasons	L	H

Technical complexity for intervention study	Aparito have deployed their technology to support global, multilingual studies. Their in-house team are well trained and can be supported with additional contractors if needed	M	H

15.3 COVID-19 infection risk

15.3.1 Hospital visits

Staff conducting Research visits within "The NIHR / Wellcome Trust Clinical Research Facility" (CRF) based at the Queen Elizabeth Hospital Birmingham will follow local guidance on the use of PPE. Patients attending for research visits will be sent detailed instructions prior to their appointment.

15.3.2 Domiciliary visits

Staff conducting Research visits in the home will follow Public Health EnglandCOVID-19 guidance on Personal protective equipment (PPE) – for care workers delivering homecare (domiciliary care) during sustained COVID-19 transmission in the UK.

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/931636/How_to_work_safely_in_domiciliary_care_v7_2_11_2020.pdf

15.3.3

Staff conducting home visits will follow local lone worker policy protocols.

16.0. Research Ethics Committee (REC) and other regulatory review and reports

The study will be conducted in accordance with the principles of good clinical practice (GCP) that have their origin in the Declaration of Helsinki and are consistent with the UK Policy Framework for Health and Social Care Research. These principles protect and promote the interests of patients, service users and the public in health and social care research, by describing ethical conduct and proportionate, assurance-based management of health and social care research, to support and facilitate high-quality research in the UK that has the confidence of patients, service users and the public.

It is for organisations and individuals that have responsibilities for health and social care research. This includes funders, sponsors, researchers and their employers, research sites and care providers.

All investigators within the UK and trial site staff will comply with the requirements of the General Data Protection Regulation and Data protection Act 2018 with respect to the collection, storage, processing and disclosure of personal information and will uphold the Act's core principles.

The CI will send Annual Progress Reports to the Main REC and the Sponsor and report any serious breaches of the study protocol/research governance within 7 days. Any substantial amendments to the research will be submitted to the Main REC and the Sponsor before implementation and copied to the R&D, except in the case of urgent safety measures which may be taken without prior approval.

16.1 Regulatory Review & Compliance

The study will be conducted in accordance with the Research Governance Framework for Health and Social Care, the applicable UK Statutory Instruments, (which include the Data Protection Act 2018 and Human Tissue Act 2008) and GCP. The protocol will be submitted to and approved by the main REC prior to circulation.

Before any patients are enrolled into the bio sampling substudy, the Principal Investigator is required to obtain written confirmation of local R&D approval. Sites will not be permitted to enrol patients until confirmation of R&D approval is received by the Study Office and written confirmation from the Study Office confirms that recruitment may commence.

It is the responsibility of the Principal Investigator to ensure that all subsequent amendments gain the necessary local approval. This does not affect the individual clinicians' responsibility to take immediate action if thought necessary to protect the health and interest of individual patients.

17.0 Peer review

This study was peer reviewed internally and externally by the NIHR Peer review panel.

18.0 Patient & Public Involvement and Engagement

18.1 Overall aim

The overall aim of Patient and Public Involvement and Engagement (PPIE) for the TLC study is to obtain patient perspectives to inform and foster coproduction of the research.

18.2 PPIE membership

The PPIE group will comprise of up to 15 individuals with Long COVID. Our goal is to recruit individuals from diverse backgrounds following the NIHR Include guidance.

In order to ensure that we have a gender and ethnically diverse and representative group, we will recruit as required:

- individuals with Long COVID who reach out directly to the research team expressing an interest in joining the PPIE group (We are currently establishing a database of these individuals)
- members of the LongCOVID SOS and Long COVID Wales through their respective group coordinators, as well as other Long COVID support groups.

Recruitment and consent

Potential PPIE members will receive a phone call or email from the PPIE lead who will explain the role of PPIE members for the TLC project and send the PIS document via email to provide more detail. Consent will be obtained electronically using the PPIE consent forms sent once the individual agrees to join the group. If an individual does not have access to the internet or electronic devices, we will send paper copies of the form for completion with a pre-paid return envelope enclosed.

18.3 PPIE Input

Specific contributions:

- For all the TLC work packages, members of the PPIE group will be invited to contribute to study design, review language/content and contribute to the co-development of patient-facing documentation (e.g., participant information leaflets, consent forms) and online information.
- In Work Package 3 of the TLC Study, they will work collaboratively with clinicians and other key stakeholders to prioritise evidence-based interventions and co-produce a virtual intervention and participant resources for evaluation in Work Package 4.
- Members of the group will also assist with the evaluation of patient and public contributions and input in the entire project.
- Members of the PPIE group will participate in the dissemination and engagement of the public through their networks, national media, plain English campaigns, and local events to ensure the study findings reach individuals with lived experience and their families, government and policymakers, and other stakeholders in health and social care and the third sector. They will also facilitate dissemination findings through co-presentation at relevant conferences and patient groups and contributions to the study website and social media platforms.

18.4 Communication with PPIE group

We will consult with the group, through face-to-face meetings, videoconferencing, and online/email or telephone communication. Methods will be flexible and accessible to promote inclusivity and ensure that we obtain and incorporate diverse views while complying with COVID-19 guidance.

The PPIE group and the broader research team will meet at least twice a year, with dates timed to coincide with relevant work/milestones. We will also communicate updates on study progress/findings at meetings, seeking feedback on interpretation and significance of findings and methods for optimal dissemination. Summaries of results from each work package within the TLC project will be sent to PPIE group members and posted on the TLC Twitter account.

18.5 Reimbursement

Group members will be reimbursed for their time and any expenses incurred according to the 2016 INVOLVE guidelines.

18.6 Support of PPI members

Training - We will provide training and support for patient partners as required through our NIHR PPI infrastructure (training courses delivered via the NIHR Applied Research Collaborative and Biomedical Research Centre).

Health and mental wellbeing – The study team cannot inform participant care. Patient partners will be provided links to the NHS Your COVID Recovery website for information on recovering from Long COVID. We will encourage patients who have concerns about their health to discuss with members of their medical team in the first instance. For non-emergency, mental health issues, we will signpost them to the MIND website. We will suggest that patient partners with severe symptoms contact 111 or 999 if considered life threatening.

18.7 Evaluation of PPI

We will evaluate PPI activities to ensure members feel empowered to meaningfully contribute to the study. The Guidance for Reporting Involvement of Patients and the Public (GRIPP) 2 checklist will be used for the reporting and evaluation of PPI for the TLC project[47]. Feedback on PPI activity outcomes will be shared with the PPI group.

19.0 Data protection and patient confidentiality

The University of Birmingham is the sponsor for this study based in the United Kingdom. The University of Birmingham will be using information from patients in order to undertake this study and will act as the data controller for this study.

All investigators and study site staff will comply with the requirements of the General Data Protection Regulation and Data Protection Act 2018 with regards to the collection, storage, processing, and disclosure of personal information and will uphold the Act's core principles. Permission for the transfer, storage and use of personal identifiable data (PID) in the trial will be provided by consenting participants. Patients will be identified by a unique study identification number. Prior to consent, only members of the participant's care-giving team will have access to personal identifiable data (PID).

All electronic transfers of person-identifiable data will meet industry and NHS-mandated standards including encryption to at least Advanced Encryption Standard (AES). Data will be added to the study database in a linked anonymised manner where the participant's identifiable information is replaced by a study ID number (Global ID).

All PID will be securely handled and maintained in controlled access locations and follow local NHS policies and procedures for information security. All samples leaving sites and received by the lead site and third-party laboratories will carry a unique identifier (Global ID) excluding any PID.

Participants will be aware that their PID might be accessed by external regulatory bodies and sponsor representatives for the purposes of assessing legal compliance and meeting relevant regulatory obligations. PID will be stored at local sites for 10 years following completion of the study. All research data will be stored for a period of 10 years after the end of the study.

19.1 Personal information storage

Aparito software platform Atom5™ is deployed on Microsoft Azure servers in Cardiff. The platform is accessed by study staff allocated to specific user access/rights, with password control.

No data will be stored on the patients' devices. Although patients will be asked to complete the study questionnaires on the Atom5™ app, the data will not be retained on their device and participants will be informed of this. In the event of a loss of connectivity, the data already entered into the Atom5™ app will be encrypted and stored securely on the participants' device until connectivity is restored at which time it will be transferred securely to the Aparito cloud. At the end of the study, the data held by Aparito will be downloaded and stored securely on the University of Birmingham server for ten years and for 20 years in the case of the SBQ-LC™ psychometric evaluation. Once transferred, the data will then be deleted from the Aparito cloud.

The Aparito Atom5™ system complies with ISO 13485 and ISO 27001. This means that they have several procedures and policies in place as required to meet UK Data Protection Act legislation from May 2018 that address the following key compliance requirements for healthcare data capture:

1. The ability to maintain comprehensive electronic records when collecting patient data

2. Being able to take explicit consent from patients regarding the capture of their details

3. Being able to take explicit consent from patients for cross-border transfer of their data

4. Being "Disclosure Ready" should a patient submit a Subject Access Request (SAR)

Any data for analysis will be fully anonymised. All anonymised data will be analysed on encrypted computers.

All essential documentation and study records will be stored by the study team in conformance with the applicable regulatory requirements and access to stored information will be restricted to authorised personnel including sponsor representatives and regulatory authorities. The main clinical study portfolio will be kept in the Institute of Applied Health Research, University of Birmingham.

20.0 Indemnity

The study is being run by the University of Birmingham. The University of Birmingham will act as sponsor. As sponsor, the University is responsible for the general conduct of the study and shall indemnify the Centre against any claims in the UK arising from any negligent act or omission by the University in fulfilling the sponsor role in respect of the study. The university is under no obligation to indemnify the Centre against any claims arising from the conduct of the study centre.

21.0 Research Governance

21.1 Sponsor

Dr Birgit Whitman (Head of Research Governance and Ethics Manager) from the University of Birmingham will act as Sponsor Representative on behalf of the University of Birmingham (Sponsor).

21.2 End of Study

This will be defined as when the last participant completes the study , withdraws from the study and/or death. If a participant withdraws from the study, the participant will be asked if any samples collected up until that time may be used for current and/or future research.

The study will be stopped prematurely if:

- Mandated by the Ethics Committee
- Funding for the study ceases

The Ethics Committee will be notified in writing if the study has been concluded or terminated early.

22.0 Access to the final study dataset

Members of the UoB research team, including the chief investigators, co-investigators and researchers in the team involved in the data analysis, will have access to the final dataset. The interim and final datasets, which will be fully anonymised, will also be accessible to PhD, Masters and undergraduate students under the supervision of members of the research team. UoB will have ownership of the final locked dataset for all consented patients.

23.0 Dissemination policy

The results will be published, presented at conferences, and made available for use by other researchers. We will host a study specific website and Twitter account and will share lay summaries coproduced with our patient partners.

24.0 References

1. Daily summary | Coronavirus in the UK. Accessed March 24, 2021. https://coronavirus.data.gov.uk/
2. Coronavirus (COVID-19) latest insights - Office for National Statistics. Accessed April 29, 2021. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19/latestinsights#overview
3. Greenhalgh T, Knight M, A'Court C, Buxton M, Husain L. Management of post-acute covid-19 in primary care. *BMJ*. 2020;370. doi:10.1136/bmj.m3026
4. Galeotti C, Bayry J. Autoimmune and inflammatory diseases following COVID-19. *Nat Rev Rheumatol*. 2020;16(8):413-414. doi:10.1038/s41584-020-0448-7
5. Mahase E. Long covid could be four different syndromes, review suggests. *BMJ*. 2020;371:m3981. doi:10.1136/bmj.m3981
6. Ploumpidis D. Living with covid-19. *Psychiatrike*. 2020;31(3):197-200. doi:10.22365/jpsych.2020.313.197
7. Sudre CH, Murray B, Varsavsky T, et al. Attributes and predictors of long COVID. *Nat Med*. Published online March 10, 2021:1-6. doi:10.1038/s41591-021-01292-y
8. Brodin P. Immune determinants of COVID-19 disease presentation and severity. *Nat Med*. 2021;27(1):28-33. doi:10.1038/s41591-020-01202-8
9. Nehme J, Borghesan M, Mackedenski S, Bird TG, Demaria M. Cellular senescence as a potential mediator of COVID-19 severity in the elderly. *Aging Cell*. 2020;19(10):e13237. doi:10.1111/acel.13237
10. Mondelli V, Pariante CM. What can neuroimmunology teach us about the symptoms of long-COVID? *Oxford Open Immunol*. 2021;2(1). doi:10.1093/oxfimm/iqab004
11. Nuffield Department of Population Health. Welcome — RECOVERY Trial. Nuffield Department of Population Health. Published 2020. Accessed April 30, 2021. https://www.recoverytrial.net/
12. PHOSP-COVID Home Page - PHOSP-COVID. Accessed April 30, 2021. https://www.phosp.org/
13. Maxwell E. *Living with Covid19 A Dynamic Review of the Evidence around Ongoing Covid19 Symptoms (Often Called Long Covid)*.
14. LongCovidSOS. Accessed April 30, 2021. https://www.longcovidsos.org/
15. FDA, HHS. *Guidance for Industry Use in Medical Product Development to Support Labeling Claims Guidance for Industry*.; 2009. Accessed April 30, 2021. https://www.fda.gov/media/77832/download
16. *Principles for Selecting , Developing , Modifying , and Adapting Patient- Reported Outcome Instruments for Use in Medical Device Evaluation Draft Guidance for Industry and Food and Drug Administration Staff , And Other Stakeholders: Draft Guidance for Ind*.; 2020. Accessed April 30, 2021. https://www.fda.gov/media/141565/download
17. Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res*. 2018;27(5):1159-1170. doi:10.1007/s11136-018-1829-0

18. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1147-1157. doi:10.1007/s11136-018-1798-3
19. Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity - Establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: Part 1 - Eliciting concepts for a new PRO instrument. *Value Heal*. 2011;14(8):967-977. doi:10.1016/j.jval.2011.06.014
20. Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity - Establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: Part 2 - Assessing respondent understanding. *Value Heal*. 2011;14(8):978-988. doi:10.1016/j.jval.2011.06.013
21. Williams SA, Kivimaki M, Langenberg C, et al. Plasma protein patterns as comprehensive indicators of health. *Nat Med*. 2019;25(12):1851-1857. doi:10.1038/s41591-019-0665-2
22. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*. 2005;120(1):15-20. doi:10.1016/j.cell.2004.12.035
23. Di Pietro V, O'Halloran P, Watson CN, et al. Unique diagnostic signatures of concussion in the saliva of male athletes: The Study of Concussion in Rugby Union through MicroRNAs (SCRUM). *Br J Sports Med*. Published online 2021. doi:10.1136/bjsports-2020-103274
24. Watson CN, Belli A, Di Pietro V. Small non-coding RNAs: New class of biomarkers and potential therapeutic targets in neurodegenerative disease. *Front Genet*. 2019;10(APR):364. doi:10.3389/fgene.2019.00364
25. Hawkins J, Madden K, Fletcher A, et al. Development of a framework for the co-production and prototyping of public health interventions. *BMC Public Health*. 2017;17(1):689. doi:10.1186/s12889-017-4695-8
26. *Final Priority Setting Method | James Lind Alliance*. Accessed April 30, 2021. https://www.jla.nihr.ac.uk/jla-guidebook/chapter-8/final-priority-setting-method.htm
27. Michie S, Atkins L, West R. The behaviour change wheel. A guide to designing interventions. Published online 2014.
28. Cella D. The functional assessment of cancer therapy-anemia (FACT-An) scale: A new tool for the assessment of outcomes in cancer anemia and fatigue. In: *Seminars in Hematology*. Vol 34. ; 1997:13-19. Accessed May 6, 2021. https://europepmc.org/article/med/9253779
29. Hajiro T, Nishimura K, Tsukino M, Ikeda A, Koyama H, Izumi T. Analysis of clinical methods used to evaluate dyspnea in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 1998;158(4):1185-1189. doi:10.1164/ajrccm.158.4.9802091
30. Tong A, Baumgart A, Evangelidis N, et al. Core Outcome Measures for Trials in People With Coronavirus Disease 2019: Respiratory Failure, Multiorgan Failure, Shortness of Breath, and Recovery. *Crit Care Med*. 2021;49(3):503-516. doi:10.1097/CCM.0000000000004817
31. Sapra A, Bhandari P, Sharma S, Chanpura T, Lopp L. Using Generalized Anxiety Disorder-2 (GAD-2) and GAD-7 in a Primary Care Setting. *Cureus*. 2020;12(5). doi:10.7759/cureus.8224
32. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: A systematic review and diagnostic metaanalysis. *Gen Hosp Psychiatry*. 2016;39:24-31. doi:10.1016/j.genhosppsych.2015.11.005
33. Kroenke K, Spitzer RL, Williams JBW. The patient health questionnaire-2: Validity of a two-item depression screener. *Med Care*. 2003;41(11):1284-1292. doi:10.1097/01.MLR.0000093487.78664.3C
34. Lang AJ, Stein MB. An abbreviated PTSD checklist for use as a screening instrument in primary care. *Behav Res Ther*. 2005;43(5):585-594. doi:10.1016/j.brat.2004.04.005
35. Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas*. 2002;3(1):85-106.
36. Bond T, Yan Z, Heene M. *Applying the Rasch Model: Fundamental Measurement in the Human Science*. 4th ed. Routledge; 2020. Accessed May 1, 2021. https://www.routledge.com/Applying-the-Rasch-Model-Fundamental-Measurement-in-the-Human-Sciences/Bond-Yan-Heene/p/book/9780367141424
37. Wright BD, Linacre JM. Reasonable mean-square fit values. Rasch Measurement Transactions. Published

1994. Accessed April 30, 2021. https://www.rasch.org/rmt/rmt83b.htm
38. Linarce JM. What do Infit and Outfit, Mean-square and Standardized mean? Rasch Measurement Transactions. Published 2002. Accessed April 30, 2021. https://www.rasch.org/rmt/rmt162f.htm
39. Fan J, Bond T. Applying Rasch measurement in language assessment. In: *Quantitative Data Analysis for Language Assessment Volume I*. Routledge; 2019:83-102. doi:10.4324/9781315187815-5
40. Linacre JM. *A User's Guide to W I N S T E P S® M I N I S T E P Rasch-Model Computer Programs Program Manual 4.8.0*.; 2021. Accessed May 1, 2021. https://www.winsteps.com/a/Winsteps-Manual.pdf
41. Smith AB, Wright P, Selby PJ, Velikova G. A Rasch and factor analysis of the Functional Assessment of Cancer Therapy-General (FACT-G). *Health Qual Life Outcomes*. 2007;5(1):1-10. doi:10.1186/1477-7525-5-19
42. Wolins L, Wright BD, Masters GN. *Rating Scale Analysis: Rasch Measurement*. Vol 78.; 1983. doi:10.2307/2288670
43. Bond TG, Fox CM. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences Second Edition*.; 2007. Accessed May 1, 2021. www.erlbaum.coOl
44. Streiner DL, Norman GR, Cairney J. *Health Measurement Scales: A Practical Guide to Their Development and Use - Oxford Medicine*. Oxford University Press; 2015. doi:10.1093/med/9780199685219.001.0001
45. Mokkink LB, Terwee CB, Patrick DL, et al. *COSMIN Checklist Manual*.; 2012. Accessed May 1, 2021. https://fac.ksu.edu.sa/sites/default/files/cosmin_checklist_manual_v9.pdf
46. Vindrola-Padros C, Chisnall G, Cooper S, et al. Carrying Out Rapid Qualitative Research During a Pandemic: Emerging Lessons From COVID-19. *Qual Health Res*. 2020;30(14):2192-2204. doi:10.1177/1049732320951526
47. Staniszewska S, Brett J, Simera I, et al. GRIPP2 reporting checklists: Tools to improve reporting of patient and public involvement in research. *BMJ*. 2017;358:3453. doi:10.1136/bmj.j3453

25.0 APPENDICES

Baseline data collection in Atom5™

Consent forms

- **TLC study**
- **TLC BioWear Sub-Study**
- **Co-production workshops**
- **PPI evaluation**

Participant Information Sheets

- **TLC study**
- **TLC BioWear Sub-Study**
- **Co-production workshops**
- **PPI evaluation**

Invitations to participate

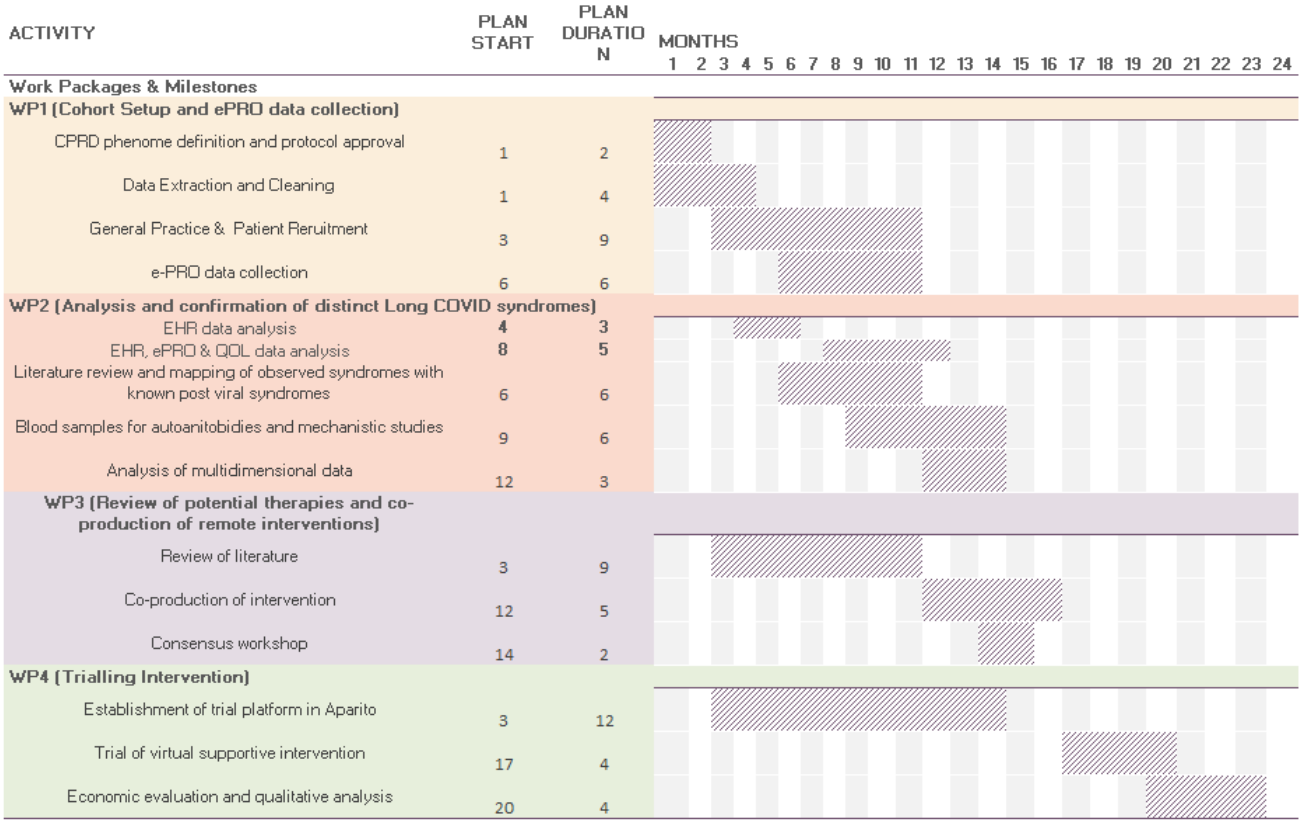
- **TLC study text**
- **TLC study email/ letter**
- **TLC BioWear Sub-Study app notification**
- **TLC BioWear Sub-Study email/ letter**
- **Co-production workshops**

Other

- **Invitation to blood and saliva sample appointment**
- **Co-production workshops advert**
- **Co-production workshops demographic questionnaire**
- **TLC thank you**
- **Voucher text**
- **Atom5™ questionnaire notification and reminders**
- **Withdrawal form**
- **Atom5™ app 'about this study' text**
- **Apartio Atom5™ app screen shots**
- **For all participants text**
- **Baseline data collection in Atom5™**
- **Atom5™ Questionnaires**

Timeframes

TLC Study


ACTIVITY
PLAN START
PLAN DURATION
MONTHS
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Work Packages & Milestones
WP1 (Cohort Setup and ePRO data collection)
CPRD phenome definition and protocol approval 1 2
Data Extraction and Cleaning 1 4
General Practice & Patient Reruitment 3 9
e-PRO data collection 6 6
WP2 (Analysis and confirmation of distinct Long COVID syndromes)
EHR data analysis 4 3
EHR, ePRO & QOL data analysis 8 5
Literature review and mapping of observed syndromes with known post viral syndromes 6 6
Blood samples for autoanitobidies and mechanistic studies 9 6
Analysis of multidimensional data 12 3
WP3 (Review of potential therapies and co-production of remote interventions)
Review of literature 3 9
Co-production of intervention 12 5
Consensus workshop 14 2
WP4 (Trialling Intervention)
Establishment of trial platform in Aparito 3 12
Trial of virtual supportive intervention 17 4
Economic evaluation and qualitative analysis 20 4