ASTERIX

Adaptive **s**tratification of COVID19 **t**o facilitate

Endotype-di**r**ected **I**ntervention Studies

Running title:	ASTERIX
Protocol Version:	1.0
Date:	11th May 2020
REC Reference Number:	20/WS/0077
ISRCTN Number:	To be confirmed
Sponsor:	NHS Greater Glasgow & Clyde
Sponsor Reference Number	GN20ID158
Protocol Number:	ASTERIX-2020

CONTACTS

Co-Chief Investigator: **Prof Kevin Blyth** Consultant Respiratory Physician & Honorary Professor NRS Senior Research Fellow Department of Respiratory Medicine Queen Elizabeth University Hospital 1345 Govan Road, Glasgow G51 4TF Tel: 0141 451 6099 Email**:** kevin.blyth@ggc.scot.nhs.uk	**Co-Chief Investigator:** **Prof Iain McInnes** Director of Institute of Infection, Immunity and Inflammation University of Glasgow University Place, Glasgow G12 8TA Tel: 0141 330 8412 Email: Iain.McInnes@glasgow.ac.uk
Co-Investigator: **Prof Andrew Biankin** Director of Translational Research Wolfson Wohl Cancer Research Centre Garscube Estate, Switchback Road, Glasgow, G61 1QH Tel: 0141 330 5670 Email: Andrew.biankin@glasgow.ac.uk	**Co-Investigator:** **Prof Rob Jones** Professor of Clinical Cancer Research Director of CRUK Clinical Trials Unit Beatson West of Scotland Cancer Centre 1053 Great Western Road Glasgow, G12 0YN Tel: 0141 301 7062 Email: r.jones@beatson.ac.uk
Co-Investigator: **Dr David Lowe** Consultant Emergency Medicine Director EmQuire Queen Elizabeth University Hospital 1345 Govan Rd, Glasgow G51 4TF Tel: 07817446334 Email: david.lowe@glasgow.ac.uk	**Co-Investigator:** **Dr Michael Murphy** Consultant Microbiologist Glasgow Royal Infirmary 84 Castle Street, Glasgow G4 0SF Tel: 0141 4515668 Email: michaelmurphy1@nhs.net
Co-Investigator: **Prof Emma Thomson** MRC Centre for Virus Research University of Glasgow 464 Bearsden Road, Glasgow G61 1QH Tel: 0141 330 2928 Email: emma.thomson@glasgow.ac.uk	**Co-Investigator:** **Dr Janet Scott** MRC Centre for Virus Research University of Glasgow 464 Bearsden Road, Glasgow G61 1QH Tel: 0141 339 8251 Email: janet.scott@glasgow.ac.uk

Co-Investigator: **Prof Duncan Porter** Honorary Professor and Consultant Rheumatologist Queen Elizabeth University Hospital 1345 Govan Road Glasgow G51 4TF Tel: 0141 211 3429 Email: Duncan.Porter@glasgow.ac.uk	**Co-Investigator:** **Prof Carl Goodyear** Professor of Translational Immunology Institute of Infection, Immunity and Inflammation University of Glasgow Glasgow G12 8TA Tel: 0141 330 3865 Email: carl.goodyear@glasgow.ac.uk
Co-Investigator: **Dr Donogh Maguire** Emergency Medicine Consultant Glasgow Royal Infirmary 84 Castle Street, Glasgow G4 0SF Tel: 0787 6624711 Email: donogh.maguire@glasgow.ac.uk	**Co-Investigator:** **Professor Patrick Mark** Professor of Nephrology Institute of Cardiovascular and Medical Sciences, University of Glasgow Glasgow G12 8TA Tel: 0141 330 8218 Email: Patrick.mark@glasgow.ac.uk
Co-Investigator and Lead Study Statistician: **Samantha Hinsley** Cancer Research UK Glasgow CTU Beatson West of Scotland Cancer Centre 1053 Great Western Road Glasgow G12 0YN, UK Tel: 0141 211 3740 Email: samantha.hinsley@glasgow.ac.uk	**Project Manager:** **Carol Evans** Cancer Research UK Glasgow CTU Beatson West of Scotland Cancer Centre 1053 Great Western Road Glasgow G12 0YN, UK Tel: 0141 301 7189 Email: carol.evans@glasgow.ac.uk
Clinical Trial Co-Ordinator: **Laura Divers** Cancer Research UK Glasgow CTU Beatson West of Scotland Cancer Centre 1053 Great Western Road GLASGOW G12 0YN Tel: 0141 301 7246 Email: Laura.Divers@glasgow.ac.uk	**Bio-Repository Contact:** **Clare Orange** NRS GG&C Bio-Repository L3/B/012 Laboratory Medicine Building Queen Elizabeth University Hospital 1345 Govan Road Glasgow G51 4TF Tel: 0141 354 9495 Email: clare.orange@ggc.scot.nhs.uk
Sponsor Contact: **Joanne McGarry** NHS Greater Glasgow and Clyde Clinical Research and Development Dykebar Hospital, Ward 11 Grahamston Road, Paisley, PA2 7DE Tel: 01471 314 4001 Email: joanne.mcgarry@ggc.scot.nhs.uk	

Funders:
ASTERIX is funded by the Chief Scientist Office, with support from the CRUK Glasgow CTU (core funded by Cancer Research UK) and underpinning infrastructure provided by NHS GG&C, NRS Glasgow Bio-Repository and NHS Safe Haven data archive

This study will be performed according to the Research Governance Framework for Health and Community Care (Second edition, 2006) and WORLD MEDICAL ASSOCIATION DECLARATION OF HELSINKI Ethical Principles for Medical Research Involving Human Subjects 1964 (as amended).

PROTOCOL APPROVAL

ASTERIX

Adaptive **s**tratification of COVID19 **t**o facilitate

Endotype-di**r**ected **I**ntervention Studies

Co-Chief Investigator

Prof Kevin Blyth
Department of Respiratory Medicine
Queen Elizabeth University Hospital
1345 Govan Road
Glasgow G51 4TF

Signature: _____

Date: ____ /____ /_____

Co-Chief Investigator

Prof Iain McInnes
Director of Institute of Infection, Immunity and Inflammation
University of Glasgow
University Place,
Glasgow
G12 8TA

Signature: _____

Date: ____ /____ /_____

Lead Statistician

Samantha Hinsley
Lead Biostatistician
Cancer Research UK Glasgow Clinical Trials Unit
Beatson West of Scotland Cancer Centre
1053 Great Western Road
Glasgow
G12 0YN

Signature: _____

Date: ____ /____ /_____

Sponsor's Representative

Joanne McGarry
Research Co-ordinator
NHS Greater Glasgow and Clyde
Clinical Research & Development
Dykebar Hospital, Ward 11
Grahamston Road, Paisley
PA2 7DE

Signature: _____

Date: ____ /____ /_____

TABLE OF CONTENTS

ABBREVIATIONS

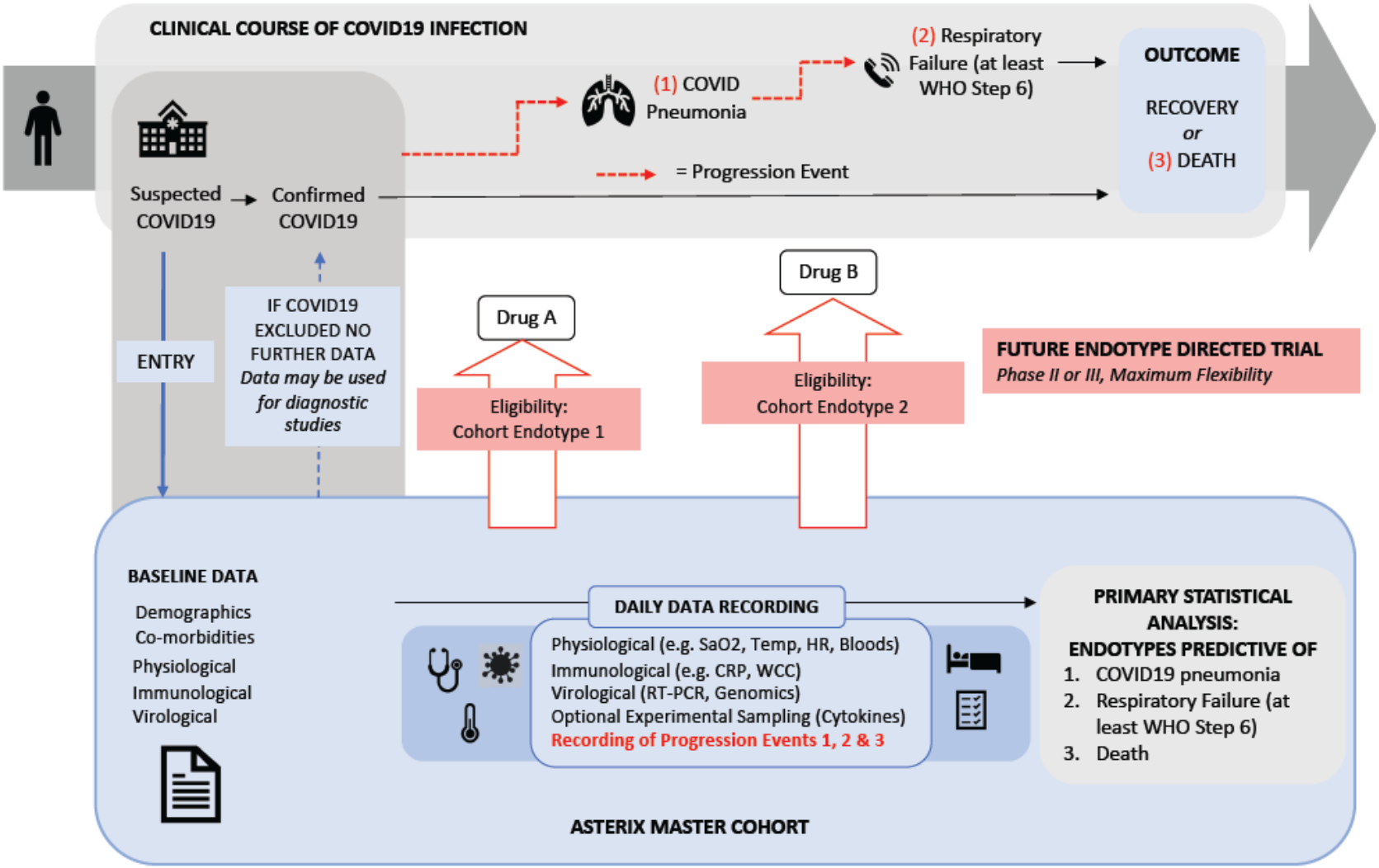
AI	Artificial Intelligence
CHI	Community Health Index
CI	Chief Investigator
COVID19	Coronavirus 19 Disease
CT	Computed Tomography
CTU	Clinical Trials Unit
CVR	Centre for Virus Research
EMR	Electronic Medical Record
GRI	Glasgow Royal Infirmary
iCAIRD	Industrial Centre for Artificial Intelligence Research in Digital Diagnostics
ICU	Intensive Care Unit
MRC	Medical Research Council
NHSGGC	NHS Greater Glasgow and Clyde
NIV	Non-invasive Ventilation
NRS	NHS Research Scotland
PBMC	Peripheral Blood Mononuclear Cells
PI	Principal Investigator
QUEH	Queen Elizabeth University Hospital
REC	Research Ethics Committee
SH	Safe Haven
SNOMED	Systematized Nomenclature of Medicine
TMF	Trial Master File
WHO	World Health Organisation

Study Synopsis

Title	ASTERIX: **A**daptive **s**tratification of COVID19 **t**o facilitate **E**ndotype-di**r**ected **I**ntervention Studies
Population	Between 600 and 1200 patients hospitalised with suspected COVID19
Rationale	The novel betacoronavirus, SARS-CoV-2, is a rapidly emerging global health threat. Health care systems are on the brink of collapse across the world, reflecting the absence of effective therapy for severe coronavirus 19 disease (COVID19). Early data infers significant disease heterogeneity, including minimal or no symptoms in 80%, but a case fatality rate of 1-2%, usually as the result of progressive COVID pneumonia[1,2]. This necessitates subtype targeting for maximum impact and minimum risk, ideally using biological signatures (endotypes, e.g. viral replication v. immune activated) to determine eligibility. COVID19 trials will also be uniquely challenged by time- and resource-dependant variation in key outcome measures (Intensive Care Unit (ICU) admission, mortality) that have the potential to confound results[3,4], making an adaptable platform, regularly calibrating to the clinical environment an essential tool.
Design	ASTERIX is prospective observational cohort study
Coordination	CRUK Glasgow Clinical Trials Unit
Study Hypothesis	That it is possible to identify discrete sub-groups of patients (endotypes), based on a range of demographic, clinical and laboratory biomarkers, that predict: • progression though different stages of disease severity • the likelihood of response to therapies
Study Objectives	• To recruit between 600 and 1200 participants with appropriate approvals • To acquire and record the necessary biological and outcome data using direct electronic medical record (EMR) capture and a bespoke data collection tool embedded in COVID19 clinical care • To provide a structured biobanking infrastructure, with surplus biological samples linked to clinical data, and a resource for observational, translational and interventional studies. • To develop different statistical models based on the data collected for progression to three adverse outcomes: 1) COVID19 Pneumonia 2) Development of respiratory failure, sufficient to prompt escalation to at least Step 6 of the World Health Organisation (WHO) 10-point ordinal scale 3) Death from any cause These models will generate critical information needed to design high impact clinical trials, including the most suitable mechanism(s) of action at different phases of the illness and the event rate at each timepoint based on current standard of care.
Eligibility	Inclusion: Admitted to hospital with suspected COVID19 Exclusion: Age < 8 years

Identification Consent & Approvals	Participants will be identified by clinical staff in emergency departments, acute receiving units and hospital wards. Clinical data and surplus biological samples will be collected with documented verbal consent under study specific REC approval (submitted). This will supplement and extend current NHS GG&C Bio-Repository COVID19 REC approvals (Ref 16/WS/0207). Approval for data linkage and storage will be covered by NHS GG&C Safe Haven (SH) project approval (GSH20RM003). No trial specific activities will occur.
Data Collection	Data will be extracted from the NHS GG&C SH. Automated processes will capture most data fields. However, daily clinical data (e.g. oxygen requirement, clinical observations, vital status) will be collected by ward staff, using a bespoke bedside data collection tool embedded in clinical care (see Appendix 1). Linked anonymized data will subsequently be transferred to the CRUK Glasgow CTU for analysis.
Biological Sample Processing and Banking	Surplus blood +/- any surplus respiratory secretions and urine will be collected at Baseline (Day 0) and where possible in a series of follow-up windows that span the course of admission (up to 14 days). If blood samples are sent from any hospital follow-up visit within 1-year these will also be banked. Immediate blood sample processing will allow storage of plasma, serum and Peripheral Blood Mononuclear Cells (PBMCs).
Utilisation and Sharing	• Tier 0 data will comprise the minimum data set for the primary statistical analyses. This will be made up of clinical data and the results of laboratory tests requested by the clinical team (see Appendix 2). Tier 0 data will be supplemented by viral genomic data generated by separate studies at the MRC University of Glasgow CVR • Tier 1 data will include Tier 0 data plus the results of additional assays available in NHS GG&CC labs and implicated in COVID19 series, but not routinely measured by clinical teams (e.g. Troponin, D-Dimer, see Appendix 3). • Tier 2 data will include Tier 0 (+/- 1) data plus the results of exploratory biomarker analyses planned by University of Glasgow co-investigators (e.g. cytokine/chemokines, metabolomics). • Tier 3 data will only be generated by approved translational studies, which may include host genomic profiling. Approval will be via established governance infrastructure including the NHS GG&C Biorepository Management Group.
Statistical Analyses	The target sample size of 600-1200 will yield 400-800 confirmed cases. These numbers will be sufficient for the primary statistical analyses (generation of reliable models for each progression event). Candidate predictors for each event will be identified *a priori* and by univariate regression. Multivariable predictive models (endotypes) will be generated for each progression event using logistic regression, including a purposeful selection approach. The performance of any developed models will be assessed using a Receiver Operator Characteristic curve, validated in an independent data set.

Figure 1. Overview of the ASTERIX study (blue) in relation to the clinical course of COVID19 infection (grey) and future endotype-directed clinical trials (red).



1. INTRODUCTION

1.1 Background

The novel betacoronavirus (SARS-CoV-2) outbreak that emerged in Wuhan, Hubei province, China in December 2019 is a rapidly emerging threat to global health. Early data regarding the human illness caused by SARS-CoV-2 infection (Coronavirus 19 disease (COVID19)) infers significant disease heterogeneity, including minimal or no symptoms in 80%, but a case fatality rate of 1-2%. Death appears to result primarily from progressive COVID pneumonia[1,2] and respiratory failure, but with varying additional factors including a hyper-inflammatory cytokine storm, fulminant myocarditis, pulmonary thrombo-embolism and acute kidney injury. Poor prognosis is associated with being older, male sex, and co-morbidities, including diabetes, hypertension and obesity[1,2].

1.2 Scientific Rationale

COVID19 trials are currently confounded by a combination of the disease heterogeneity described above, and time/resource-dependent variation in key outcome measures (e.g. ICU admission criteria, mortality)[3,4]. However, it is highly likely that phenotypic clusters ('endotypes') can be defined that explain the likelihood of COVID19 progression, and that once identified, these endotypes could be efficiently targeted in clinical trials deploying drugs with appropriate mechanisms of action, at the right timepoint in the illness. The ASTERIX study will establish the basic regulatory and logistical requirements needed for characterization, phenotyping and data linkage in COVID19 hospital inpatients, and will use the data collected to immediately define endotype signatures (e.g. viral replication, immune activated, microangiopathic) than can be used to design and deploy endotype targeted COVID19 trials.

1.3 Study Hypothesis

That it is possible to identify discrete sub-groups of patients (endotypes), based on a range of demographic, clinical and laboratory biomarkers, that predict progression though different stages of disease severity and the likelihood of response to therapies according to their mode of action.

2. STUDY OBJECTIVES

1) To recruit between 600 and 1200 participants with appropriate ethical and governance approvals.
2) To acquire and record the necessary biological and outcome data using direct EMR capture and a bespoke data collection tool embedded in clinical care.
3) To provide a structured biobanking infrastructure, which utilises surplus biological materials efficiently and can be deployed in observational, translational and interventional studies.
4) To develop different statistical models based on the data collected for progression to three clinically relevant adverse outcomes:
    - COVID19 Pneumonia
    - Development of respiratory failure, sufficient to prompt an escalation in care to at least Step 6 of the WHO 10-point ordinal scale
    - Death from any cause

These models will generate critical information needed to design high impact clinical trials, including the most suitable mechanism(s) of action to deploy at different timepoints and event rates based on current standard of care.

3. STUDY DESIGN

ASTERIX is a prospective single-centre observational cohort study, which effectively functions as a 'Master Protocol' (see Figure 1 for overview). Modelled on other Scottish-led initiatives, such as Precision-Panc[5], ASTERIX will leverage Scotland's strengths in clinical research and data science to define the natural history of COVID19 and clinico-molecular subtypes (endotypes) associated with adverse outcomes. ASTERIX thereby provides a framework to direct patients to future high impact endotype-directed COVID19 clinical trials via separate protocols.

3.1 Study Population

ASTERIX will recruit up to 1200 patients with suspected COVID19 from acute hospitals in NHS GG&C, yielding up to 800 confirmed cases. The lead centre will be the Queen Elizabeth University Hospital (QUEH). At least 10,000 COVID19 admissions

are expected in NHS GG&C (which serves the West of Scotland and a population of 2.7 million people) during the COVID19 pandemic, with Imperial modelling suggesting recurring waves of admissions for at least a further year[6].

3.2 Eligibility

3.2.1 Inclusion Criteria

Admitted to hospital with suspected COVID19.

3.2.2 Exclusion Criteria

Age < 8 years

This aligns with current NHS GG&C Bio-Repository REC approval for storage and use of surplus tissue.

3.3 Identification and Cohort Entry and Exit

Participants that are admitted to hospital with 'suspected COVID19' will be identified by clinical staff in emergency departments, acute receiving units and hospital wards. Entry to the cohort will require confirmation by those staff that the clinical presentation that has resulted in hospitalization is consistent with COVID19, as defined by any of the following:

- Identification on Trakcare that the patient has suspected COVID19, by activation of the COVID19 pathway icon (in green, second from the right in panel below)

- Sending respiratory samples for SARS-CoV-2 testing
- Identification of suspected COVID19 on the NHS GGC COVID19 assessment tool (see daysix.covid.co). This requires the input by the user of 'Yes' or 'Unsure' to the Question 'Is this a COVID19 presentation?'.

Following activation of any of these access points, clinical data (see Section 4.1) and surplus biological samples will be collected (see Section 4.3).

Patients will exit the cohort if subsequent clinical investigations fail to confirm a diagnosis of COVID19. This will require explicit confirmation by the clinical team, defined as input by the NHS GGC COVID19 assessment tool user that the answer to the above question has changed to 'No'. Receipt of a negative SARS-CoV-2 laboratory test will not be sufficient to remove participants from the cohort given the documented false negative rate of up to 25%. No further data will be collected in patients who exit the cohort, but data and samples already banked will be stored and may be used for subsequent diagnostic studies, in which they can serve as control subjects.

3.4 Consent and Approvals

Verbal consent for use of surplus data and samples will be sought and documented using a specific field within the NHS GGC COVID19 assessment tool. This process will be documented in study specific REC approval, which supplements and extend a current COVID19-specific amendment to NHS GG&C Bio-Repository REC approvals (16/WS/0207) that allow for limited COVID sample banking over a short period of time. Consenting participants will be provided with a brief information leaflet with guidance on how to withdraw from the cohort at a later stage, using an established NHS GG&C process. In the case of withdrawal, no further data will be collected, and no further surplus biological samples will be banked. However, data already on file and samples already banked will be retained unless the participant specifically withdraw consent for this. Approval for data linkage and storage, and NHS GG&C R&D approval will be granted via NHS GG&C Safe Haven Project Approval (GSH20RM003). This linkage will include data and some surplus biological samples from historical patients involved in the first wave of COVID19 admissions in NHSGGC (from 1st Jan 2020). These patients will not have given verbal consent in the manner described above for new participants. However, we feel their inclusion is justified as it critically important that we learn quickly and utilise as much data and information as possible to do this. Importantly, any participants who enter the cohort via this

retrospective route will have the same protections in terms of linked anonymised data collection only, and any surplus blood samples that are linked will have been collected under the temporary COVID19 amendment made to the existing Glasgow Bio-repository REC approval (Ref 16/WS/0207, March 2020).

3.5 Registration

Participants will be registered to the study database using a unique study identifier provided by the Bio-Repository and transmitted via the NHS GG&C Safe Haven. Registration will occur on secure transmission of electronic data from the Safe Haven to the CRUK Glasgow Clinical Trials Unit (CTU), which will occur at least every 7 days. This secure transmission will include all available datapoints for every patient meeting the study eligibility criteria.

3.6 Co-enrolment

There will be no co-enrolment restrictions, ASTERIX cohort participants will be permitted to participate in other research studies or concurrent trials and this information will be collected as part of the dataset.

4. STUDY PROCEDURES

4.1 Data Collection

Baseline data (day of admission) and serial data (every subsequent day until discharge or death) will be collected and stored in the NHS GG&C Safe Haven, with regular downloads to the CRUK Glasgow CTU for statistical analysis. Most data fields will be populated by automated extraction from the EMR using established Safe Haven architecture. This will be supplemented by the recording of daily clinical data (e.g. oxygen requirement, clinical observations, vital status) by clinical staff where clinical pressures allow, using a bespoke, bedside data collection tool embedded in clinical care (the NHS GGC COVID19 assessment tool (see daysix.covid.co)). This clinician facing data collection tool will have a data dictionary aligned with the WHO ISARIC study[16], with each term bound to the systematized Nomenclature of medicine (SNOMED). This will ensure interoperability and facilitate data sharing.

Baseline data will include demographics, co-morbidities and drug history. Serial data will include all routine laboratory test results, supplemented by viral genomic data via separate studies being conducted at the MRC-University of Glasgow CVR. The ASTERIX clinical dataset (see Appendix 1 for details) has been aligned with the WHO ISARIC platform[16], but has been restricted to a core data set critical to COVID19.

4.2 Definition of Clinical Progression Events

The ASTERIX dataset includes recording of 3 clinical progression events that will be used to build endotype models (definitions for each are shown in Figure 2, overleaf):

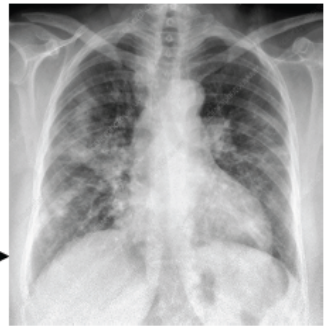
1) Development of COVID pneumonia
2) Development of respiratory failure sufficient to prompt an escalation in care to *at least* Step 6 of the WHO 10-point ordinal scale
3) Death from any cause (step 10 of the WHO 10-point ordinal scale)

Progression (1) will be defined using radiology reports currently being formatted as COVID pneumonia present or absent for an NHS GG&C service improvement project (see below). For progression (2) a clinical deterioration to at least step 6 of the WHO ordinal scale will be used, since this will be immune to resource variation that could bias a definition based on ICU admission *per se*. Daily data, including oxygen saturation, oxygen dose and use of non-invasive ventilation (NIV) or high flow oxygen will be recorded by clinical staff using a bespoke COVID19 ward round tool. This facilitates electronic recording of clinical activity and upload to the EHR (offering clinical efficiency to the user) but also requires the user to input key ASTERIX data points (e.g. oxygen saturation, oxygen dose, temperature, use of any form of respiratory support). This tool has been developed in conjunction with NHS GG&C eHealth & Clinical Informatics (led by Lowe), and will require governance approval at NHS GG&C Board level. Clinical progression (3) will be defined as death from any cause and is equivalent to Step 10 on the WHO ordinal scale. It is obviously expected that some patients will meet more than one of the progression endpoints defined, and some will achieve all three. It is therefore expected that patients will have membership of multiple progression groups. It is also possible that progression will

Figure 2. ASTERIX Clinical Progression Events

Patient State	Descriptor	Score
Uninfected	Uninfected; no viral RNA detected	0
Ambulatory	Asymptomatic; viral RNA detected	1
	Symptomatic; Independent	2
	Symptomatic; Assistance needed	3
Hospitalized: Mild disease	Hospitalized; no oxygen therapy	4
	Hospitalized; oxygen by mask or nasal prongs	5
Hospitalized: Severe disease	Hospitalized; Oxygen by NIV or High flow	6
	Intubation & Mechanical ventilation, $pO_2/FIO_2 \geq 150$ or $SpO_2/FIO_2 \geq 200$	7
	Mechanical ventilation $pO_2/FIO_2 < 150$ ($SpO_2/FIO_2 < 200$) **or** vasopressors	8
	Mechanical ventilation $pO_2/FIO_2 < 150$ **and** vasopressors, dialysis, or ECMO	9
Death	Dead	10

The WHO 10-point ordinal scale



Clinical Progression (1): COVID Pneumonia
- Defined by existing NHSGGC radiology reporting for iCAIRD COVID study

Clinical Progression (2): Severe Hypoxaemic Respiratory Failure sufficient to prompt an escalation in care to *at least* Step 6
- 'NIV' will include CPAP (continuous positive airway pressure) or Bi-PAP (bi-level positive airway pressure)
- 'High Flow' oxygen will include an $FiO_2 \geq 0.6$ or use of high flow nasal oxygen therapy ($HFNO_2$)

Clinical Progression (3): Death

= Cohort Entry

be non-linear, e.g. some patients may progress from cohort entry (hospitalization with suspected COVID19) to Progression 3 (Death). Patients who achieve Progression 2 (clinical deterioration to at least step 6 of the WHO ordinal scale) and who subsequently die will be allocated membership of both Progression 2 and 3. Those who clinically deteriorate to at least step 6 but who do not die (i.e. steps 6-9) will only be allocated membership of Progression Group 2.

4.3 Biological Sample Processing and Storage

No study specific (i.e. no additional) samples will be taken. Surplus material from laboratory tests (blood +/- respiratory and urine samples where available) will be screened, processed and stored for immediate or later translational studies (see Figure 3, overleaf for an overview). Initial screening of samples for membership of the cohort (as defined in Section 3.3) will involve all samples that have completed routinely requested assays. Samples within the cohort will then be aliquoted and immediately stored under the oversight of NHSGGC Biorepository, including appropriate governance and data security arrangements, or transferred for aliquoting to a registered satellite biorepository laboratory (e.g. Institute of Infection, Immunity & Inflammation, University of Glasgow). Any samples transferred to satellite laboratories will be returned to the NHSGGC biorepository for storage. Patients who exit the cohort will have no further surplus samples collected, but those already banked will be stored for future research.

4.3.1 Blood Samples

Surplus blood will be collected at Baseline (Day 0) and where possible in a series of follow-up windows that span the course of admission (up to 14 days). If blood samples are sent from any hospital follow-up visit within 1-year these will also be banked. Blood samples will be centrifuged, and plasma, serum and PBMCs aliquoted from all timepoints, where resources and storage allow (see Figure 3 overleaf). For serum and plasma, up to 5 x 500 microlitre aliquots will be stored per patient per timepoint. For PBMCs, the buffy coat layer will be frozen as a cell pellet, using an established protocol from the Goodyear lab.

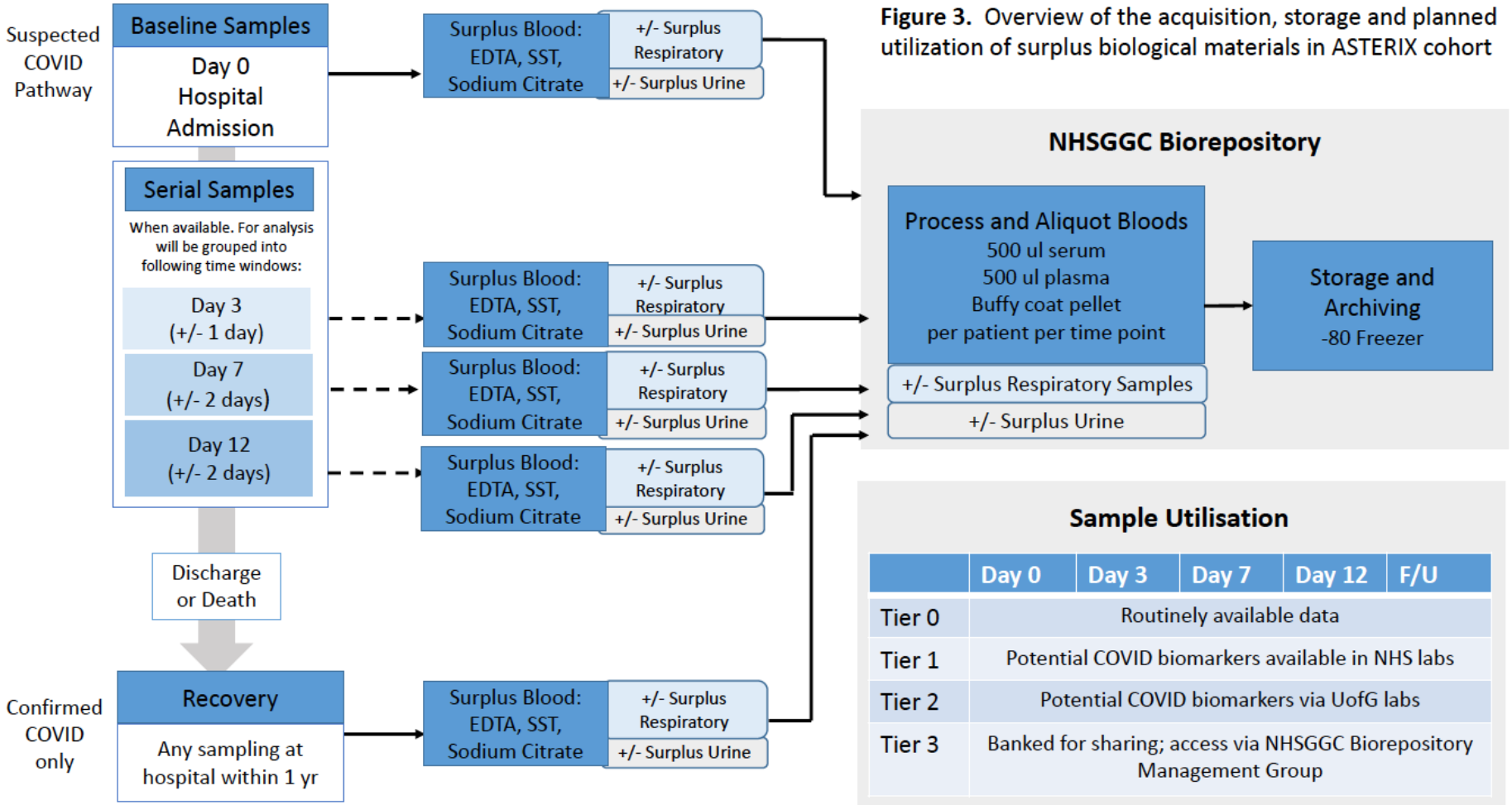


Figure 3. Overview of the acquisition, storage and planned utilization of surplus biological materials in ASTERIX cohort

All blood samples will be stored at -80C under the oversight of NHSGGC Bio-Repository, an NHS Research Scotland (NRS) accredited tissue bank.

4.3.2 Respiratory Secretions

Surplus respiratory samples will be scarce in patients not intubated in ICU, since COVID19 pneumonia is typified by a dry, non-productive cough. Respiratory samples will be more abundant in ventilated patients. Any surplus respiratory samples that are available will be stored appropriately in freezers in the containment level 3 laboratories at Glasgow Royal Infirmary and QEUH under the governance of the NHS GGC Bio-Repository

4.3.3 Urine Samples

Surplus urine samples (at least 500 microlitre aliquots) will also be banked, where these samples are available, and resources and storage allow.

4.4 Sample Utilisation and Sharing

Sample utilisation and prioritisation is summarised in Figure 3 (page 19).

- Tier 0 data will comprise the minimum data set for the primary statistical analyses. This will include clinical data and the results of laboratory tests requested by the clinical team. This is captured within the Trakcare COVID19 order set deployed across NHS GG&C (see Appendix 2). Tier 0 data will be supplemented by viral genomic data generated by separate ongoing studies at the MRC-University of Glasgow CVR
- Tier 1 data will include Tier 0 data plus the results of additional assays available in NHS GG&C labs and implicated in COVID19 series, but not routinely measured by clinical teams (see Appendix 3). These will include N-terminal pro-Brain Natriuretic Peptide (NT-proBNP), high sensitivity Troponin I (hsTnI), Haemoglobin-A1c (HbA1c), Lipids, High density lipoprotein-cholesterol (HDL-c), Ferritin, Creatine Kinase (CK), D-Dimer. Tier 1 assays will be performed on samples collected at Baseline (Day 0) and Day 3 (+/-1 day).

- Tier 2 data will include Tier 0 (+/- 1) data plus the results of exploratory biomarker analyses planned by University of Glasgow co-investigators (e.g. cytokine/chemokines, metabolomics)
- Tier 3 data will only be generated by approved translational studies, which may include host genomic profiling. Approval will be via established governance infrastructure including the NHS GG&C Bio-Repository Management Group.

4.5 Imaging Data and other aligned COVID19 datasets

In parallel with ASTERIX, all COVID19 imaging data is being curated within iCAIRD, primarily for development of Artificial Intelligence (AI) tools for detection of COVID pneumonia. Chest imaging is also being labelled within the clinical record by NHS GG&C radiologists as 'COVID pneumonia present or not'. These NHS GG&C text reports will be used to define COVID pneumonia as described in Section 4.2.

The chest radiograph and computed tomography (CT) images stored in this archive will also be available and will be analysed using qure.ai (India) and Vida LungPrint Discovery software (USA). LungPrint is a fully automated, AI-powered analysis that will generate categorical input variables describing the extent of COVID lung injury for input to ASTERIX models, subject to appropriate approvals (Blyth and Lowe are current lead/co-investigators on iCAIRD projects). A range of other linked data will also be available across the continuum of care. Embedded within clinical workflows, recently established data collection workflows facilitate data collection from the patient's home (via the Scottish Ambulance Service), to community care (via COVID Community Assessment Centres) through to the secondary care ASTERIX environment.

5. ASSESSMENT OF SAFETY

There will be no risk to the patient from ASTERIX study activities, therefore no formal safety reported will be performed.

6. STATISTICS AND DATA ANALYSIS

6.1 Sample Size

The target sample size of 600-1200 suspected COVID19 cases will yield 400-800 confirmed cases (assuming 1/3 cases receive an alternative diagnosis, based on current clinical experience in NHS GG&C). Based on estimates reported in recent series, the following frequencies of clinical progression events are expected:

(1) COVID19 pneumonia: 75%, based on the reported prevalence of COVID19 pneumonia in a snap-shot audit of NHSGGC COVID19 admissions (via Dr Mark Hall, Consultant Radiologist, NHS GG&C, 4th May 2020)

(2) Development of respiratory failure, sufficient to prompt an escalation in care to *at least* Step 6 of the WHO 10-point ordinal scale: 22.5% of total, based on 30% of COVID pneumonia (event (1))

(3) Death: 13.3% of total, based on 59% of event (2).

The following table shows, for various sample sizes within the target range, the number of candidate predictor parameters that can be included in modelling to build reliable models for each progression event which minimise overfitting and maintain a global shrinkage factor >0.9.

Total # of suspected cases	**Total # confirmed cases***	**Number of patients with each event type**			**Number of predictor parameters for reliable models for each event type**		
		(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
1200	800	600	180	106	10	9	8
900	600	450	135	80	7	7	6
600	400	300	90	53	5	5	4

*Assuming 1/3 of cases receive an alternative diagnosis

Effective model building may be possible with a sample size at the lower end of the target range but this will depend on the event rates observed and proportion of suspected cases that are subsequently confirmed. These event rates and COVID19 diagnosis rates will be monitored throughout the study. The current predictions represent the best available data in what is a highly dynamic clinical environment. The following table shows the power to detect differences in event rates at the 5% 2-

sided level of significance based on a dichotomous predictor that splits 50:50 in the patient group.

Total N (confirmed cases)	N per group	Power to compare…		
		70% vs 80%	18% vs 27%	9% vs 17.5%
1000	500	96%	93%	98%
800	400	91%	86%	95%

6.2 Statistical analysis plan

Three 'clinical progression events' will be used to identify high-risk populations for later targeting in multiple high impact clinical trials. The identification of these high-risk populations will be based on multivariable models in which the dependent variables will be one of the following:

1) Development of COVID pneumonia
2) Development of respiratory failure, sufficient to prompt an escalation in care to *at least* Step 6 of the WHO 10-point ordinal scale
3) Death from any cause.

Definitions of each 'clinical progression event' are summarized in Figure 3. Given the likely variation in ICU resource we have explicitly chosen not to use ICU admission nor mechanical ventilation *per se* for progression (2). The definition used, based on objective evidence of severe respiratory failure will be defined using data collected by the bedside on a daily basis by clinical teams (see Section 4.1). Importantly, this definition will also capture patients not deemed suitable for ICU due to age or comorbidities. Such patients will certainly have deteriorated, and their data could contribute significantly to accurate model building. Moreover, such patients would also be potentially eligible for intervention trials.

A variety of model building approaches may ultimately be applied to the data, but the primary approach outlined below. Candidate predictors for each Clinical Progression event will be identified *a priori* and checked by univariate regression, and will include routinely available data such as demographics, virus characteristics and laboratory measures (baseline and change from admission) that provide direct

or surrogate measures of implicated physiological processes (e.g. CRP for IL-6[9,10], Troponin for myocardial injury[11], arterial PaO_2 and chest radiographs for acute lung injury[12]). Multivariable models will be built using logistic regression techniques [13-15] using a purposeful selection approach. Candidate predictors associated with p-value <0.25 in univariate regression will be included in these models, assuming no strong collinearity exists. Fractional polynomial techniques will be used to explore how to best include any continuous variables in the model. The fit of the final model will be assessed using the Hosmer Lemeshow Test. A bootstrapped calibration slope will be used as a shrinkage factor to adjust the model for potential overfitting and to calculate adjusted regression co-efficiences. The predicted performance of any developed models will be assessed using a Receiver Operator Characteristic curve. Any model developed will be subject to validation in an independent data set. Models may subsequently be updated by results emerging from embedded translational studies.

7. STUDY CLOSURE

The study will end 6 months after the last patient is recruited.

7.1 End of Study Notification/Declaration of the End of a Study Form

An end of study notification will be submitted to the ethics committee within 90 days using the 'Declaration of the end of a study' form. However if the study is terminated either before the planned conclusion of the trial, as specified in the protocol, the ethics committee will be notified in writing of the termination of the trial within 15 days of the date of termination with a clear explanation of reasons and details of follow-up measures needed.

7.2 Temporary Halt of the Study

If recruitment to the study needs to be temporarily halted for reasons not specified in the protocol the Sponsor will inform the REC immediately and at the latest within 15 days from when the study is temporarily halted. This includes studies where the stoppage was not envisaged in the approved protocol and where there is an

intention to resume it. It does not include trials where recruitment may be temporarily halted for logistical reasons such as trial team availability. The notification will be made as a substantial amendment and will clearly state what activities have been halted and the reasons for this. To restart a trial that has been temporarily halted the Sponsor will make a request as a substantial amendment providing evidence that it is safe to restart the trial. If the Sponsor decides not to recommence the trial the REC will be notified in writing within 15 days of the decision, using the end-of-trial declaration form.

7.3 Early Termination of a Study

In the case of early termination, the Chief Investigator (CI) will notify the end of a study to the Sponsor immediately and at the latest within 15 days after the trial is halted, explaining the reasons and describing the follow-up measures. This does not include trials that complete early because full recruitment has been achieved.

8. DATA HANDLING

8.1 Data Capture and Linkage

ASTERIX data will be stored in a format suitable for transfer to an electronic Case Report Form for commercial trial purposes or use by translational researchers using linked translational samples in the ASTERIX biobank (see below). An agreed data dictionary from NHS GG&C Safe Haven (a mirrored clinical data repository) will be made available to all collaborators receiving linked anonymised data.

8.2 Record Retention

Archiving of the essential documents should be performed by both the participating site and Sponsor/CRUK CTU. Participating sites are responsible for archiving their documentation and should follow the requirements of their Research and Development Office in conjunction with advice from the CRUK CTU and Sponsor regarding the duration of document retention. Sites should not archive their trial documentation until they have been instructed by the CRUK CTU or Sponsor that they are able to do so. Where possible, at the time of archiving, sites will be notified of the archiving retention period. If this is not confirmed at the time of archiving,

sites should not destroy archived documentation until authorisation is given from the Sponsor. The Sponsor and CRUK CTU will be responsible for archiving the Trial Master File (TMF) and all other essential trial documentation that is not held at participating trial sites as per their applicable Standard Operation Procedures.

9. STUDY MANAGEMENT

Sites wishing to participate in the trial should contact the Chief Investigator. A Principal Investigator (PI) must lead the study at each site and they will be responsible for providing CRUK Glasgow CTU with all core documentation. Protocol training will be given to sites via initiation slides that will be provided to sites prior to the study opening at that site. Once all the documentation is received at CRUK Glasgow CTU an initiation call will be performed and after this the site will be contacted by email or fax when they are activated and are able to initiate the study.

9.1 Trial Management Group (TMG)

The study will be coordinated from CRUK Glasgow CTU by the TMG. The TMG normally includes those individuals responsible for the day-to-day management of the study. Members of the TMG include the Chief Investigator, Co Investigators, Project Manager, Trial Statistician, Clinical Trial Coordinator and Patient Representative as well as other key individuals central to the management, oversight and delivery of the study. The role of the group is to monitor all aspects of the conduct and progress of the study, ensure that the protocol is adhered to and take appropriate action to safeguard participants and the quality of the trial itself.

9.2 COVID-19 Umbrella Trial Steering Committee (COVID-19 UTSC)

The role of the COVID-19 UTSC is to provide overall supervision of the study and ensure that it is being conducted in accordance with the principles of GCP and the relevant regulations. The UTSC should agree any significant protocol amendments, provide advice to the investigators on all aspects of the trial and have members who are independent of the investigators, in particular, an independent chairperson. Decisions about continuation or termination of the trial or substantial amendments to the protocol are usually the responsibility of the UTSC.

10. STUDY MONITORING

This study will be risk assessed by NHS GG&C R&D department and audit of this study will be determined by the results of the risk assessment. There will be no routine study monitoring for this study.

11. PROTOCOL AMENDMENTS

Any change in the study protocol will require an amendment. Any proposed protocol amendments will be initiated by the CI following discussion with the sponsor and any required amendment forms will be submitted to the appropriate ethical approval granting body and sponsor. The CI will liaise with study sponsor to determine whether an amendment is non-substantial or substantial. All amended versions of the protocol will be signed by the CI and Sponsor representative. Before the amended protocol can be implemented favourable opinion/approval must be sought from the original reviewing REC and R&D office(s).

12. ETHICAL CONSIDERATIONS

The study will be carried out in accordance with the World Medical Association Declaration of Helsinki (1964) and its revisions (Tokyo [1975], Venice [1983], Hong Kong [1989], South Africa [1996] and Edinburgh [2000]). The study will not begin until R&D approval has been granted by the NHSGGC Safe Haven delegated approval system (GSH20RM003), study specific REC approval and NHS R&D Management Approval has been granted. The ASTERIX REC approval will cover use of the routinely collected data and surplus biological materials described herein (see page 20) for a) the research purposes described in Tiers 0–2 and b) for banking for use later studies (Tier 3). Additional approval for access to banked Tier 3 samples will however require approval by the NHS GG&C Bio-Repository Management Group.

13. CONFIDENTIALITY

All information collected during the course of the study will be kept strictly confidential. Information will be held securely on paper and electronically at the

CRUK Glasgow CTU. The CRUK CTU will comply with all aspects of the 1998 Data Protection Act and operationally this will include:

- Appropriate storage, restricted access and disposal arrangements for patient's personal and clinical details
- Where anonymisation of documentation is required, sites are responsible for ensuring only the instructed identifiers are present before sending to CRUK Glasgow CTU.

If a participant withdraws verbal consent from further collection of data, their samples will remain on file and will be included in the final analysis unless they specifically withdraw consent for this.

14. INSURANCE AND INDEMNITY

No special insurance is in place for patients in this study other than standard NHS liability insurance providing indemnity against clinical negligence. This does not provide cover for non-negligence e.g. harm caused by an unexpected side effect of participating in a trial. The sponsors have responsibility for ensuring that financial cover for damages or compensation arising from no fault harm is available to patients, where applicable. Cover for this study has been agreed under the current policy. The Hospital Trust/Health Board at each participating site is responsible for:

1. Acts and omissions of its own staff and others engaged by it, including the Clinical Trials Unit and PI
2. Ensuring the appropriate insurance administered by the National Health Service Litigation Authority is in place
3. Ensuring non-NHS employees involved in the study have Honorary Contracts with the Trust/Board to cover access to patients and liability arrangements.

15. FUNDING

ASTERIX is funded by the Chief Scientist Office, with support from the CRUK Glasgow CTU (core funded by Cancer Research UK), and underpinning infrastructure provided by NHS GG&C, the NRS Glasgow Biorepository and NHS Safe Haven data archive.

16. ANNUAL REPORTS

An annual progress report will be submitted to the funder. Annual reports will be submitted to the ethics committee and sponsor with the first submitted one year after the date that all trial related approvals are in place.

17. DISSEMINATION OF FINDINGS

Study findings will be disseminated via peer-reviewed scientific publications and by presentation at national and international meetings. The Co-CIs are responsible for approving the content and dissemination of all publications, abstracts and presentations and for assuring the confidentiality and integrity of the study.

18. REFERENCES

1. Weiss P, Murdoch DR. Clinical course and mortality risk of severe COVID-19. Lancet. 2020;395(10229):1014–5.

2. Wu Z, McGoogan JM. Characteristics of and Important Lessons from the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases from the Chinese Center for Disease Control and Prevention. Jama. 2020;323(13):1239–42.

3. Cube M von, Grodd M, Wolkewitz M, Hazard D, Lambert J. Harmonizing heterogeneous endpoints in COVID-19 trials without loss of information - an essential step to facilitate decision making. Medrxiv. 2020;2020.03.31.20049007.

4. Jin X, Pang B, Zhang J, Liu Q, Yang Z, Feng J, et al. Core Outcome Set for Clinical Trials on Coronavirus Disease 2019 (COS-COVID). Engineering-london. 2020;

5. Valle JW, Evans TRJ, McKay CJ, Dixon-Hughes J, Paul J, Stobo J, et al. Precision-Panc Master Protocol: Personalizing treatment for pancreatic cancer ISRCTN14879538—Part of Precision-Panc United Kingdom. J Clin Oncol. 2019;37(4_suppl):TPS460–TPS460.

6. Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf.

7. Wang M, Cao R, Zhang L, Yang X, Liu J, Xu M, et al. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. Cell Res. 2020;30(3):269–71.

8. Chen C, Huang J, Cheng Z, Wu J, Chen S, Zhang Y, et al. Favipiravir versus Arbidol for COVID-19: A Randomized Clinical Trial. Medrxiv. 2020;2020.03.17.20037432.

9. Zhang C, Wu Z, Li J-W, Zhao H, Wang G-Q. The cytokine release syndrome (CRS) of severe COVID-19 and Interleukin-6 receptor (IL-6R) antagonist Tocilizumab may be the key to reduce the mortality. Int J Antimicrob Ag. 2020;105954.

10. Mehta P, McAuley DF, Brown M, Sanchez E, Tattersall RS, Manson JJ, et al. COVID-19: consider cytokine storm syndromes and immunosuppression. Lancet. 2020;395(10229):1033–4.

11. Chen T, Wu D, Chen H, Yan W, Yang D, Chen G, et al. Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. Bmj. 2020;368:m1091.

12. Matthay MA, Zemans RL, Zimmerman GA, Arabi YM, Beitler JR, Mercat A, et al. Acute respiratory distress syndrome. Nat Rev Dis Primers. 2019;5(1):18.

13. Hosmer JDW, Lemeshow S. Applied Logistic Regression. 2004. (John Wiley & Sons).

14. Riley RD, Snell KI, Ensor J, Burke DL, Harrell FE, Moons KG, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. Stat Med. 2018;38(7):1276–96.

15. Bursac Z, Gauss CH, Williams DK, Hosmer DW. Purposeful selection of variables in logistic regression. Source Code Biology Medicine. 2008;3(1):17.

16. Dunning JW, Merson L, Rohde GGU, Gao Z, Semple MG, Tran D, et al. Open source clinical science for emerging infections. Lancet Infect Dis. 2014;14(1):8–9.

APPENDIX 1. ASTERIX CLINICAL DATASET

The following data will be collected via the NHS GGC COVID19 assessment tool (see daysix.covid.co).

Data Field	Initial Assessment (ED / SATA)	Daily Ward Round
Demographics (including age, postcode)	X	X
Verbal Consent		X
Location within the Hospital	X	X
Living Arrangements (Lives alone, Carer etc.)	X	X
Clinical Frailty Score (Rockwood)	X	X
Documentation of Capacity +/- any Anticipatory Care Plan	X	X
Co-morbidities (aligned with ISARIC)	X	X
Home Respiratory Support (e.g. LTOT, NIV)	X	X
Smoking Status (including vaping)	X	X
COVID Symptom Start	X	X
COVID Result	X	X
COVID Trial Summary		X
Nutrition Details		X
Palliative Care / MDT		X
COVID19 Symptoms (aligned with ISARIC)	X	X
SARS-CoV-2 Exposure (Co-habitation/Health Care Worker)	X	X
Evidence of COVID19 Pneumonia	X	X
Clinical Examination Findings	X	X
Vital Signs (e.g. HR, BP, RR)	X	X
Medication (Key Summary)	X	X
COVID19 Escalation Plan		X
Blood / Imaging Requests	X	X
Blood Gas Findings ($Pa0_2$ / FiO_2 etc.)	X	X
Respiratory Interventions (e.g. Proning, O_2 (with FiO_2), CPAP, Invasive Ventilation, ECMO)	X	X
Circulation Interventions (e.g. IV fluids, Inotropes)	X	X
COVID Clinical Trial Participation	X	X

APPENDIX 2. TIER 0 ROUTINELY AVAILABLE DATA

The following investigations constitute the NHS GG&C COVID19 Trak Care Order Set that will be available in all ASTERIX participants:

Blood Tests

Haematology

- FBC
- ESR
- Coagulation

Biochemistry

- U&E
- LFT
- CRP
- Bone profile
- Magnesium
- LDH
- Bicarbonate
- Glucose
- Lactate

Radiology

- Chest Radiograph

Respiratory Secretions

Microbiology

- Blood cultures

Virology

- COVID swabs

APPENDIX 3. TIER 1 ASSAYS

The following assays will be conducted in NHS GG&C laboratories using banked surplus blood samples. Tier 1 assays will be performed on samples collected at Baseline (Day 0) and Day 3 (+/-1 day) only.

- N-terminal pro-Brain Natriuretic Peptide (NT-proBNP)
- high sensitivity Troponin I (hsTnI)
- Haemoglobin-A1c (HbA1c)
- Lipid Profile
- High density lipoprotein-cholesterol (HDL-c)
- Ferritin
- Creatine Kinase (CK)
- D-Dimer
- LDH
- Vitamins D, B1, B2, B6, B12
- Folate
- Magnesium