

## 1. Protocol details

### **1.1 PROTOCOL TITLE:**

Personalised simulation technologies for optimising treatment in the inter

### **1.2 Names (titles), roles and contact details of:**

#### **Sponsor**

Guy's and St Thomas' NHS Foundation Trust

Elizabeth Bruna  
16<sup>th</sup> floor Tower Wing  
Guy's Hospital  
Great Maze Pond  
London  
SE1 9RT

[r&d@gstt.nhs.uk](mailto:r&d@gstt.nhs.uk)

02071887188

#### **Chief Investigator**

Dr Luigi Camporota

Intensive Care Research Office  
Ground Floor South Wing Block C,  
Lane Fox Unit,  
St Thomas' Hospital,  
Westminster Bridge Road,  
London,  
United Kingdom  
SE1 7EH

[luigi.camporota@gstt.nhs.uk](mailto:luigi.camporota@gstt.nhs.uk)

020 7188 3046

### ***1.3 Protocol details***

Version number 1.0

Draft

Date 05/02/2020

## 2 Signature Page

The Chief Investigator and the R&D (sponsor office) have discussed this and agree to perform the investigations and to abide by this protocol

The investigator agrees to conduct the trial in compliance with the approved Data Protection Act (1998), the Trust Information Governance Policy (or Research Governance Framework (2005' 2<sup>nd</sup> Edition; as amended), the regulatory requirements as amended.

Chief investigator

Dr Luigi Camporota

---

Signature

Sponsor Representative

R&D to Add

GSTFT

---

Signature

## Contents Page

1.	Protocol details.....	
1.1	PROTOCOL TITLE:.....	
1.2	Names (titles), roles and contact details of: .....	
1.3	Protocol details.....	
2	Signature Page .....	
	Contents Page .....	
3	List of Abbreviations and Definitions .....	
4	Summary/Synopsis.....	
5	Introduction.....	
6	Study objectives and purpose .....	
7	Study design & Flowchart .....	
7.1	Study Design.....	
7.2	Flowchart.....	
8	Subject selection .....	
8.1	Inclusion Criteria.....	
8.2	Exclusion Criteria .....	
9	Study procedures .....	
9.1	Subject recruitment and screening.....	

11.1.1	Demographics .....
11.1.2	Physiological .....
11.1.3	Ventilator.....
11.1.4	Blood Gas Analysis.....
11.1.1	ECMO .....
11.2	Format and Scale of Data .....
11.3	Methodologies of Data Collection and Generation .....
11.4	Data Quality and Standards .....
11.5	Data handling and record keeping.....
11.5.1	GSTFT .....
11.5.2	Data Repository.....
11.5.3	Academic Partners.....
11.6	Metadata standards and data documentation .....
11.6.1	Data Set Directory Structures and Filenames.....
11.6.2	Metadata Descriptive Files .....
11.7	Data preservation strategy and standards .....
11.7.1	GSTFT .....
11.7.2	Data Repository.....
11.7.3	Academic Collaborators .....
11.8	Data Security and Confidentiality of Potentially Disclosive Data .....

14	Financing and Insurance.....
15	Reporting and dissemination.....
	References.....

### 3 List of Abbreviations and Definitions

GSTFT	Guy's and St Thomas' NHS Foundation Trust
UON	University of Nottingham
UWAR	University of Warwick

## 4 Summary/Synopsis

Title	<i>Personalised simulation technology for intensive care unit</i>
Protocol Short Title/Acronym	PSTOTICU
Protocol Version number and Date	0.1.4 21/08/2019
IRAS Number	266780
REC Reference	NA
Study Duration	<i>2 years</i>
Sponsor name	Guy's and St Thomas' NHS Foundation Trust
Chief Investigator	Dr Luigi Camporota
Funder	Engineering and Physical Sciences Research Council
Medical condition or disease under investigation	Respiratory Failure
Purpose of research	This study aims to further develop our models, by the fitting of those models to real data of critical care patients.
Primary objective	To further develop our simulation models of recently proposed therapeutic strategies
Secondary objective (s)	To integrate data-streams and combine them with our existing modelling technology to simulate individual patients and treatment strategies for imminent problems. To develop mathematical models that can explore the "design space" of possible treatments for individual patient in simulation, and which can be suggested for evaluation
Number of Episodes	5,000
Study Type	Observational
Endpoints	Not Applicable
Main Inclusion Criteria	This project will include critical care patients



## 5 Introduction

This project aims to further develop and refine in silico physiological models to the physiological and treatment data of critical care patients. The development is fitting the model to, and its description of, the physiological specific interventions. An example would be how a patient's physiological increase in positive end expiratory pressure (PEEP). Data for this example is physiological and treatment data pre and post ventilator changes (e.g., PEEP) the clinical team, for a clinically appropriate time period.

Computer simulation offers a new approach to traditional medical research suited to investigating treatment of critical illness. Critically ill patients provide great detail, providing extensive, high quality data streams for model development and patient-matching. Models based on this data can incorporate very complex physiology and be validated against responses of individual patients, for use as investigative tools.

In contrast to trials on animal models and humans, in silico models of disease pathology are completely configurable, reproducible and reusable. Combinations of treatments, can be applied to the same patient or subset of patients to understand mode of action, quantitatively compare effectiveness in various interventions for particular clinical objectives and particular patient groups.

It is widely accepted that modification of existing mechanical ventilation strategies and novel physical treatment strategies could significantly reduce pulmonary impairment, and mortality rates, particularly if combined with novel pharmacological interventions.

To achieve this we are undertaking a collaborative project between three different and complimentary areas of expertise.

The three institutions involved in this project are:

## 6 Study objectives and purpose

This project is a collaborative endeavour between Guy's and St Thomas academic collaborators at the University of Nottingham and the University of Manchester. The project aims to further develop and refine in silico physiological models, by the fitting of clinical physiological and treatment data of critical care patients.

The project's primary objective is to further develop our simulation platform to include a range of commonly used therapeutic interventions within the intensive care unit (ICU). The interventions include prone positioning, optimising the ventilator settings to reduce breathing effort with mechanical ventilation, application of recruitment manoeuvres to cause and maintain an "open lung", and Extra-corporeal membrane oxygenation (ECMO). These interventions have been chosen on the basis of amenability to simulation, optimisation, possessing a strong potential for providing improvement in current usage in many ICU's worldwide.

Integration of available data within the ICU with our existing modelling tools will enable a real time simulation of individual patients and treatments. Sufficient data sets for the models would enable systematic and efficient exploration of the possible outcomes for an individual patient in simulation, allowing optimised treatment strategies to be evaluated by the clinician. Possible therapeutic interventions to be explored include: designed pulmonary recruitment (alveolar re-opening) manoeuvres, ventilation strategies (e.g. airway-pressure release ventilation, low tidal volume, high positive end-expiratory pressure), inspiratory flow waveform modification, and closed-loop ventilatory control.

To facilitate the further development of the simulation platform requires the collection of curated data sets representative of the therapeutic interventions of interest. A data repository will be created to catalogue and share collated data sets between collaborators. The repository will enforce the oversight and governance ensuring the requirements for data protection against unlawful or unauthorised processing, access, loss, destruction, and retention.

At this stage of the project, the simulation platform will not be accessible to the public.

## 7 Study design & Flowchart

### 7.1 Study Design

This study aims to further develop and refine *in silico* physiological models to the physiological and treatment data of critical care patients.

Of specific interest to model development is fitting the model to, physiological response of patients to specific interventions. An example physiological parameters respond to an increase in PEEP. Data for this example physiological and treatment data pre and post ventilator changes (e.g., PEEP) from the clinical team, for a clinically appropriate time period. This would come from

Data required for this study will therefore be truncated case report data from short time periods, on the order of 12 hours duration. The exact length of time will be determined by the clinical scenario, but would be for a clinically appropriate

No attempt will be made to capture data from entire patient cohorts or time periods.

There is no intention to implement any study specific interventions or treatments. The study will be retrospective, pragmatic and limited to that data recorded through the clinical team, routinely provided to all patients.

Within this project, there is no intention for physiological models development. Clinicians will have no access or exposure to the simulations and predictions generated by the models that this project aims to develop.

For the purposes of this document, a **data set** is defined as a collated set of data for a single subject and one intervention.

### 7.2 Flowchart

The collaborative nature of this study requires that data be managed across multiple environments:

## GSTFT

Electronic Health Records  
&  
Physiological Data Capture

Capture &  
Collation

GSTFT Network Share Drive

Anonymisation

GSTFT Network Share Drive

Upload

## Data Repository

GSTFT SharePoint Site

Download

## Academic Centres

Project Computer

## 8 Subject selection

Subjects will be drawn from the adult critical care population at GSTFT. Subjects will be receiving mechanical ventilation.

It is anticipated that this project will require several iterations of development, with each iteration examining a different treatment intervention. The results of each iteration will inform the exact nature of data collection in the subsequent iteration. The number of development and analysis iterations that will be required is dependent on the results of the first iteration of development.

As noted in section 7.1 collated data sets will be episodic in nature with each episode lasting for the order of twelve hours. It is entirely possible that the admission and discharge of a patient may span several episodes of interest, or the patient may receive a similar intervention across several episodes. These episodes may therefore be contributed under differing study ids to ensure each episode is catalogued in the data repository.

### 8.1 Inclusion Criteria

- At least 18 years of age
- Patients admitted to GSTFT Intensive Care between the 1st of January 2019 and 31/03/2019
- Receiving mechanical ventilation

### 8.2 Exclusion Criteria

- Pregnancy or lactation.

## **9 Study procedures**

### ***9.1 Subject recruitment and screening***

Episodes of interest will be identified retrospectively by querying existing attributes. As an example, patient treatment data could be queried for episode setting is increased from 5 to 10.

Data subjects will not be approached as data is collected retrospectively from existing data sources. We will not be requested any additional data or resources.

Consent will not be sought from individual patients due to the non-interventive nature of the data collection.

Data collection will be limited to that data recorded as part of routine care from 01/01/2010 and 31/03/2019

### ***9.2 Schedule of assessments for each visit***

Data is available from the electronic health records and the data export technology employed in the routine care of critically ill patients.

No schedule of assessment is planned due to the retrospective nature of the health care records.

### ***9.3 Follow up Procedures***

There are no follow up procedures planned for this study.

### ***9.4 End of Study Definition***

Completion of the project is defined as the time point five years after completion of activities.

## **10 Assessment of Safety**

Study procedures for this study consist of data collection from the electronic health records (EHR) which are part of the routine care provided to patients. It is anticipated that this study poses no greater risk to patient safety than that caused by their routine clinical care.

### ***10.1 Study Steering Committee***

There will be no monitoring, steering or safety committees set up for this study.

### ***10.2 Ethics & Regulatory Approvals***

Appropriate HRA and REC will be in place before the study commences.

## 11 Data

### ***11.1 Data to be collected***

Data will be collected within four domains:

- Demographics
- Physiological
- Ventilation
- Blood gas analysis

Demographic data will be collected from the health care record, and is of

Physiological data is continuously captured by monitoring equipment (Philips) within ICU, and recorded on a regular basis into an electronic health record. Data can also be captured remotely from the continuous stream captured

Ventilator parameters and patient ventilatory parameters are measured at the bedside in the ICU by the patient ventilators (Dräger V500). This data is recorded on a regular basis into an electronic health record system by clinical staff. Data can also be captured from the ventilators.

Blood gas analysis results are collected routinely as part of the care of ICU patients and are recorded in their electronic health record. Data will be collected retrospectively from the electronic health record.

Data points that will be collected from each domain are detailed below. This list is intended to identify specific data fields, alongside a short description of that data field. It is not realistic that only subsets of the entire list of potential data fields will be collected. Data will be present in individual data sets, with availability dependent on differing patient characteristics.

#### **11.1.1 Demographics**

STUDY\_ID

Unique study data set identifier



CENTRAL_TEMP_MONITOR_C	Central temperature measurement
PERIPHERAL_TEMP_MONITOR_C	Peripheral temperature measurement
ECG_HEART_RATE_MIN	ECG Heart Rate
HEART_RATE_MONITOR_MIN	Monitor Heart Rate
ECG_RHYTHM_MONITOR	ECG cardiac rhythm
PVC_RATE_MONITOR_MIN	Premature Ventricular Contractions
ECTOPIC_BEAT_MONITOR_STR	Premature Ventricular Contractions
ECT_RATE_MONITOR_MIN	ECG Ectopic Rate
SPO2_MONITOR monitoring	Photoplethysmogram - Non-invasive
RESP_RATE_MONITOR_MIN	Respiratory Rate
NIBP_SYS_MMHG	Non-invasive systolic blood pressure
NIBP_DIA_MMHG	Non-invasive diastolic blood pressure
NIBP_MEAN_MMHG	Non-invasive mean blood pressure
ABP_SYS_MMHG	Invasive systolic blood pressure
ABP_DIA_MMHG	Invasive diastolic blood pressure
ABP_MEAN_MMHG	Invasive mean blood pressure
ART_SYS_MMHG	Invasive systolic blood pressure
ART_DIA_MMHG	Invasive diastolic blood pressure
ART_MEAN_MMHG	Invasive mean blood pressure

CENTRAL_TEMP_ICIP_C	Central temperature measurement
PERIPHERAL_TEMP_ICIP_C	Peripheral temperature measurement
HEART_RATE_ICIP	Heart Rate
CARDIAC_RHYTHM_ICIP	Cardiac Rhythm
SPO2_ICIP	Photoplethysmogram
NIBP_SYS_ICIP_MMHG	Non-invasive systolic blood pressure
NIBP_DIA_ICIP_MMHG	Non-invasive diastolic blood pressure
NIBP_MEAN_ICIP_MMHG	Non-invasive mean blood pressure
ABP_SYS_ICIP_MMHG	Invasive systolic blood pressure measurement
ABP_DIA_ICIP_MMHG	Invasive diastolic blood pressure measurement
ABP_MEAN_ICIP_MMHG	Invasive mean blood pressure measurement
ART_SYS_ICIP_MMHG	Invasive systolic blood pressure measurement
ART_DIA_ICIP_MMHG	Invasive diastolic blood pressure measurement
ART_MEAN_ICIP_MMHG	Invasive mean blood pressure measurement
PABP_SYS_ICIP_MMHG	Systolic pulmonary artery blood pressure
PABP_DIA_ICIP_MMHG	Diastolic pulmonary artery blood pressure
PABP_MEAN_ICIP_MMHG	Mean pulmonary artery blood pressure
CVP_MEAN_ICIP_MMHG	Invasive central venous pressure measurement
CARDIAC_OUTPUT_ICIP_L_MIN	Cardiac Output
CARDIAC_INDEX_ICIP_L_MIN_M2	Cardiac Index

P_LOW_CMH2O	Pressure Low
P_HIGH_CMH2O	Pressure High
P_MEAN_CMH2O	Pressure mean
P_MIN_CMH2O	Pressure minimum
P_PEAK_CMH2O	Pressure peak inspiratory
P_PLATEAU_CMH2O	Pressure plateau
MIN_VOL_EXP_L_MIN	Expired minute volume
MIN_VOL_EXP_SPON_L_MIN	Expired minute volume - spontaneous
MIN_VOL_EXP_MAND_L_MIN	Expired minute volume - mandatory
MIN_VOL_LEAK_L_MIN	Minute volume leak
TIDAL_VOLUME_MLS	Tidal Volume
TIDAL_VOLUME_INSP_MLS	Tidal Volume Inspired
TIDAL_VOLUME_EXP_MLS	Tidal Volume Expired
TIDAL_VOLUME_SPON_MLS	Tidal Volume Spontaneous
TIDAL_VOLUME_INSP_SPON_MLS	Tidal Volume Inspired Spontaneous
TIDAL_VOLUME_EXP_SPON_MLS	Tidal Volume Expired Spontaneous
TIDAL_VOLUME_MAND_MLS	Tidal Volume Mandatory
TIDAL_VOLUME_INSP_MAND_MLS	Tidal Volume Inspired Mandatory
TIDAL_VOLUME_EXP_MAND_MLS	Tidal Volume Expired Mandatory
RESISTANCE_CMH2O_L_S	Resistance

E_IE	Expiratory component IE ratio
I_IE_SPON	Inspiratory component IE ratio -
E_IE_SPON	Expiratory component IE ratio -
ELASTANCE_CMH2O_L	Elastance
RSB_S_L	Rapid shallow breathing index
COMP_DYN_MLS_CMH2O	Dynamic compliance
COMP_20_MLS_CMH2O	Compliance 20
TIME_CONSTANT_S	Time constant
TIME_CONSTANT_EXPIRED_S	Time constant expired
P01_CMH2O	Airway occlusion pressure
NIF_CMH2O	Negative Inspiratory Force
EIP_CMH2O	End inspiratory Pressure
FIO2_PC	Fractional Inspired Oxygen
VT_CO2_MLS	Tidal CO2 production
V_CO2_MLS_MIN	CO2 Production
VOLUME_DS_MLS	Deadspace Volume
ET_CO2_KPA	End tidal CO2 kPa
ET_CO2_VOLPC	End tidal CO2 vol%
CO2_SLOPE_KPA_L	CO2 Slope
VENTILATOR_MODE	Ventilator Mode

T_LOW_SETTING_S	Time low - Ventilator Setting
T_INSP_MAX_SETTING_S	Maximum Inspiratory Time - Ventilator Setting
T_LOW_MAX_SETTING_S	Maximum time low - Ventilator Setting
SLOPE_SETTING_S	Slope setting - Ventilator Setting
FLOW_SETTING_L_MIN	Flow setting - Ventilator Setting
CONST_FLOW_L_MIN	Constant flow - Ventilator setting
RESPIRATORY_RATE_SETTING_MIN	Respiratory rate - Ventilator Setting
P_SUPPORT_SETTING_CMH2O	Pressure support - Ventilator Setting
TIDAL_VOLUME_SETTING_MLS	Tidal volume setting - Ventilator Setting
FLOW_TRIGGER_SETTING_L_MIN	Flow trigger - Ventilator Setting
ATC_COMPENSATION_SETTING_PC	ATC compensation - Ventilator Setting
FIO2_SETTING_PC	Fractional Inspired Oxygen - Ventilator Setting
INSP_TERM_SETTING_PIF_PC	Inspiratory term % Peak Inspiratory Flow
EXP_TERM_SETTING_PIF_PC	Expiratory term % Peak Expiratory Flow

#### 11.1.4 Blood Gas Analysis

STUDY_ID	Unique study data set identifier
DATETIME	ISO 8601 Date & Time
BLOOD_GAS_TYPE	Blood gas type
PCO2_KPA	Partial Pressure CO2

NA_MMOL_L	Sodium concentration
CL_MMOL_L	Chloride concentration
HAEMOGLOBIN_G_L	Haemoglobin concentration
OXY_HAEMOGLOBIN_PC	Oxyhaemoglobin
CARBOXY_HAEMOGLOBIN_PC	Carboxyhaemoglobin
H_HAEMOGLOBIN_PC	Deoxyhaemoglobin
METHE_HAEMOGLOBIN_PC	Methaemoglobin
GLUCOSE_MMOL_L	Glucose concentration
LACTATE_MMOL_L	Lactate concentration
TEMP_BLOOD_GAS_C	Temperature blood gas
FIO2_BLOOD_GAS_F	FiO2 blood gas
R_BLOOD_GAS	Respiratory quotient
P50_BLOOD_GAS_KPA	P50
BASE_EXCESS_MMOL_L	Base Excess
BASE_EXCESS_ACT_MMOL_L	Actual Base Excess
HCO3_MMOL_L	Bicarbonate concentration
HCO3_STAND_MMOL_L	Standardised bicarbonate concentration

### 11.1.1 ECMO

STUDY_ID	Unique study data set identifier
----------	----------------------------------

SWEEP_FO2_ECMO_ICIP_F	Sweep Gas Fractional O2 concentration
SWEEP_FLOW_ECMO_ICIP_L_MIN	Sweep gas flow
P_ACCESS_ECMO_CHELP_MMHG	Access pressure - Cardiohelp data
P_PREOXY_ECMO_CHELP_MMHG	Pre membrane pressure - Cardiohelp data
P_POSTOXY_ECMO_CHELP_MMHG	Post membrane pressure - Cardiohelp data
P_TMP_ECMO_CHELP_MMHG	Transmembrane pressure - Cardiohelp data
Q_BLOOD_ECMO_CHELP_L_MIN	ECMO Blood Flow - Cardiohelp data
PUMP_ECMO_CHELP_RPM	ECMO Pump RPM - Cardiohelp data
SVO2_ECMO_CHELP_PC	Pre Membrane Haemoglobin saturation

## ***11.2 Format and Scale of Data***

The requirements and nature of this project mean that all study data will

Data will be collated to comma separated value (CSV) format. CSV is accessible by all major data processing software platforms e.g. spreadsheets, database software. The vast majority of modern programming languages possess tools to read and write CSV format files.

The near universal interoperability of the CSV format ensures ease of data exchange and validity of collated data sets.

It is anticipated that this project will require several iterations of development, with each iteration examining a different treatment intervention. The results of each iteration will inform the exact nature of data collection in the subsequent iteration. The scope of this section is limited to the scope defined in the preceding section, 11.1

Collated data will then be subject to the defined anonymisation process. Once the anonymisation process is complete, data will be made available to academic research repository.

### ***11.4 Data Quality and Standards***

Data retrieved from the EHR has been validated by a qualified clinician at the point of retrieval and will be considered representative of patient condition and treatment.

Data retrieved from medical equipment, e.g. ventilators, is not validated. Different data domains exhibit different qualities in this regard. Lab value measurements are made on equipment that is calibrated on a basis appropriate for high quality and safe care. Data from sources such as ventilators are not standardised in that the equipment models are standardised across GSTFT.

### ***11.5 Data handling and record keeping***

The collaborative nature of this study requires that data be managed across different environments:

- GSTFT
- Data Repository
- Academic partners

Please note the data pathway defined in section 7.2 of this document.

Each environment differs in available data management facilities and practices and compliance that must be observed. Each environment will be managed separately.

#### **11.5.1 GSTFT**

Data available at GSTFT can be of a personal and confidential nature, and is part of a research project where the highest risk of breaching patient confidentiality exists.



- Confidentiality & Data Protection Policy (Version 2)
- Information Security Policy (Version 4)
- Information security (Principle f)

### **11.5.2 Data Repository**

This project will set up a central data repository for data storage and distribution.

Clinical investigators at GSTFT will upload data sets to the data repository. The project will ensure the local and data repository indexes are maintained and updated.

GSTFT operates a SharePoint site, to which read only access can be given to clinical investigators at GSTFT.

The duplication between data sets stored within IT systems at GSTFT and the data repository will act as an extra layer of redundant storage and mutual back up.

Data storage, management and security practices for the data repository will be defined in the Data Repository Operating Procedures (SOP):

- PSTOTICU\_SOP\_003 - Standard Operating Procedure for Data Repository Administrator

### **11.5.3 Academic Partners**

The University of Warwick (UWAR) and the University of Nottingham (UON) are the academic partners of this study and are the end recipients of compiled data sets.

Data storage and management practices for study participants at UWAR and UON will be defined in the Standard Operating Procedure (SOP):

- PSTOTICU\_SOP\_002 - Data Storage and Management by Academic Partners

## **11.6 Metadata standards and data documentation**

A log of study subjects linked to unique study data set identifiers and investigators at GSTFT. This record will be securely stored within the site file system within GSTFT. This log will not be shared to the central repository or a blank log is provided by study document:

- PSTOTICU\_SD\_004 – Subject Log

A full index of the collated subject data set files will be maintained by site. A local GSTFT copy will be kept, and duplicated within the central data repository. The index is included in **Error! Reference source not found.**, and a blank document:

- PSTOTICU\_SD\_005 – Data Set Index

### 11.6.2 Metadata Descriptive Files

To provide context for the CSV files of a data set, text files will be generated describing the intervention captured. These descriptions will be entirely linked to the intervention and provide no data which might provide an increased risk to subjects. File format will be of a txt format, and thus universally readable. The textual descriptions will be apparent on examination of the quantitative data to provide a more accessible context for the data.

Example descriptions:

- 'PEEP increased from 5cmH2O to 10cmH2O'
- '250ml Crystalloid Bolus given at 02:00' \*
- 'Prone positioning started 02:00. Prone positioning discontinued'

\* - Dates and times are subject to anonymisation as per PSTOTICU\_SOP Data De-Identification Procedure (SOP) Data De-Identification

## 11.7 Data preservation strategy and standards

### **11.7.2 Data Repository**

The data repository will be maintained for a five year period after data collection is completed. After this period, access will be revoked for those users with no further role in the study. Data will be retained as part of the study archival process at GSTFT and duplication will no longer be required.

### **11.7.3 Academic Collaborators**

Upon completion of data collection activities, it is anticipated that academic collaborators will have an extended period. Data will be available to project collaborators for five years after data collection activities are completed.

## ***11.8 Data Security and Confidentiality of Potentially Sensitive Information***

The level of data security is dependent upon the nature of the data – personal data will require higher levels of security.

Data collected for this study will be subjected to an anonymisation process to reduce to a negligible risk of subject re-identification.

Despite the low risk of subject identification, health related data is sensitive. It is appropriate to ensure study data is handled in a way that ensures security and protection against unlawful or unauthorised processing, access, loss, destruction or disclosure.

### **11.8.1 Data Anonymisation**

The aim of anonymisation, is to ensure that the risk of potential re-identification is minimised to negligible levels.

Clinical data will be collected by investigators at GSTFT and subject to an anonymisation process, before being made available to academic partners. The anonymisation process employed is defined in the SOP:

- PSTOTICU\_SOP\_004 – Standard Operating Procedure (SOP) for Data Collection and De-identification

### **11.8.2.1 GSTFT**

Data will be stored on GSTFT network drives within a project folder at S:\IntensiveCare\Research folder. This ensures that access to data is controlled by the Intensive Care clinical research team at GSTFT. Access is authorised through the authorisation system administered by GSTFT IT services and therefore to ensure the security of clinical services and data. This method of project specific storage is consistent across the research portfolio of this research team.

Usage of GSTFT network drives also ensures the durability and availability of data to the same standard as that provided for clinical services and data.

GSTFT data management practice and information governance policy will be implemented and is resident within the GSTFT environment.

### **11.8.2.2 Data Repository**

Access to the data repository requires a username and password, and an additional one time pass code is required. One time pass codes are generated by an authentication app set up at the time of account registration.

Authorisation and permissions are described in the Standard Operating Procedure.

- PSTOTICU\_SOP\_003 - Standard Operating Procedure for Data Repository Administrator

The clinical investigator team at GSTFT will have permissions to upload and delete data. Collaborators will be granted read only permissions to the repository, and can only download data sets and indexes only.

Auditing of all file operations and specifically downloads will be implemented. A full audit trail of repository activity is available. This will be collated, reviewed and reported on a monthly basis.

Data sets stored in the repository will be duplicates of data stored within the clinical services and data.

Data will only be processed on specific computing devices maintained with the project. Storage on project computers will be encrypted with BitLocker software and data will be transmitted over an encrypted transport layer. Access to project computers will be limited to project staff.

Potential loss of data from the computing environments of study collaborators will be minimized by duplication of data within the repository and GSTFT environments.

### ***11.9 Data Collection and Processing Responsibilities***

The chief investigator for this project has ultimate responsibility for ensuring that data is properly administered in accordance with the processes outlined by this SOP.

The chief investigator may delegate responsibility for management of the data to other clinical investigators. A record of delegated responsibilities will be maintained in the study document:

- PSTOTICU\_SD\_008 – Staff Delegation Log

## 12 Statistical considerations

### 12.1 *Sample size calculation*

The subjects will comprise a convenience sample only. This study is not sized for size or incidence, and so a formal sample size calculation to ensure power is inappropriate. Power calculation is not meaningful in this context, since the use of prospective matching technology rather than statistical comparison.

As described in section 8, the admission of one patient may contribute to a large catalogue of data sets within the repository.

The iterative nature of model development was described previously in section 7, such that there will be a large but variable number of episodes, fitting the requirements for iterative development, available from the electronic health record.

Although the precise number of iterations or available episodes cannot be predicted, we anticipate 20 development iterations for which a maximum of 250 approaches per iteration. We thus put an upper bound of collected episodes at 5,000 episodes.

### 12.2 *Statistical analysis*

The core models have been designed to represent a dynamic in vivo cardiac model.

The pulmonary model is comprised of conducting airways and a respiratory system with alveolar compartments, with each compartment having a corresponding stiffness, threshold opening pressures and extrinsic pressures as well as vascular resistances. This allows for a wide spectrum of ventilation patterns to be replicated. The model includes inherent physiological reflex mechanisms such as vasoconstriction.

The cardiovascular model consists of 19 vascular compartments, each with a momentary pressure and volume and its non-linear compliance. Ventricular output is using a time-dependent elastance, which implements pulsatile blood flow. This flow is variably transmitted to all intrathoracic compartments, including the

## **13 Ethical considerations**

HRA approval will be sought for this study. Local R&D approval will also be sought.

## **14 Financing and Insurance**

Finance for this project is provided through an EPSRC grant, reference number EP/S010111/1.

Insurance is provided by the sponsor, Guys & St Thomas' NHS Foundation Trust.

## **15 Reporting and dissemination**

Progress of the study will be disseminated and discussed at collaborator meetings.

The findings will be presented at national and international conferences and published in peer-reviewed journals.

## References

- Chikhani, M., Das, A., Haque, M., Wang, W., Bates, D., & Hardman, J. (2019). A model of acute respiratory distress syndrome: quantitative evaluation of lung recruitment, oxygenation and decreased oxygen delivery. *British Journal of Anaesthesia*, 123, 103–113.
- Das, A., Cole, O., Chikhani, M., Wang, W., Ali, T., Haque, M., . . . Hardman, J. (2018). Lung recruitment maneuvers in acute respiratory distress. *Critical Care Medicine*, 46, 1–10.
- Das, A., Haque, M., Chikhani, M., Cole, O., Wang, W., Hardman, J., & Bates, D. (2019). The effects of lung recruitment maneuvers in acute respiratory distress. *Respiratory Medicine*, 17, 34.
- Hardman, J., Bedforth, N., Ahmed, A., Mahajan, R., & Aitkenhead, A. (1998). Validation of its respiratory components and its ability to predict changes in mechanical ventilation. *British Journal of Anaesthesia*, 81, 103–113.
- Laviola, M., Das, A., Chikhani, M., Bates, D., & Hardman, J. (2017). Investigation of lung oscillations and deadspace gas mixing during apnea using computational modeling. *IEEE Eng Med Biol Soc*, 337–340.
- Laviola, M., Das, A., Chikhani, M., Bates, D., & Hardman, J. (2019). Computational modeling of the mechanisms of carbon dioxide clearance. *British Journal of Anaesthesia*, 123, 103–113.
- McCahon, R., Columb, M., Mahajan, R., & Hardman, J. (2008). Validation of a computational model of acute respiratory distress syndrome: the indices of oxygenation at constant lung-state. *British Journal of Anaesthesia*, 101, 365.
- McNamara, M., & Hardman, J. (2005). Hypoxaemia during open-airway ventilation: a modelling analysis. *Anaesthesia*, 60, 741–746.