

1. Protocol details

1.1 **PROTOCOL TITLE:**

Personalised simulation technologies for optimising treatment in the intensive care unit

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

Sponsor

Guy's and St Thomas' NHS Foundation Trust

Elizabeth Bruna

16th floor Tower Wing

Guy's Hospital

Great Maze Pond

London

SE1 9RT

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

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 protocol and agree to perform the investigations and to abide by this protocol.

The investigator agrees to conduct the trial in compliance with the approved protocol, the National Research Ethics Service (NRES) Code of Practice, the Data Protection Act (1998), the Trust Information Governance Policy (online), the Research Governance Framework (2005' 2nd Edition; as amended), the relevant legislation and any other applicable 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 Personal 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 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 and validate simulation models, by the fitting of those models to the data of critical care patients.
Primary objective	To further develop our simulation models and to evaluate of recently proposed therapeutic interventions.
Secondary objective (s)	To integrate data-streams and clinical data into our existing modelling technology to predict the outcome of individual patients and treatment options for imminent problems. To develop mathematical models that can be used to explore the "design space" of potential interventions for an individual patient in simulation, and to suggest which interventions can be suggested for evaluation.
Number of Episodes	5,000
Study Type	Observational
Endpoints	Not Applicable
Main Inclusion Criteria	This project will include critically ill patients with respiratory failure.

5 Introduction

This project aims to further develop and refine in silico physiological models to the physiological and treatment data of critical care patients. development is fitting the model to, and its description of, the physiological specific interventions. An example would be how a patients physiological increase in positive end expiratory pressure (PEEP). Data for this example 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 are described in great detail, providing extensive, high quality data streams for model development and patient-matching. Models based on this data can incorporate very complex physiological processes and be validated against responses of individual patients, for use as investigation tools.

In contrast to trials on animal models and humans, in silico models of disease pathology are completely configurable, reproducible and reusable. Different combinations of treatments, can be applied to the same patient or subsequently to different 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, or the development of novel physical treatment strategies could significantly reduce pulmonary impairment, and mortality rates, particularly if combined with novel pharmacological treatments.

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 Exeter. The project aims to further develop and refine in silico physiological models, by the fitting of these models to physiological and treatment data of critical care patients.

The project's primary objective is to further develop our simulation platform to enable real time simulation of individual patients and treatments. Sufficient physiological models would enable systematic and efficient exploration of the possible interventions 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).

Integration of available data within the ICU with our existing modelling tools will enable the real time simulation of individual patients and treatments. Sufficient physiological models would enable systematic and efficient exploration of the possible interventions 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 creation of curated data sets representative of the therapeutic interventions of interest. A centralised data repository will be created to catalogue and share collated data sets between centres. The data repository will enforce the oversight and governance ensuring the requirements of data protection against unlawful or unauthorised processing, access, loss, damage or distortion.

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 would be collected from a single patient, and then used by the clinical team, for a clinically appropriate time period. This would compare the model's response to the patient's response.

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

No attempt will be made to capture data from entire patient cohorts or to compare different cohorts.

There is no intention to implement any study specific interventions or treatments. The data will be retrospective, pragmatic and limited to that data recorded throughout the clinical care of the patient, and will not be provided to the clinical team, but will be routinely provided to all patients.

Within this project, there is no intention for physiological models developed to be used in clinical practice. 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 physiological and treatment data for one 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 receiving mechanical ventilation.

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

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

8.1 Inclusion Criteria

- At least 18 years of age
- Patients admitted to GSTFT Intensive Care between the 01/03/2019 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-interaction nature of the data collection.

Data collection will be limited to that data recorded as part of routine care between 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 electrocardiogram (ECG) in the routine care provided to patients. It is anticipated that this study will pose less 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 often

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 by the ventilators.

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

Blood gas analysis results are collected routinely as part of the care of ICU patients, and recorded on a regular basis into 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. It is important to note that the following list is not exhaustive, and it is realistic that only subsets of the entire list of potential data fields will be present in individual data sets, with availability dependent on differing patient characteristics and clinical context.

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
ABP_DIA_ICIP_MMHG	Invasive diastolic blood pressure
ABP_MEAN_ICIP_MMHG	Invasive mean blood pressure
ART_SYS_ICIP_MMHG	Invasive systolic blood pressure
ART_DIA_ICIP_MMHG	Invasive diastolic blood pressure
ART_MEAN_ICIP_MMHG	Invasive mean blood pressure
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
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
VTCO2_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 Inspiration
EXP_TERM_SETTING_PIF_PC	Expiratory term % Peak Expiration

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	Methehaemoglobin
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

SWEET_FO2_ECMO_ICIP_F	Sweep Gas Fractional O2 concentration
SWEET_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 be collated into a single data file.

Data will be collated to comma separated value (CSV) format. CSV is a standard file format that is widely used and easily accessible by all major data processing software platforms e.g. spreadsheets, databases and statistical software. The vast majority of modern programming languages possess the ability to read and write CSV format files.

The near universal interoperability of the CSV format ensures ease of data exchange and the 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 data collection will be 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, 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 time of collection and can be considered representative of patient condition and treatment.

Data retrieved from medical equipment, e.g. ventilators, is not validated in the same way as EHR data. Different data domains exhibit different qualities in this regard. Lab values and measurements are made on equipment that is calibrated on a basis appropriate for the equipment. The data is used to ensure high quality and safe care. Data from sources such as ventilators or other medical equipment is 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 data protection 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 will be managed in a 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. A process will be in place for ensuring 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 be used 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 Management and Security Operating Procedures (SOP):

- PSTOTICU_SOP_003 - Standard Operating Procedure for Data Management and Security by Data Administrator

11.5.3 Academic Partners

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

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

- PSTOTICU_SOP_002 - Data Storage and Management by Academic Partner

11.6 Metadata standards and data documentation

A log of study subjects linked to unique study data set identifiers investigators at GSTFT. This record will be securely stored within the site form within GSTFT. This log will not be shared to the central repository or 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. An 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 and will 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 disc...

* - Dates and times are subject to anonymisation as per PSTOTICU_SO 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 activities are completed. After this period, access will be revoked for those users with no longer required. Data will be retained as part of the study archival process at GSTFT and duplicates no longer required.

11.7.3 Academic Collaborators

Upon completion of data collection activities, it is anticipated that academic collaborators will be available to project collaborators for five 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 Identifiable 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 minimise the risk of a negligible risk of subject re-identification.

Despite the low risk of subject identification, health related data is sensitive and 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 damage.

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 data will undergo an anonymisation process, before being made available to academic partners. The data anonymisation process employed is defined in the SOP:

- PSTOTICU_SOP_004 – Standard Operating Procedure (SOP) for Data Anonymisation

11.8.2.1 GSTFT

Data will be stored on GSTFT network drives within a project folder, such as S:\\IntensiveCare\\Research folder. This ensures that access to data is controlled by the Intensive Care clinical research team at GSTFT. Access is authorised via a centralised authorisation system administered by GSTFT IT services and therefore to the data is controlled by the security of clinical services and data. This method of project specification and access 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 consistent with clinical services and data, as the data is resident within the GSTFT environment.

11.8.2.2 Data Repository

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

Authorisation and permissions are described in the Standard Operating Procedure (SOP) document:

- PSTOTICU_SOP_003 - Standard Operating Procedure for the Data Repository Administrator

The clinical investigator team at GSTFT will have permissions to upload and download data sets. External collaborators will be granted read only permissions to the repository, allowing them to download data sets and indexes only.

Auditing of all file operations and specifically downloads will be implemented. A detailed 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 system.

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

Potential loss of data from the computing environments of study collaborators will be mitigated by ensuring data is backed up regularly, limiting the amount of data stored on any one device, and preventing 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 data is collected and processed in accordance with the processes outlined by this SOP.

The chief investigator may delegate responsibility for management of the data to 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 size or incidence, and so a formal sample size calculation to ensure inappropriate. Power calculation is not meaningful in this context, since prospective matching technology rather than statistical comparison.

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

The iterative nature of model development was described previously in section 2.2.2. In this section, we will describe the iterative nature of the model development process. The iterative nature of the model development process is that there will be a large but variable number of episodes, fitting the required number of episodes, available from the electronic health record.

Although the precise number of iterations or available episodes cannot anticipate 20 development iterations for which a maximum of 250 approaches. 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* card

The pulmonary model is comprised of conducting airways and a series of alveolar compartments, with each compartment having a corresponding set of parameters for 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 momentary pressure and volume and its non-linear compliance. Ventricular contraction is simulated using a time-dependent elastance, which implements pulsatile blood flow. The pressure signal is variably transmitted to all intrathoracic compartments, including the lungs.

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/N008642/1.

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 academic journals.

References

Chikhani, M., Das, A., Haque, M., Wang, W., Bates, D., & Hardman, J. (1998). The effects of lung recruitment maneuvers in acute respiratory distress syndrome: quantutative evaluation by oxygenation and decreased oxygen delivery. *British Journal of Anaesthesia*, 81, 342-347.

Das, A., Cole, O., Chikhani, M., Wang, W., Ali, T., Haque, M., . . . Hardman, J. (1998). The effects of lung recruitment maneuvers in acute respiratory distress. *Critical Care Medicine*, 26, 171-176.

Das, A., Haque, M., Chikhani, M., Cole, O., Wang, W., Hardman, J., & Bates, D. (2000). The effects of lung recruitment maneuvers in acute respiratory distress syndrome. *Respiratory Medicine*, 94, 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, 342-347.

Laviola, M., Das, A., Chikhani, M., Bates, D., & Hardman, J. (2017). Investigation of oscillations and deadspace gas mixing during apnea using computational models. *IEEE Eng Med Biol Soc*, 337-340.

Laviola, M., Das, A., Chikhani, M., Bates, D., & Hardman, J. (2019). Computational mechanisms of carbon dioxide clearance. *British Journal of Anaesthesia*, 122, 10-16.

McMahon, R., Columb, M., Mahajan, R., & Hardman, J. (2008). Validation of a fidelity, computational model of acute respiratory distress syndrome using the indices of oxygenation at constant lung-state. *British Journal of Anaesthesia*, 101, 365.

McNamara, M., & Hardman, J. (2005). Hypoxaemia during open-airway modelling analysis. *Anaesthesia*, 60, 741-746.