1 Full study title

Process mapping and data collection to inform a computer simulation model of hospitalised patients with bloodstream infection, sepsis and systemic infection

2 Short study title and acronym

Computer Simulation Model of patients with Bloodstream infection, Sepsis and systemic Infection (CSM-BSI)

Also known as Hospital-in-a-Box

3 Protocol version and date

1.0

Final version

2023-10-20

4 Research reference numbers

IRAS Number: 334396

SPONSORS Number: LHS0211

FUNDERS Number: 226691/Z/22/Z

5 Signature page

The undersigned confirm that the following protocol has been agreed and accepted and that the Chief Investigator agrees to conduct the study in compliance with the approved protocol and will adhere to the principles outlined in the Declaration of Helsinki, the Sponsor's SOPs, and other regulatory requirement.

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

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

For and on behalf of the Study Sponsor:

Signature: Date: ...../...../.....

.....

Name (please print):

.....

Position:

.....

Chief Investigator:

Date: 12/07/2023

Signature:

Name: (please print):

ALESSANDRO GERADA

6 Table of contents

7 List of abbreviations

A&E	Accident and emergency
ABM	Agent-based model
AMR	Antimicrobial resistance
BSI	Bloodstream infection
CI	Chief Investigator
CIPHA	Combined Intelligence for Population Health Action
DAAG	Data Asset and Access Group (CIPHA)
DES	Discrete event simulation
GDP	Gross domestic product
HES	Hospital episode statistics
HIV	Human immunodeficiency virus
ICU	Intensive care unit
IT	Information Technology
LCL	Liverpool Clinical Laboratories
LUFT	Liverpool University Hospitals Foundation Trust
NHS	National Health Service
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SD	System dynamics
SMG	Study Management Group
STRESS-DES	Strengthening the Reporting of Empirical Simulation Studies
TRE	Trusted Research Environment
UK	United Kingdom
UoL	University of Liverpool
USD	United States Dollar

8 Key study contacts

Chief Investigator	Alessandro Gerada Antimicrobial Pharmacodynamics and Therapeutics Group Pharmacology Department Institute of Systems, Molecular, and Integrative Biology University of Liverpool Ground Floor William Henry Duncan (WHD) Building 6 West Derby Street L7 8TX United Kingdom +44 7469 910 748 alessandro.gerada@liverpool.ac.uk
Sponsor	Liverpool University Hospitals NHS Foundation Trust is the Sponsor for this Study. It is recognised that as an employee of the Trust the Chief Investigator has been delegated specific duties, as detailed in the Sponsorship Approval letter.
Funder(s)	Wellcome Trust Gibbs Building 215 Euston Road London NW1 2BE +44 20 7611 8888
Key Protocol Contributors	Dr Alessandro Gerada Prof William Hope Dr Alexander Howard

9 Study summary

Study Title	Computer Simulation Model of patients with Bloodstream infection, Sepsis and systemic Infection
Internal ref. no. (or short title)	CSM-BSI / Hospital-in-a-Box
Research Question/Aim(s)	The aim of this project is to design, construct, parameterise and calibrate a computer simulation model of patients with presumed bloodstream infection/sepsis/systemic infection (the study population) to ensure that the simulation of patients with bloodstream infection is realistic and correct.
Study Design	Divided into three sub-studies: 1) Qualitative study (structured interviews and focus groups) – staff participants 2) Analysis of retrospective routinely collected data 3) Observational study (direct observation without intervention) of the management of patients with suspected bloodstream infection
Study Participants	1) Staff members with specific roles of interest 2) Patients being investigated with blood cultures 3) Patients with suspected bloodstream infection and staff caring for them (indirectly)
Planned Size of Sample	1) 10 staff participants 2) All patients associated with blood culture investigation request over a retrospective 5-year study period 3) All patients and staff within the ward environment during the observational study days
Planned Study Period	1) Structured interviews and focus group sessions to be organised over a 6-month period 2) Retrospective 5-year study period 3) Thirty days of observation (not necessarily in sequence)

10 Funding and support in kind

Funder(s)	Financial and non-financial support given
Wellcome Trust	Funding of staff and expenses

11 Role of study sponsor and funder

Study Sponsor (Liverpool University Hospitals NHS Trust):

The sponsor will assume overall responsibility for proportionate, effective arrangements being in place to set up, run and report the project. They will have overall responsibility for the research, including:

1. Identifying and addressing problems with the proposal, protocol and applications and ensuring that they take into account systematic reviews of relevant existing research evidence and other relevant research in progress, make appropriate use of patient, service user and public involvement, and are scientifically sound (e.g., through independent expert review), safe, ethical, legal and feasible and remain so for the duration of the research, taking account of developments while the research is ongoing
2. Satisfying itself that the investigators, research team and research sites are suitable
3. Ensuring that roles and responsibilities of the parties involved in the research and any delegation by the sponsor of its tasks are agreed and documented
4. Ensuring adequate provision is made for insurance or indemnity to cover liabilities which may arise in relation to the design, management and conduct of the research project
5. Ensuring appropriate arrangements are made for making information about the research publicly available before it starts (unless a deferral is agreed by or on behalf of the research ethics committee)
6. Agreeing appropriate arrangements for making data and tissue accessible, with adequate consent and privacy safeguards, in a timely manner after it has finished
7. Ensuring arrangements for information about the findings of the research to be made available, including, where appropriate, to participants)
8. Ensuring that, where expected or required, the research has approval from a research ethics committee (Whether outright or following a provisional opinion, re-submission, or appeal) and any other relevant approval bodies before it begins
9. Verifying that regulatory and practical arrangements are in place, before permitting the research to begin in a safe and timely manner

10. Putting and keeping in place arrangements for adequate finance and management of the research project, including its competent risk management and data management
11. Ensuring that effective procedures and arrangements are kept in place and adhered to for reporting (e.g., progress reports, safety reports) and for monitoring the research, including its conduct and the ongoing suitability of the approved proposal or protocol considering adverse events or other developments.

Study funder (Wellcome Trust):

The role of the Wellcome Trust as study funder is limited to provision of funds to conduct the research. It has no role in the design or implementation of the study.

12 Protocol contributors

Dr Alessandro Gerada – study conceptualisation, study design, protocol writing, Chief Investigator

Prof William Hope – review and editing

Dr Alexander Howard – subject material from ADAPT-AST protocol where overlap is present, with permission

13 Keywords

Bloodstream infection, antimicrobial resistance, computer simulation, computer modelling, sepsis.

14 Study Gantt chart

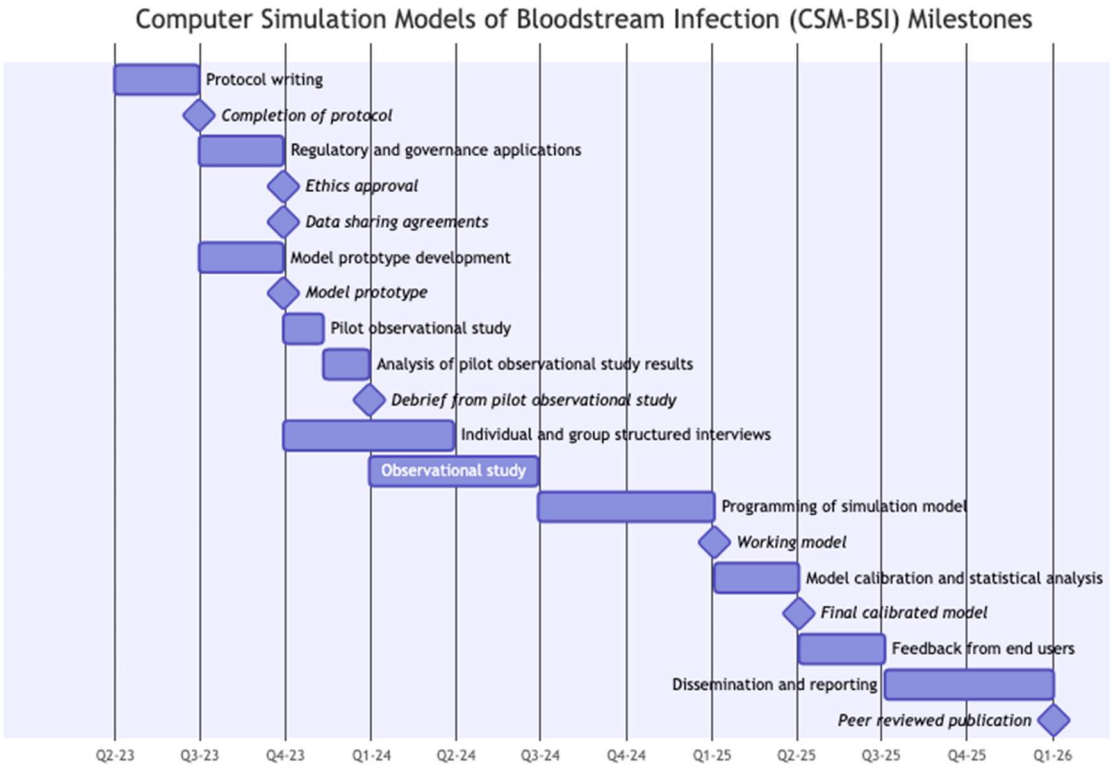


Figure 1 - Study Gantt chart of timelines and milestones (latter represented by diamond shape)

15 Study protocol

15.1 Background

15.1.1 Antimicrobial resistance – a global threat

Antimicrobial resistance (AMR) is an established and critical threat to modern healthcare. Various UK and international reports have highlighted the threat to human health posed by AMR (Davies et al., 2013; WHO, 2016). Recent estimates of the magnitude of AMR burden in the European region have reported 541,000 deaths associated with bacterial AMR in 2019, of which the largest burden arose from bloodstream infections – 195,000 deaths (European Antimicrobial Resistance Collaborators, 2022). AMR also has a staggering economic cost – even maintenance of the status quo (i.e., if current AMR rates to remain constant) the cumulative world GDP cost would be 5.8 trillion USD (Taylor et al., 2014). This compares to the combined GDP of Germany and the United Kingdom. Although the discovery of novel antimicrobials is important, antimicrobial optimisation is crucial for the control of AMR, and has been identified as one of the key themes in global and national AMR guidance documents (Mitchell et al., 2022). This protocol addresses this theme, but also contributes to the optimisation of AMR surveillance, data management and dissemination, by expanding the use of surveillance data for computer simulation modelling.

The core method by which antimicrobial stewardship aims to help institutions and individuals tackle the problem of AMR is through an overall reduction in antimicrobial usage.

However, this measure is the antithesis to the general principles of the management of sepsis, which mandate administration of broad-spectrum antimicrobial therapy within 1 hour of initial assessment. Sepsis is a major worldwide patient safety issue, with an estimated 66,096 deaths per year in the United Kingdom. Frontline clinicians and institutional policy makers must strike a delicate balance between appropriate treatment of patients with sepsis and the control of broad-spectrum antimicrobial usage (Academy of Medical Royal Colleges, 2022).

Further complicating antimicrobial decision making is the following three clinical phenotypes can be encountered:

1) Sepsis,
2) Systemic infection,
3) Bacteraemia.

The phenotypes have varying levels of organ dysfunction and diagnostic microbiology results, with significant overlap that adds to diagnostic uncertainty and complicates antimicrobial decision making.

15.1.2 Antimicrobial use in complex systems

At a system level, there are many interventions that can be considered to optimise antimicrobial usage:

1) Empirical antimicrobial treatment recommendations (e.g., antimicrobial formulary recommendations)
2) Sensitivity and specificity of sepsis diagnosis (e.g., diagnostic cut-offs of early warning systems)
3) Prescriber behaviour through teaching and training
4) Specialist teams (e.g., sepsis and antimicrobial stewardship teams)
5) Changes in diagnostic methodology (e.g., clinical use of a novel biomarker)
6) Clinician decision support tools

Although it is likely that the above actions would lead to “better” prescribing in a particular setting, the decision of which, if any, to implement can be challenging. Interventions must be weighed against the risk of under-treating serious infections and sepsis. Additionally, although resources are finite – investment costs must be considered within the context of patient outcomes and a broad endpoint of total system resources expended, rather than within a cost silo.

Furthermore, healthcare systems are complex systems and prediction of the downstream impact of the above decisions can be difficult. Such decisions are very often made based on retrospective aggregate data analysis and generally their impact is predicted in isolation. This drives many of the challenges in optimisation of antimicrobial usage in a healthcare environment (Ratnapalan and Lang, 2020). Table 1 lists features that are typical of complex systems alongside corresponding examples in AMR.

Table 1 - AMR-specific examples of complex system features

Complex system features	**Examples in AMR**
Nonlinearity	The effects of antimicrobial resistance can have a compounding effect - resistant isolates often trigger laboratory subprocesses such as second-line antimicrobial testing. Aggregate data can mask the delays in reporting and administration of optimal antimicrobials in such cases.
Interaction with rest of environment	Availability of staff to perform interventions depends on workload allocation that includes demand from other clinical and non-clinical requirements
Uncertainty	Decisions must be made despite some critical states being hidden from agents, e.g., the infecting pathogen is not known early in the disease course
Inherent pattern	Most empirical antibiotic choices follow an institutional-level formulary
Unpredictability	Prescribers can deviate from formulary recommendations

Complex systems are difficult to understand through analysis of individual components in isolation. Most decisions are currently made based on retrospective data without considering future population and AMR changes. Even when published data are available, inferring the results onto the healthcare system of interest can be misleading, since key baseline metrics (such as population AMR rates and prescribing practices) are almost always different. Novel approaches are required to make better decisions on which interventions to implement by considering their impact on the healthcare system as a whole, and on a representative patient population and healthcare system.

15.1.3 The role of computer simulation in AMR

Computer simulation involves the *in silico* reconstruction of a virtual system modelled after the real-life complex system of interest. The virtual system is then dynamically analysed and mathematical models used to gain insights into the behaviour of the real system. Computer simulation is a powerful tool and approach that can be used to manage the challenges described in the previous section. It has been extensively used in management and policy to predict events and behaviour, in complex situations, e.g., in traffic management, defence, manufacturing, design of supply chains and engineering. However, such simulation is relatively underused in healthcare.

Simulation has been applied to the following healthcare problems (Mielczarek and Uziałko-Mydlikowska, 2012), which have similarities to the challenges encountered in the precise use of antimicrobials and AMR:

Table 2 - Literature healthcare simulation examples and potential applications in AMR

Application area	**Literature examples**	**AMR applications**
Health policy evaluation	Policy simulation to predict drug expenditure (Dormuth et al., 2005)	Predicted individual and system risk of AMR development
Intervention and treatment programs	Simulation of effects of HIV intervention programs (Rauner, 2001)	Effects of AMR rapid diagnostics
Spread of infectious diseases	Use of agent-based modelling to inform interventions for control of SARS-CoV-2 (Kerr et al., 2021)	Modelling of horizontal transmission of AMR colonisation
Estimating the effect of organisational changes	Simulation of ICU bed availability (Cahill and Render, 1999)	Simulating the effect of dedicated sepsis and AMR teams Effects of changes in antimicrobial formularies
System diagnosis	Accident and emergency department performance (Gunal and Pidd, 2006)	Understanding causes of inappropriate or delayed antimicrobial prescriptions
Staff scheduling	Nursing staffing levels (Saville et al., 2021)	Optimisation of nursing, clinician and pharmacy staffing levels to improve expedite time to antimicrobial delivery
Resource optimisation	Optimisation of patient beds, physicians and nurses in a hospital simulation (Ordu et al., 2021)	Optimisation of laboratory resources to produce AMR diagnostics such as blood cultures
Medical decision making	System dynamics model of infection risk, expectations and perceptions on antibiotic prescribing (Kianmehr et al., 2020)	Simulation of the downstream AMR impacts of changes in prescriber behaviour patterns secondary to education programmes
Extreme events planning	Response of rural acute health-care system to a bioterrorism attack (Miller et al., 2004)	Simulation of emergence of novel drug resistance Simulation of impact of antibiotic shortages

Computer simulation can reveal profound insights by allowing the components of the system to virtually interact while observing outcomes. Although deterministic models of such systems are possible (i.e., components are allowed to interact with each other in predictable ways), stochastic simulation models (i.e., interactions are probabilistic) can be more appealing due to the ability to incorporate randomness and direct simulation of agent behaviour. Models are constructed from the ground up, by defining the behaviour of individuals when presented with certain situations (e.g., a clinician reviewing a patient suspected of having infection). Repeated stochastic iterations of the simulation are analysed to produce aggregate observations, presented in the form of probability distributions and proportions (Monte Carlo simulation).

The core antimicrobial management of patients with the sepsis/bloodstream infection/systemic infection paradigm can be simplified into the following decision nodes:

- Initial therapy with empirical antibiotics
- Collection and processing of diagnostic microbiology tests
- Refinement of antimicrobial therapy using clinical progress and diagnostic microbiology results

We propose that computer simulation modelling can be an invaluable research and strategic tool to guide the precise use of antimicrobials and facilitate resource allocation, and can be achieved by incrementally developing application at these three core nodes.

There are three main paradigms of stochastic simulation (Borshchev and Filippov, 2004):

1) Agent-based modelling
2) Discrete event simulation
3) System dynamics

Each paradigm is suited to a particular level of model abstraction. System dynamics (SD) models are relatively high-level simulations that look at general dynamic interactions such as flow of resources between parts of the system. In the context of AMR simulations, SD models could be useful for studies interested in looking at high-level interactions, such as between antimicrobial supply chains and hospital usage at a regional or national level. Although agent-based models (ABM) are highly useful for modelling communicable diseases, as exemplified by SARS-CoV-2, the process-driven and patient-centred nature of sepsis/bloodstream infection/severe infection management lends itself well to discrete event simulation (DES). DES has been applied to hospital simulation problems with similar challenges, using data from Hospital Episode Statistics (HES) and Patient Administration Systems (PAS) to create a whole hospital simulation (Günal and Pidd, 2011). A brief introduction to the defining features of ABM and DES is provided in Appendix 1 – .

Although the use of stochastic simulation models, including DES, is well established in healthcare, to our knowledge there is no available model that specifically simulates patients with infection that are treated with broad-spectrum antibiotics. The availability of such a model would allow the simulation of various “what if?” scenarios and examine, *in silico*, the effects of institutional changes and resource allocation. The downstream impact of decisions can be first tested virtually, which departs from the traditional approach of making empirical changes based on retrospective evidence, and continuing to prospectively observe the impact via surveillance.

The availability of such a model will be an asset to tackling the AMR emergency and therefore has strong population health potential. Although the model will help improve the patient care and antimicrobial usage in a tertiary care setting, there will be cascading benefits to the wider population. Optimising the use of tertiary care antimicrobials, in particular broad-spectrum agents that are high risk for generation of AMR, will lead to a reduction in AMR prevalence in the regional population.

15.1.3 Building an advanced AMR simulator

For the simulator to be useful in addressing the needs discussed in the earlier sections, it must accurately model the real healthcare system of interest, while allowing sufficient flexibility to allow for deployment in a variety of settings.

The principle aim of this study is to support the development and parameterization of a Discrete Event Simulation model of adult patients with presumed sepsis/bloodstream infection/systemic infection (the study population) to ensure that the simulation of patients with bloodstream infection is realistic and correct.

This observational study will generate and collect the following:

1) Qualitative information about patient pathways (process mapping) including resource requirement (e.g., hospital staff)
2) Simulation parameter estimation using routinely collected patient clinical and administrative data
3) Fine-tuning of important parameters through direct observation

This study will run in parallel to parts of the model software coding, to inform the model structure design.

15.2 Overview of clinical study and rationale

To construct this model, we will harness the expertise of key hospital staff responsible for the management of patients with bloodstream infection, and carefully parameterise the model using routinely collected clinical data and direct observation data where indicated. Simulation models are often criticised for a lack of reproducibility. We intend that the model will be generalisable to other populations and healthcare settings. Therefore, this research will follow the Strengthening the Reporting of Empirical Simulation Studies guideline to enable others to apply this model to wider problems in AMR (Monks et al., 2019). The model report will address all the STRESS-DES checklist items.

By reporting the model using established good practice principles for simulation studies, external users can apply the model to their population of interest, using local parameters and data.

The study design and methods in later sections have been informed by two pilot studies:

- Prototype DES model that simulated patients with bacteraemia. This study was not intended to produce realistic simulations but provide feasibility for a DES BSI from a software development perspective, and that calibration of the model to real-world observed data is possible. Aggregate outcome data (not individual level) was used as a calibration metric. The time to administration of first dose of antibiotic that was active against the infecting pathogen was identified as a key and measurable metric that can be used for model calibration
- Laboratory process mapping for blood culture specimens. This study defined a system for process mapping and flow chart annotation for laboratory specimens. Both routinely collected retrospective data and direct observation were used to inform parameters.

Later sections of this protocol are presented as three discrete sub-studies, as informed by the above pilot studies. Table 3 provides a high-level outline of these sub-studies. Although the studies are designed to run broadly in series, it is expected that some parts may occur in parallel (e.g., direct observations may be used to update process maps and feed back to staff participants).

Table 3 - Outline of sub-studies

	Title	Study format	Core aim/deliverable
Sub-study 1	Qualitative process mapping	Qualitative study Focus groups	Detailed process map of patients with BSI, sepsis, systemic infection
Sub-study 2	Retrieval of retrospective routinely collected clinical data	Retrospective data collection and analysis	Individual-level anonymised dataset for use as model parameters and for calibration
Sub-study 3	Direct observation study	Direct observation of events in treatment of patients with BSI, sepsis, systemic infection	Detailed dataset of important event timings that are not available in Sub-study 2 dataset

15.3 Overall research question/aim

15.3.1 Objectives

The aim of this observational study is to design and calibrate a computer simulation model for hospitalised adults that require investigation and treatment for suspected serious infection.

15.3.2 Outcome

The main outcome of this study is a computer model that is parameterised using real-world data. The model will be customisable and address critically important questions such as:

- Optimisation of healthcare resources
- Impact of system-wide decisions on antimicrobial policy
- Options appraisal for laboratory testing methodology

15.4 Sub-study 1 – Qualitative process mapping

15.4.1 Aim

Generate a detailed process map of a theoretical patient presenting to the healthcare environment for treatment of suspected infection, which will be used to inform the qualitative structure and logic of the simulation model (i.e., pathways, decisions, etc) during the software development process.

15.4.2 Participants

Patient participants are not required for this sub-study.

The participants of this qualitative study are:

- Staff participants from NHS partner organisation:
  - Accident & Emergency or Acute Medicine Physician
  - Patient flow manager
  - Infectious Diseases Physician or Clinical Microbiologist
  - Accident & Emergency or Acute Medicine senior nurse
  - Pharmacist
  - Bacteriology senior biomedical scientist
- Meeting convenors (representatives from the clinical study team)
- Administrator (from clinical study team)

Invitations to participate will be sent out through an all-user trust email, and compensation will be offered for participants' time, within the study sponsor allowances. Attempts will be made to organise sessions at a time and place convenient for the participants. Participation in this research is outside time dedicated for clinical care.

15.4.3 Consent

Although consent for this study is implied by the participant seeking to participate in the study, formal written consent will be required prior to enrolment. The consent process will also clarify details around the study – the participant will be invited to keep a copy for their records. The sub-study consent form is available in 0

Staff consent form.

Participants will be presented with the written consent form during the first meeting/session, and opportunity is given to ask questions. Participants must have capacity to give consent for the study prior to proceeding.

15.4.4 Outcomes and objectives

Comprehensive flow-chart of the management of a patient presenting with, and treated for, suspected infection. The flow-chart will be used to inform the simulation logic during the software development process.

15.4.5 Methods

Participants will be invited to a series of individual and group meetings to gather qualitative data on the management of patients with suspected bloodstream infection, as summarised in Table 4.

Table 4 - Summary of Sub-study 1 (qualitative study) focus group meetings

Meeting/Event name	Attendees	Core aim/outcome	Meeting resources	Duration
One-to-one data generation session	Individual staff participant Meeting facilitator/ convenor	Open-ended questioning on how patients with bloodstream infection are managed, with a focus on the participant's area of expertise	Conversation prompts (see 0 Table 6 - Raw data fields (pre-processing)	90 minutes

Database	Field	Example
Demographic data	NHS number	111 111 1111
Demographic data	Hospital number	RQ6111111 1
Demographic data	Postcode	L1 1AA
Demographic data	Date of birth	01/01/2000
Prescribing data	NHS number	111 111 1111
Prescribing data	Hospital number	RQ6111111 1
Prescribing data	Drug name	amoxicillin
Prescribing data	Location	Liverpool Medical Centre
Prescribing data	Route	oral
Prescribing data	Dose	500mg
Prescribing data	Frequency	every 8 hours
Prescribing data	Start date & time (prescription)	01/01/2022 10:00
Prescribing data	End date & time (prescription)	07/01/2022 08:00

Prescribing data	Start date & time (administration)	01/01/2022 10:00			
Prescribing data	End date & time (administration)	07/01/2022 08:00			
Prescribing data	Drug allergies	penicillin			
Prescribing data	Allergy reaction type	Rash			
Microbiology data	NHS Number	111 111 1111			
Microbiology data	Hospital number	RQ6111111 1			
Microbiology data	Date & Time Collected	01/01/2022 00:00			
Microbiology data	Date & Time Received	02/01/2022 12:00			
Microbiology data	Date & Time Authorised	03/01/2022 20:00			
Microbiology data	Specimen type	blood			
Microbiology data	Report code	YUPO			
Microbiology data	Comment code	IUTI			
Microbiology data	Specimen site	right arm			
Microbiology data	Location	Ward 4			
Microbiology data	Specimen number	c,22.111111 1.A			
Microbiology data	Clinical Details	unstructured free text			
Microbiology data	Epithelial cells	+++			
Microbiology data	White cell count	> 50			

			Microbiology data	Red cell count	> 50	
			Microbiology data	Organism count	> 10^7	
			Microbiology data	Organism code	ESCO	
			Microbiology data	Organism name	E coli	
			Microbiology data	AML	S	
			Microbiology data	AMP	I	
			Microbiology data	OX	R	
			Microbiology data	MET	S	
			Microbiology data	MEL	I	
			Microbiology data	AMC	R	
			Microbiology data	P/T	S	
			Microbiology data	CL	I	
			Microbiology data	CXM	R	
			Microbiology data	CPD	S	
			Microbiology data	CRO	I	
			Microbiology data	FOX	R	
			Microbiology data	CTX	S	
			Microbiology data	CAZ	I	

			Microbiology data	CZA	R
			Microbiology data	CIP	S
			Microbiology data	LEV	I
			Microbiology data	CN	R
			Microbiology data	AK	S
			Microbiology data	TE	I
			Microbiology data	TGC	R
			Microbiology data	MER	S
			Microbiology data	ETP	I
			Microbiology data	TEC	R
			Microbiology data	VA	S
			Microbiology data	DA	I
			Microbiology data	E	R
			Microbiology data	ATM	S
			Microbiology data	LZ	I
			Microbiology data	TSU	R
			Microbiology data	W	S
			Microbiology data	NIT	I

			Microbiology data	FOT	R		
			Microbiology data	MUP	S		
			Microbiology data	DAP	I		
			Microbiology data	RD	R		
			Microbiology data	OB	S		
			Microbiology data	MTZ	I		
			Microbiology data	C	R		
			Microbiology data	FD	S		
			Microbiology data	FLC	I		
			Microbiology data	MCA	R		
			Microbiology data	ANI	S		
			Microbiology data	CAS	I		
			Microbiology data	VOR	R		
			Microbiology data	AMB	S		
			Other pathology data	NHS Number	111 111 1111		
			Other pathology data	Hospital number	RQ6111111 1		
			Other pathology data	PaO2	7.0kPa		
			Other pathology data	PaCO2	7.0kPa		

			Other pathology data	pH	7.0	
			Other pathology data	HCO3	7.0	
			Other pathology data	FiO2	32%	
			Other pathology data	Lactate	1.5	
			Other pathology data	Bilirubin	15	
			Other pathology data	ALT	15	
			Other pathology data	Alkaline phosphatase	15	
			Other pathology data	Urea	15	
			Other pathology data	Creatinine	85	
			Other pathology data	Hb	85	
			Other pathology data	Platelets	130	
			Other pathology data	White cell count	2.4	
			Other pathology data	Neutrophils	2.4	
			Other pathology data	Eosinophils	2.4	
			Other pathology data	Lymphocytes	2.4	
			Other pathology data	Basophils	2.4	
			Other pathology data	Monocytes	2.4	
			Other pathology data	CRP	34	

			Other pathology data	PT	34	
			Other pathology data	APTT	34	
			Other pathology data	Fibrinogen	34	
			Other pathology data	Date & time collected	02/01/2022 12:00	
			Other pathology data	Date & time reported	03/01/2022 20:00	
			Admission data	NHS Number	111 111 1111	
			Admission data	Hospital number	RQ6111111 1	
			Admission data	Admission date & time	01/01/2022 00:00	
			Admission data	Admission location	A&E	
			Admission data	Transfer date & time	02/01/2022 12:00	
			Admission data	Transfer location	ward 4	
			Admission data	Discharge date & time	03/01/2022 20:00	
			Admission data	Discharge status	discharged home	
			Admission data	Discharge location	nursing home	
			Admission data	Vital status	alive, dead	
			Episode data	NHS Number	111 111 1111	
			Episode data	Consultation date & time	01/01/2022 00:00	
			Episode data	Coded diagnosis	UTI	

Episode data	Diagnosis date	01/01/2022 00:00			
Electronic healthcare record data	NHS Number	111 111 1111			
Electronic healthcare record data	Hospital number	RQ6111111 1			
Electronic healthcare record data	Clinical form date & time	01/01/2022 00:00			
Electronic healthcare record data	Clinical notes	unstructured free text			
Electronic healthcare record data	Clinical form date & time	02/01/2022 12:00			
Electronic healthcare record data	Clinical form	semi-structured free text			
Electronic healthcare record data	Observation date & time	03/01/2022 20:00			
Electronic healthcare record data	Heart rate	100			
Electronic healthcare record data	Respiratory rate	20			
Electronic healthcare record data	Systolic blood pressure	120			
Electronic healthcare record data	Diastolic blood pressure	80			
Electronic healthcare record data	Temperature	36.5			

<table>
<tr>
<td></td>
<td></td>
<td></td>
<td>
<table>
<tr><td>Electronic healthcare record data</td><td>AVPU score</td><td>A</td></tr>
<tr><td>Electronic healthcare record data</td><td>Oxygen saturation</td><td>96%</td></tr>
<tr><td>Electronic healthcare record data</td><td>FiO2</td><td>32%</td></tr>
<tr><td>Electronic healthcare record data</td><td>NEWS score</td><td>4</td></tr>
<tr><td>Electronic healthcare record data</td><td>eSepsis pathway date & time</td><td>03/01/2022 20:00</td></tr>
<tr><td>Electronic healthcare record data</td><td>eSepsis pathway</td><td>semi-structured free text</td></tr>
<tr><td>Electronic healthcare record data</td><td>AKI alert</td><td>1</td></tr>
<tr><td>Diagnostic coding data</td><td>NHS Number</td><td>111 111 1111</td></tr>
<tr><td>Diagnostic coding data</td><td>Hospital number</td><td>RQ6111111 1</td></tr>
<tr><td>Diagnostic coding data</td><td>Primary diagnostic code</td><td>Urinary tract infection</td></tr>
<tr><td>Diagnostic coding data</td><td>Secondary diagnostic code</td><td>Sepsis</td></tr>
</table>

Discussion prompts for Sub-study 1 meeting convenors)
</td>
<td></td>
</tr>
<tr>
<td>One-to-one data validation session</td>
<td>Individual staff participant</td>
<td>Feedback of results of qualitative analysi</td>
<td>Draft process flowchart
Conversation prompts (see 0)</td>
<td>90 minutes</td>
</tr>
</table>

	Meeting facilitator/ convenor	s and draft process map		
Focus group mock patient scenarios and validation	All staff participants Meeting facilitator/ convenor	Assessment of the practical functionality of the process map flowchart	Process flowchart Mock patient scenarios	120 minutes
Focus group final feedback session	All staff participants Meeting facilitator/ convenor	Final validation of flowchart that incorporates full feedback from prior sessions	Process flowchart	90 minutes

Two one-on-one meetings and two group meetings will be convened with the above participants. A set of guidelines on flowchart construction will be made available in advance. The following principles will be highlighted:

- Flowchart entry points:
  - Acutely presenting patients - suspicion of infection at patient triage
  - Hospitalised patients - collection of blood cultures for patients not previously suspected of being infected
- Flowchart exit points:
  - Completion of antimicrobial course of therapy (either on clinician decision that infection is unlikely, or completion of prescribed multi-day antimicrobial course)
  - Reporting of negative blood culture when infection suspected but empirical treatment not initiated
- Appropriate use of flowchart notation as described in Appendix 5 – Flowchart Design Principles, in particular:
  - Decision steps
  - Estimated time required for time-dependent events
  - Resource requirements, where applicable (e.g., staff resource, medications and equipment)
  - Re-entrant pathways (or loops) – e.g., as triggered when a patient clinically deteriorates

Flowchart construction will be adapted from a process modelling study of the ABCDE primary survey in trauma resuscitation (Lodemann et al., 2022). However, since our study is intended to be developed primarily as a DES model rather than an ABM one, additional notation is introduced to clarify which staff member is required for a particular step.

It is acknowledged that real patient management has high variability, and some events can be performed differently (by necessity, or by patient or staff choice), leading to the same outcome. The flowchart should capture the core events and decisions that guide the patient journey.

The study will take an iterative approach, whereby data from the sessions/meetings is collected, explored, and converted into a flowchart. At each step, the output of the qualitative analysis will be fed back to the participants and the process repeated (Figure 2).

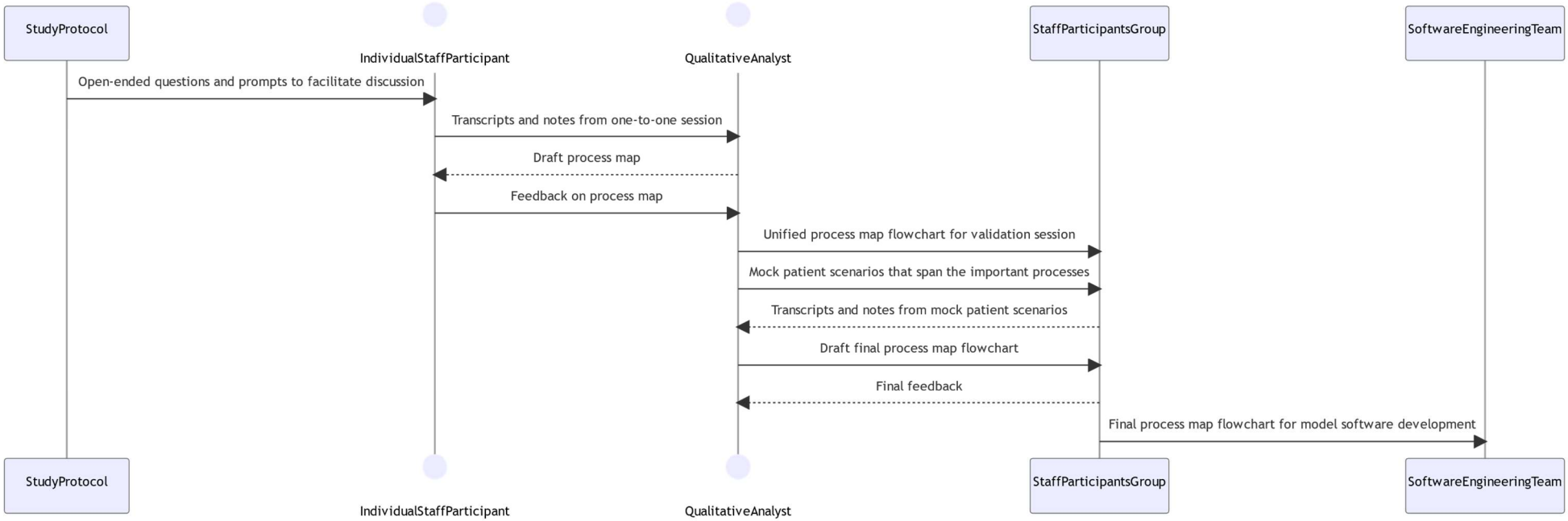


Figure 2 - Iterative process of qualitative data collection and analysis

15.4.6 Data analysis

Meetings and focus group sessions will be recorded (with consent) and transcribed. Data will be analysed using NVivo to pick out key themes and themes that overlap between participants. This qualitative data will be used to construct a process flowchart as described in the previous sections. Flowcharts will be designed based with the intension of being informative and rigorous for model software development. To ensure consistency in notation, a set of design principles have been agreed and are described in Appendix 5 – Flowchart Design Principles. Specialist visualisation software – Microsoft Visio and mermaid.js – will be used to create easy to read flowcharts.

Once the final flowchart is agreed, the study team will proceed to code the simulation logic in an appropriate simulation modelling framework and language. The prototype model was constructed using the SimPy module within the Python programming language ("SimPy," 2020). However, depending on the complexity of the final modal, an alternative language or framework may be more appropriate. This will be discussed and explored within the study group meetings.

15.5 Sub-study 2 – Retrieval of retrospective routinely collected clinical data

15.5.1 Aim

Generate an anonymised individual-level dataset that will be used to calibrate the model's parameters.

15.5.2 Study setting

This retrospective routinely collected clinical data required for this sub-study will be requested from our clinical partners – Liverpool University Hospitals NHS Foundation Trust (LUFT) and Liverpool Clinical Laboratories (LCL).

15.5.3 Study population

The aim of this study is to generate a computer simulation model of adult patients with infection requiring broad-spectrum antimicrobial therapy. The collection of blood cultures from patients signifies sufficient clinician concern for sepsis, bloodstream infection or systemic infection. The collection of blood cultures is also often followed with empirical antimicrobial therapy. Therefore, blood culture requests will be used as an objective inclusion criterion for retrospective data inclusion into the study.

Inclusion criteria:

- Patients who were managed within a LUFT acute hospital site (Aintree or Royal) and had a concurrent blood culture investigation requested
- Age ≥ 18 years at the time of the study

Exclusion criteria:

- Age < 18 years at the time of the study
- Blood culture requested but patient not managed on an acute hospital site

Study period:

A 5-year period is targeted for the data collection.

One of the anticipated challenges is a significant change to the number and management of patients with bloodstream infection during the SARS-CoV-2 pandemic. Therefore, the study period must overlap this period of disruption (e.g., 2018 – 2023) in order to ensure that the model is not biased by the impact of the pandemic. Conversely, the pandemic's impact on the data presents a unique opportunity to ensure that the model is not overfit to the status quo data and that it can continue to make good predictions in periods of system stress. This will be achieved by a secondary analysis that filters the input parameter data as follows:

1. Patients treated before the onset of the SARS-CoV-2 pandemic (baseline)
2. Patients treated during and after the SARS-CoV-2 pandemic (comparator)

Other data not covered by the study population:

Additional data is required to complete non-direct patient care elements of the model. However, since these data can be reported in an aggregate manner, they are not considered part of the study population:

- Patients presenting to A&E triage (per hour)
- Ward staffing levels (per shift)

15.5.4 Sampling

No sampling will be required, as all data within the study period will be included in the data analysis.

15.5.5 Consent

This sub-study will require the processing of retrospective data collected for routine clinical care delivery. Identifiable data will be collected initially for the purpose of linking records across the multiple database sources. This data processing will be without collecting retrospective patient consent. Once data linkage has taken place, the data will be fully anonymised prior to any further processing and analysis.

15.5.6 Methods

The raw data for this project will be requested from our NHS partner organisation (LUFT and LCL), subject to their information governance requirements. At a high level, the following data will be required to create an accurate simulation of patients with bloodstream infection:

- Presentation of patients for acute care, such as the rate of patients presenting to A&E triage and the proportion of patients triaged with a sepsis diagnosis
- Clinical status (observations such as blood pressure, temperature, respiratory rate, etc.) of patients on presentation
- Laboratory investigation data – which patients had blood cultures collected and what where the results?
- Antimicrobial treatment and allergy history – what empirical antimicrobials were administered, if any, and was treatment modified later, e.g., with blood culture results
- Ward staffing data – in order to make accurate predictions of changes in staffing, the general staffing rates of wards must be known (nursing, physicians and auxiliary staff)
- Other patient characteristics, such as basic demographics (including age), main diagnosis, and co-morbidities
- Timings of events (e.g., time of blood culture collection, time of prescription of antimicrobials)

It is expected that the data will have to be pulled from multiple databases, including:

- Coding and admissions data
- Administrative data (e.g., patient ward transfers)
- Staffing data (e.g., ward nurse staffing)
- Prescribing data (e.g., antimicrobials)
- Laboratory data (e.g., blood culture analytical data)

The required data fields are all routinely collected for the delivery of clinical care.

A complete table of the data fields that will be collected is available in Appendix 4 – Data collection fields.

15.5.7 Data analysis

Since data will be collected from disparate clinical databases, a data integration step is expected. This will be performed on the received pseudonymised data using Microsoft Excel, Microsoft Access and R (R Core Team, 2022). Records from different systems will be linked using a unique primary key (pseudonymous hash number generated by CIPHA system) in order to ensure referential integrity in later stages of data analysis. Details of the pseudonymised process can be found in section 16.7.

Data will then be analysed and summarised; the approach for this sub-study is to generate:

1) Descriptive data of patient characteristics and patient events (which will be utilised as co-variates during analyses), e.g., patient observations, blood culture bacterial isolate, blood culture susceptibility, antimicrobial agent prescribed
2) Time based events fit to an appropriate probability distribution, e.g.:
    a. Number of events per defined period of time – fit using a Poisson distribution, e.g., number of patients requiring empirical treatment for bloodstream infection per day/shift/hour
    b. Time required for events to take place – fit using an exponential distribution, e.g., time take to administer prescribed antimicrobial, time taken for incubated blood culture to report positive

Probability distributions that have a closed form solution will be fit using maximum likelihood estimation. More complex models where a closed form solution is not possible or feasible will be fit using Bayesian estimation in the Stan modelling language (Stan Development Team, 2022).

Downstream to this point, data will be utilised both in aggregate form (e.g., proportions/probability, or probability distribution parameters) and using non-identifiable patient level (e.g., to create realistic simulated patients).

Data produced from this sub-study will be used to calibrate the simulation model. The model can be expressed with the following abstraction:

1) Parameter inputs within a parameter space – e.g., patient characteristics, staffing numbers, time required to perform blood culture investigation. Some parameters may be known with high certainty (e.g., incubation times for blood cultures) while others could be estimated but with residual uncertainty (e.g., time it takes for a clinician to perform a clinical review).
2) Model agent interactions – i.e., the simulation process itself
3) Output measurements – e.g., antimicrobial usage within the simulated population, time taken to administer effective antimicrobials (defined as first dose of antimicrobial to which the infecting organism is susceptible)

Given that certain input parameters are likely to be known only with low certainty, metaheuristic stochastic optimisation methods will be required to make informed estimates of these parameters and allow accurate simulation. These methods allow the estimation of paramaters that cannot be differentiated, such as categorical variables or proportions. Examples of these methods include grid search, random search, hill climbing methods and genetic algorithms. Regardless of the optimisation method chosen, an objective function (e.g., a loss function of simulated output measurements and observed measurements) will be used to fit the model.

15.6 Sub-study 3 – Direct observation study

15.6.1 Aim

Produce detailed vertical observation data via direct patient observation to “fine tune” parameters that are otherwise insufficiently or inaccurately informed through retrospectively collected data.

15.6.2 Study setting

In order to maximise the data yield, the observational study will be performed on hospital areas expected to have the highest concentration of patients with suspected bloodstream infection and blood culture investigation requests:

- A&E
- Acute medical units

15.6.3 Study population

For the direct observation part of the study, blood culture requests are less suited as the sole primary inclusion criterion, since key events would have already occurred by the time the request is made. This particularly applies to unwell patients with sepsis. The primary inclusion criterion for the observational study will therefore also include the point at which a positive sepsis screen is triggered.

Eligibility criteria:

- LUFT patients receiving hospital care in one of the study settings (A&E or acute medical units) during one of the observational study days
- Positive sepsis screen trigger (as defined by LUFT sepsis policy) AND/OR blood culture investigation request
- Age 18 years or above at the time of positive sepsis screen

Exclusion criteria:

- Age under 18 years at the time of sepsis screen or blood culture investigation request
- Patient no longer physically located within the study setting

15.6.4 Sampling

Where possible, all events relating to the management of patients in the study population within the study setting will be observed and recorded. It is acknowledged that this may not always be possible, for example if multiple patients satisfy the inclusion criteria but are being managed concurrently. In this case, the observer is allowed discretion to determine whether observation of the multiple events is possible.

15.6.5 Recruitment

Patients and staff will not be directly recruited for this sub-study. Prior approval will be sought from the ward manager.

15.6.6 Consent

This study is a non-interventional observational study with the main observation metrics being event times. Therefore, consent requirements are proportionate to the level of risk. Prior approval will be sought from the ward manager. Brief information on the nature of the study will be made available to staff at the daily handover, and informational posters will be made available on the wards. Although individual written informed consent will not be sought, the information material will indicate that an opt out system of consent will be in place. Prior to observing the delivery of direct patient care, the observer will introduce themselves to the patient as an observer for research purposes and allow for

the opportunity to opt out of the observations. On opting out, the observer will discontinue their direct observation of the staff member/s and/or patient/s.

15.6.7 Outcomes and objectives

The main deliverable of this sub-study is an adjunctive dataset collected through direct observation of patient management. Compared to the data generated by Sub-study 2, this data is expected to be on fewer patients but of high accuracy and resolution, likely to focus on time periods with comparatively high event concentration (e.g., first 4 hours of treatment). Being a prospective observational study, the data generated will:

- Validate the process flowcharts from Sub-study 1, and reveal any events that may have been missed
- Provide higher accuracy time requirements for individual events, since the retrospectively collected data from Sub-study 2 can have inaccurately recorded time-stamps

15.6.8 Methods

The study member or collaborator will observe the clinical management of patients being treated for suspected infection, and record the following in a bespoke database:

- Brief description of event (e.g., blood culture draw)
- Resource/s required for event (e.g., staff member, blood culture collection kit)
- Time taken for event (e.g., in minutes or hours)
- Outcome of event (e.g., successful or unsuccessful blood culture draw)

Direct observation of medical care provision generates the risk of performance bias, in particular the Hawthorne (observer) effect on the staff providing the care (Sedgwick and Greenwood, 2015). Where possible, observations will be conducted with minimal impact on staff and patients.

16 Ethical and regulatory consideration

16.1 Assessment and management of risk

This study does not include an interventional component, therefore there is minimal risk of direct harm to patients or staff. The risks associated with the sub-studies are:

1) Sub-study 1 (qualitative structured interviews and focus groups) – the risk of harm associated with this study is negligible, since staff participation is entirely voluntary, and discussions will centre around policy and protocols rather than management of specific patients.
2) Sub-study 2 (retrospective data collection) – there is a risk of data protection breaches with this sub-study; mitigations discussed in later sections
3) Sub-study 3 (direct observational study):
    a. risk of data protection breaches (low risk, since patient identifiable data will not be collected during this observational study)
    b. risk of observer bias influencing the delivery of patient care

The availability of integrated retrospective data (routinely collected for clinical care) is dependent on data engineering and data sharing agreements that are not yet in place. There is a risk that data availability is not of sufficient quality or is delayed, impacting the progression through study milestones. We will review progress on the technical and governance infrastructure at the first annual milestone, and review this risk assessment. If required, a substantial amendment will be submitted to pursue an alternate route for data access.

16.2 Research Ethics Committee (REC) and other Regulatory review & reports

The study will be conducted in accordance with the principles of Good Clinical Practice (GCP). Before the start of the study, Liverpool University Hospitals Foundation Trust sponsorship will be obtained by the Chief Investigator (CI) via Liverpool's Single Point of Access to Research and Knowledge (SPARK), a joint research service that encompasses NHS Liverpool Health Partners (LHP) and University of Liverpool research support functions.

A favourable opinion will be sought from the Health Departments Research Ethics Service NHS Research Ethics Committee (REC) and Health Research Authority (HRA) via the Integrated Research Application System (IRAS). All correspondence with the REC will be retained. The CI will submit an annual progress report (APR) to the REC within 30 days of the anniversary date on which the favourable opinion was given, and annually until the study is declared ended. If the study is ended prematurely, the CI will notify the REC, including the reasons for the premature termination.

Confirmation of Capacity and Capability will be obtained from each recruiting site before any participants are recruited at that site.

16.3 Amendments

A study management group (SMG) will be convened and chaired by the CI with representation from across the study team to provide ongoing management and oversight, support amendments and deal with reporting, data issues and protocol deviations. It is the sponsor's responsibility to decide whether an amendment is substantial or non-substantial for the purposes of submission to the REC, HRA and any other applicable regulatory bodies. Amendments will be categorised according to HRA guidance (http://www.hra.nhs.uk/resources/after-you-apply/amendments/). If a substantial amendment to the protocol is required, the chief investigator (CI) will submit a valid notice of amendment via the sponsor to the relevant regulatory bodies for consideration via the appropriate route (e.g., via the IRAS Amendment Tool). Amendment history will be tracked by keeping all protocol versions in a secure password-protected folder in the University of Liverpool computer system. Substantial amendments will not be undertaken until all necessary regulatory approvals have been given.

16.4 Peer review

The design of this study was peer reviewed as part of the Wellcome Trust (the funder) funding application. The protocol was peer-reviewed by Dr George L Drusano MD (Professor of Medicine, Institute for Therapeutic Innovation, University of Florida).

The reviewers are external to the investigators' host institution, independent, not involved in the study in any way, and have sufficient knowledge of the clinical subject area to consider the study's methdological and service aspects.

16.5 Patient & Public Involvement

This study involves the integration of data from different sources (mainly from LUFT IT systems but may also include specimens sent from local General Practices) for use in AMR research. An AMR Citizen's Jury was commissioned by the University of Liverpool (UoL) in 2022 to explore the jury's support for the use of integrated pseudo-anonymised data to (Centre for New Democratic Processes, 2022):

- Inform individual treatment and hospital utilisation
- Identify trends in AMR manifesting as serious infections
- Identify unmet clinical need and shape the research and development of new medicines

Jurors were generally supportive of healthcare staff having access to this data (98% of jurors were "very supportive" or "somewhat supportive"). Overall jurors were moderately supportive of researchers having access to this data (74% of jurors were "very supportive" or "somewhat supportive"). The results of this citizen's jury informed the design of this study and protocol, in particular the prioritisation of the simulation model being used by decision makers for hospital utilisation.

The study aims and design are regularly presented to members of the public through a UoL-funded Patient and Public Involvement and Engagement (PPIE) programme.

16.6 Protocol compliance

Any and all protocol deviations will be recorded and reviewed as part of the regular SMG meetings and reported to the sponsor where appropriate.

16.7 Data protection and patient confidentiality

The retrospective data used for this study will be generated by CIPHA – governed by a Population Health data sharing agreement that is in place between CIPHA and trusts/GP practices providing the data. Data will be pseudonymised by the time access is provided to the study team.

A high-level description of the DSCRO-level anonymisation process is the following:

- Patient name removed
- Date of birth converted to age
- Address converted to first part of post code only
- Hospital or NHS numbers irreversibly hashed to maintain data integrity across data tables (no hashing key will be kept, therefore reversing the anonymisation will not be possible by the study team)

Access to the anonymised datasets will be limited to members and collaborators of the study group. Access will be password protected by NHS or UoL IT security. The anonymised data will be stored for 10 years after publication, in accordance with the UK Research and Innovation (UKRI) best practice recommendations (UK Research and Innovation, 2018). A data catalogue will be maintained. The data custodian will be the CI, and will be responsible for the safe custody, transport and storage of any aggregate data resulting from this research.

Sub-study 3 (direct observation) requires the linking of observed events to a particular source (patient) and routinely collected clinical data. Patient hospital number (RQ6) will be used as the main identifier to link observations to an individual, and observations to routinely collected clinical data. This will be stored in a password-protected database on the sponsor's NHS IT system (shared drive). The data

collected in the observational study (sub-study 3) will be anonymised prior to transfer to university (University of Liverpool) systems for analysis. This will include:

1. Replacement of hospital number (RQ6) with a random hash number that preserves data association with a single individual (research staff conducting analysis will not have access to a reversing cipher).
2. Demographic information such as name, date of birth, postcode, will not be collected and hence not available for transfer
3. Replacement of real date-time events with a random date that preserves relativity within the care episode. This will be achieved as follows:
    a. Generate a random date-time for each patient care episode, drawn from a range of 100 years
    b. Add this date-time to every date-time event for each patient respectively
    c. Date-time events will have correct interval from admission time, but it will not be possible for the research team to know exact date and time
    d. Research team will not have access to the random date-time added to each patient

Any transfer of data to university systems will be done using encrypted email and files stored on the university cloud (Office 365 – OneDrive) as recommended by university governance policy.

16.8 Data flows

A data flow chart for this study is available in Appendix 7 – The study team will make an application under existing population health data sharing agreements, led by the CI, to the CIPHA Data Asset and Access Group (DAAG) to:

1. Conduct a data gap analysis to identify which data streams need to be set up. Some data tables, such as patient demographics, are already available within the CIPHA and covered by a pre-existing data sharing agreement (DSA). Others, such as data from the Telapath microbiology system from LCL and EPMA prescribing data from LUHFT need to be prospectively incorporated. Any new data streams will be governed by existing data sharing agreements between the trust (sponsor) and CIPHA.
2. Once data streams and transfers are established, any data processing will take place within the TRE administered by CIPHA. This TRE is administered by CIPHA. CIPHA meets requirements of GDPR, The Data Protection Act 2018 and the NHS Data Security and Protection Toolkit, and is certified to ISO27001, ISO9001 and Cyber Essentials standards. Individual user access is limited, and user permissions are based on the minimum access required to conduct the processing required.
3. CIPHA data engineer to perform patient-level linkage of new data sources to existing CIPHA data sources under existing data sharing agreements
4. Permit access of researchers within the study group access to the CIPHA TRE to perform analysis of linked pseudonymised data as described in the protocol methodology. Researchers will not have access to the pseudonymisation cipher.

This approval will be sought using a DAAG data access request form completed by the CI and a Tier Two DSA completed by LUHFT data controllers for transfer of new data where required.

The legal basis for the study under General Data Protection Regulation (GDPR) is Article 9(2)(j), in that data processing and access is necessary for the purpose of scientific or statistical purposes in

accordance with Article 89(1). Data controllers, the third party (NHS CIPHA) and investigators must comply with the requirements of the Data Protection Act 1998 with regards to the collection, storage, processing and disclosure of personal information and will uphold the Act's core principles. Only the minimum amount of data pertaining to the minimum number of individuals required to facilitate the study methodology will be used. Identifiable data will be kept secure within an encrypted, password protected CIPHA data warehouse and only accessible by a CIPHA engineer assigned to data linkage and anonymisation. The data will only remain identifiable until linkage has occurred, at which point identifiable fields will be removed. The anonymisation procedure is described in Section 16.7.

16.9 Indemnity

The study will be covered by the sponsor's standard legal indemnity and insurance policies for a non-clinical trial of an investigational medicinal product (CTIMP)

16.10 Access to the final study dataset

The final study dataset will be accessible to the CI, research associates and data scientists within the approved study team, and the CIPHA engineer(s) performing data linkage and anonymisation.

17 Dissemination policy

17.1 Dissemination policy

The data arising from the study will be owned by the study authors. On completion of the study, the data will be analysed and tabulated, and a Final Study Report prepared, which will be accessible on the Centres for Antimicrobial Optimisation Network (CAMO-NET) website. Participating investigators can publish any of the study data with permission of the CI. Public participants of the Liverpool AMR Citizens' Jury will be informed of the outcome of the study by provision of the publication. All statistical and machine learning code used to generate the results will be made available open source, shared via a GitHub public repository following journal publication.

17.2 Authorship eligibility and use of professional writers

The main study results will be published as soon as a manuscript is completed, in a peer-reviewed journal, on behalf of all collaborators. The manuscript will be prepared by a writing group composed of the investigators. All investigators will be granted authorship on the final study report in line with International Committee for Medical Journal Editors criteria for individually named authors and group authorship.

18 References

Academy of Medical Royal Colleges, 2022. Statement on the initial antimicrobial treatment of sepsis V2.0.

Borshchev, A., Filippov, A., 2004. From System Dynamics and Discrete Event to Practical Agent Based Modeling: Reasons, Techniques, Tools. p. 23.

Cahill, W., Render, M., 1999. Dynamic simulation modeling of ICU bed availability 4.

Centre for New Democratic Processes, 2022. Liverpool Antimicrobial Resistance (AMR) Citizen's Jury.

Chaudhuri, A.B., 2020. Flowchart and algorithm basics: the art of programming. Mercury Learning and Information, Dulles, VA.

Davies, S.C., Fowler, T., Watson, J., Livermore, D.M., Walker, D., 2013. Annual Report of the Chief Medical Officer: infection and the rise of antimicrobial resistance. The Lancet 381, 1606–1609. https://doi.org/10.1016/S0140-6736(13)60604-2

Dormuth, C.R., Burnett, S., Schneeweiss, S., 2005. Using Policy Simulation to Predict Drug Plan Expenditure when Planning Reimbursement Changes: PharmacoEconomics 23, 1021–1030. https://doi.org/10.2165/00019053-200523100-00005

European Antimicrobial Resistance Collaborators, 2022. The burden of bacterial antimicrobial resistance in the WHO European region in 2019: a cross-country systematic analysis. Lancet Public Health S2468-2667(22)00225–0. https://doi.org/10.1016/S2468-2667(22)00225-0

Galea, S., Riddle, M., Kaplan, G.A., 2010. Causal thinking and complex system approaches in epidemiology. International Journal of Epidemiology 39, 97–106. https://doi.org/10.1093/ije/dyp296

Gunal, M., Pidd, M., 2006. Understanding Accident and Emergency Department Performance using Simulation, in: Proceedings of the 2006 Winter Simulation Conference. Presented at the 2006 Winter Simulation Conference, IEEE, Monterey, CA, USA, pp. 446–452. https://doi.org/10.1109/WSC.2006.323114

Günal, M.M., Pidd, M., 2011. DGHPSIM:: Generic simulation of hospital performance. ACM Trans. Model. Comput. Simul. 21, 1–22. https://doi.org/10.1145/2000494.2000496

Kerr, C.C., Stuart, R.M., Mistry, D., Abeysuriya, R.G., Rosenfeld, K., Hart, G.R., Núñez, R.C., Cohen, J.A., Selvaraj, P., Hagedorn, B., George, L., Jastrzębski, M., Izzo, A.S., Fowler, G., Palmer, A., Delport, D., Scott, N., Kelly, S.L., Bennette, C.S., Wagner, B.G., Chang, S.T., Oron, A.P., Wenger, E.A., Panovska-Griffiths, J., Famulare, M., Klein, D.J., 2021. Covasim: An agent-based model of COVID-19 dynamics and interventions. PLoS Comput Biol 17, e1009149. https://doi.org/10.1371/journal.pcbi.1009149

Kianmehr, H., Sabounchi, N.S., Sabounchi, S.S., Cosler, L.E., 2020. A system dynamics model of infection risk, expectations, and perceptions on antibiotic prescribing in the United States. J Eval Clin Pract 26, 1054–1064. https://doi.org/10.1111/jep.13203

Lodemann, T., Akçalı, E., Fernandez, R., 2022. Process Modeling of ABCDE Primary Survey in Trauma Resuscitations: A Crucial First Step for Agent-Based Simulation Modeling of Complex Team-Based Clinical Processes. Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare Publish Ahead of Print. https://doi.org/10.1097/SIH.0000000000000622

Medical Research Council, 2023. Health Data Access Tool Kit [WWW Document]. URL https://hda-toolkit.org/story_html5.html (accessed 4.3.23).

Mielczarek, B., Uziałko-Mydlikowska, J., 2012. Application of computer simulation modeling in the health care sector: a survey. SIMULATION 88, 197–216. https://doi.org/10.1177/0037549710387802

Miller, G., Randolph, S., Gower, D., 2004. Simulating the Response of a Rural Acute Health-care Delivery System to a Bioterrorist Attack. International Journal of Disaster Medicine 2, 24–32. https://doi.org/10.1080/15031430410024363

Mitchell, J., O'Neill, A.J., King, R., 2022. Creating a framework to align antimicrobial resistance (AMR) research with the global guidance: a viewpoint. Journal of Antimicrobial Chemotherapy 77, 2315–2320. https://doi.org/10.1093/jac/dkac205

Monks, T., Currie, C.S.M., Onggo, B.S., Robinson, S., Kunc, M., Taylor, S.J.E., 2019. Strengthening the reporting of empirical simulation studies: Introducing the STRESS guidelines. Journal of Simulation 13, 55–67. https://doi.org/10.1080/17477778.2018.1442155

Nance, R.E., 1996. A history of discrete event simulation programming languages, in: Bergin, T.J., Gibson, R.G. (Eds.), History of Programming Languages---II. ACM, New York, NY, USA, pp. 369–427. https://doi.org/10.1145/234286.1057822
Nianogo, R.A., Arah, O.A., 2015. Agent-Based Modeling of Noncommunicable Diseases: A Systematic Review. Am J Public Health 105, e20–e31. https://doi.org/10.2105/AJPH.2014.302426
Nygaard, K., Dahl, O.-J., 1978. The development of the SIMULA languages. SIGPLAN Not. 13, 245–272. https://doi.org/10.1145/960118.808391
Ordu, M., Demir, E., Davari, S., 2021. A hybrid analytical model for an entire hospital resource optimisation. Soft comput 25, 11673–11690. https://doi.org/10.1007/s00500-021-06072-x
R Core Team, 2022. R: A Language and Environment for Statistical Computing.
Ratnapalan, S., Lang, D., 2020. Health Care Organizations as Complex Adaptive Systems. Health Care Manag (Frederick) 39, 18–23. https://doi.org/10.1097/HCM.0000000000000284
Rauner, M.S., 2001. Using Simulation for AIDS Policy Modeling: Benefits for HIV/AIDS Prevention Policy Makers in Vienna, Austria 14.
Saville, C., Monks, T., Griffiths, P., Ball, J.E., 2021. Costs and consequences of using average demand to plan baseline nurse staffing levels: a computer simulation study. BMJ Qual Saf 30, 7–16. https://doi.org/10.1136/bmjqs-2019-010569
Sedgwick, P., Greenwood, N., 2015. Understanding the Hawthorne effect. BMJ h4672. https://doi.org/10.1136/bmj.h4672
SimPy, 2020.
Stan Development Team, 2022. Stan Modeling Language Users Guide and Reference Manual.
Taylor, J., Hafner, M., Yerushalmi, E., Smith, R., Bellasio, J., Vardavas, R., Bienkowska-Gibbs, T., Rubin, J., 2014. Estimating the economic costs of antimicrobial resistance: Model and Results. RAND Corporation, Santa Monica, CA. https://doi.org/10.7249/RR911
UK Research and Innovation, 2018. Guidance on best practice in the management of research data.
WHO, 2016. Global action plan on antimicrobial resistance. Geneva.

19 Appendices

Appendix 1 – Overview of Agent Based Modelling and Discrete Event Simulation

There are different types of computer simulation models that could be implemented in this research work. The core entity requiring simulation is a hospital environment within which patients are investigated and managed for sepsis/bloodstream infection/systemic infection.

Although a deterministic model of differential equations could be applied, this would require the model to either be a very high-level abstraction of the system, or otherwise require a level of complexity and detail that would not be practical. The core challenge of simulating complex systems is that system behaviour cannot be understood by describing it as a combination of independent individual components (Galea et al., 2010). Other methods of computer simulation, such as agent-based modelling and discrete event simulation take a bottoms-up approach to model construction, starting with abstract representations of system components (such as patients, staff, laboratory machines, etc) which are instructed to follow real-world behaviour. The interaction between these components is allowed to run within the simulation, collecting system-wide aggregate output data either during, or at the end of the simulation.

Agent based modelling (ABM)

Agent based modelling is characterised by three core components (Nianogo and Arah, 2015):

1) Agents (e.g., patients, staff), consisting of a set of attributes and behaviours
2) Relationships between patients (e.g., nurses and clinicians reviewing patients)
3) Spatial environment within which agents interact[1]

The behaviour of agents is simulated at every time step of the simulation, which can be designed to represent real-world time (e.g., minutes, hours, etc). Healthcare interest in ABM has been increasing over the last decade, driven by their use in communicable diseases and increased availability of the intensive computing resources required to simulate large models.

Discrete event simulation (DES)

Discrete event simulation involves a mathematical model of a system whose state is updated at discrete time points within the simulation. Again, the time points can represent real-world time. DES models are particularly useful for process-driven scenarios or queuing problems. Within healthcare, their most common application has been in Accident and Emergency operational management. The history of DES extends back to the 1950s, and has a close relationship with important milestones in computer science, such as the development of the SIMULA programming language (Nance, 1996; Nygaard and Dahl, 1978).

In DES models, agents progress through a series of discrete events, which may have a time requirement and resource requirement. For example, this may be the collection of blood cultures from a patient, which require the appropriate staff member to be present (resource) and inherently require a variable amount of time (which can be drawn from a probability distribution).

[1] Geospatial environment can be omitted if not important for the particular simulation

Appendix 2 – Amendment History

Amendment No.	Protocol version no.	Date issued	Author(s) of changes	Details of changes made

Appendix 3 – Data collection fields for Sub-study 2

Table 5 - Data fields post-processing

Field	Type
Demographic data	
Age	Date
Sex	Text
Admission details	
Patient hospital location	Text
Discharge date and time	Date-time
Triage/nurse assessment date and time	Date-time
Clinician assessment date and time	Date-time
Clinical details	
Coding diagnosis	Text
Sepsis score	Number
Medication allergies	Text
Allergy reactions	Text
Treatment details	
Antimicrobial prescription date and time	Date-time
Antimicrobial name	Text
Antimicrobial dose	Number
Antimicrobial administration date and time	Date-time
Investigation details	
Blood culture request date and time	Date-time
Blood culture collection date and time	Date-time
Blood culture laboratory booking in date and time	Date-time
Blood culture interim authorisation date and time	Date-time
Blood culture authorisation date and time	Date-time
Microbiology details	
Organism identification	Text
Antimicrobial susceptibility	Text

Table 6 - Raw data fields (pre-processing)

Database	Field	Example
Demographic data	NHS number	111 111 1111
Demographic data	Hospital number	RQ61111111
Demographic data	Postcode	L1 1AA
Demographic data	Date of birth	01/01/2000
Prescribing data	NHS number	111 111 1111
Prescribing data	Hospital number	RQ61111111
Prescribing data	Drug name	amoxicillin
Prescribing data	Location	Liverpool Medical Centre
Prescribing data	Route	oral
Prescribing data	Dose	500mg
Prescribing data	Frequency	every 8 hours
Prescribing data	Start date & time (prescription)	01/01/2022 10:00
Prescribing data	End date & time (prescription)	07/01/2022 08:00
Prescribing data	Start date & time (administration)	01/01/2022 10:00
Prescribing data	End date & time (administration)	07/01/2022 08:00
Prescribing data	Drug allergies	penicillin
Prescribing data	Allergy reaction type	Rash
Microbiology data	NHS Number	111 111 1111
Microbiology data	Hospital number	RQ61111111
Microbiology data	Date & Time Collected	01/01/2022 00:00
Microbiology data	Date & Time Received	02/01/2022 12:00
Microbiology data	Date & Time Authorised	03/01/2022 20:00
Microbiology data	Specimen type	blood
Microbiology data	Report code	YUPO
Microbiology data	Comment code	IUTI
Microbiology data	Specimen site	right arm

Microbiology data	Location	Ward 4
Microbiology data	Specimen number	c,22.1111111.A
Microbiology data	Clinical Details	unstructured free text
Microbiology data	Epithelial cells	+++
Microbiology data	White cell count	> 50
Microbiology data	Red cell count	> 50
Microbiology data	Organism count	> 10^7
Microbiology data	Organism code	ESCO
Microbiology data	Organism name	E coli
Microbiology data	AML	S
Microbiology data	AMP	I
Microbiology data	OX	R
Microbiology data	MET	S
Microbiology data	MEL	I
Microbiology data	AMC	R
Microbiology data	P/T	S
Microbiology data	CL	I
Microbiology data	CXM	R
Microbiology data	CPD	S
Microbiology data	CRO	I
Microbiology data	FOX	R
Microbiology data	CTX	S
Microbiology data	CAZ	I
Microbiology data	CZA	R
Microbiology data	CIP	S
Microbiology data	LEV	I
Microbiology data	CN	R
Microbiology data	AK	S
Microbiology data	TE	I
Microbiology data	TGC	R
Microbiology data	MER	S

Microbiology data	ETP	I
Microbiology data	TEC	R
Microbiology data	VA	S
Microbiology data	DA	I
Microbiology data	E	R
Microbiology data	ATM	S
Microbiology data	LZ	I
Microbiology data	TSU	R
Microbiology data	W	S
Microbiology data	NIT	I
Microbiology data	FOT	R
Microbiology data	MUP	S
Microbiology data	DAP	I
Microbiology data	RD	R
Microbiology data	OB	S
Microbiology data	MTZ	I
Microbiology data	C	R
Microbiology data	FD	S
Microbiology data	FLC	I
Microbiology data	MCA	R
Microbiology data	ANI	S
Microbiology data	CAS	I
Microbiology data	VOR	R
Microbiology data	AMB	S
Other pathology data	NHS Number	111 111 1111
Other pathology data	Hospital number	RQ61111111
Other pathology data	PaO2	7.0kPa
Other pathology data	PaCO2	7.0kPa
Other pathology data	pH	7.0
Other pathology data	HCO3	7.0
Other pathology data	FiO2	32%

Other pathology data	Lactate	1.5
Other pathology data	Bilirubin	15
Other pathology data	ALT	15
Other pathology data	Alkaline phosphatase	15
Other pathology data	Urea	15
Other pathology data	Creatinine	85
Other pathology data	Hb	85
Other pathology data	Platelets	130
Other pathology data	White cell count	2.4
Other pathology data	Neutrophils	2.4
Other pathology data	Eosinophils	2.4
Other pathology data	Lymphocytes	2.4
Other pathology data	Basophils	2.4
Other pathology data	Monocytes	2.4
Other pathology data	CRP	34
Other pathology data	PT	34
Other pathology data	APTT	34
Other pathology data	Fibrinogen	34
Other pathology data	Date & time collected	02/01/2022 12:00
Other pathology data	Date & time reported	03/01/2022 20:00
Admission data	NHS Number	111 111 1111
Admission data	Hospital number	RQ61111111
Admission data	Admission date & time	01/01/2022 00:00
Admission data	Admission location	A&E
Admission data	Transfer date & time	02/01/2022 12:00
Admission data	Transfer location	ward 4
Admission data	Discharge date & time	03/01/2022 20:00
Admission data	Discharge status	discharged home
Admission data	Discharge location	nursing home
Admission data	Vital status	alive, dead
Episode data	NHS Number	111 111 1111

Episode data	Consultation date & time	01/01/2022 00:00
Episode data	Coded diagnosis	UTI
Episode data	Diagnosis date	01/01/2022 00:00
Electronic healthcare record data	NHS Number	111 111 1111
Electronic healthcare record data	Hospital number	RQ61111111
Electronic healthcare record data	Clinical form date & time	01/01/2022 00:00
Electronic healthcare record data	Clinical notes	unstructured free text
Electronic healthcare record data	Clinical form date & time	02/01/2022 12:00
Electronic healthcare record data	Clinical form	semi-structured free text
Electronic healthcare record data	Observation date & time	03/01/2022 20:00
Electronic healthcare record data	Heart rate	100
Electronic healthcare record data	Respiratory rate	20
Electronic healthcare record data	Systolic blood pressure	120
Electronic healthcare record data	Diastolic blood pressure	80
Electronic healthcare record data	Temperature	36.5
Electronic healthcare record data	AVPU score	A
Electronic healthcare record data	Oxygen saturation	96%
Electronic healthcare record data	FiO2	32%
Electronic healthcare record data	NEWS score	4
Electronic healthcare record data	eSepsis pathway date & time	03/01/2022 20:00

Electronic healthcare record data	eSepsis pathway	semi-structured free text
Electronic healthcare record data	AKI alert	1
Diagnostic coding data	NHS Number	111 111 1111
Diagnostic coding data	Hospital number	RQ61111111
Diagnostic coding data	Primary diagnostic code	Urinary tract infection
Diagnostic coding data	Secondary diagnostic code	Sepsis

Appendix 4 – Discussion prompts for Sub-study 1 meeting convenors

1. One-to-one data generation session
    - Please state your job title and key roles/responsibilities.
    - This study is exploring the key processes relating to the management of patients with sepsis, bloodstream infection and/or severe infection. These are patients that generally involve the following:
        - Collection of blood cultures (and other microbiology samples if appropriate, such as sputum or urine cultures)
        - Administration of empirical antibiotics (antibiotics given prior to blood culture or test results)
        - Review of empirical antibiotics, either with microbiology test results, or for the purpose of "step-down"
    - Within your role, how do you become aware of patients that fit the above criteria?
    - Considering a hypothetical patient that fits the above criteria, describe their general management (focusing from your perspective).
    - What key decisions or events do you think are most important?
    - Do you involve other staff members (clinical or non-clinical) within these decisions or events?
    - How does your involvement with the patient finish?
2. One-to-one validation session
    - We have analysed your comments from our first meeting and constructed the following draft flowchart of a patient being managed for sepsis, bloodstream infection and/or severe infection. What are your initial thoughts on this flowchart?
    - If you try tracing through the flowchart (thinking of a hypothetical patient may help), are you able to reach the end?
    - Does the flowchart capture all the key decisions and events from your perspective?
    - Do you see any obvious flaws in the flowchart logic (e.g., possibility of entering an infinite loop)?
    - Do you have any other feedback or changes to recommend?

Appendix 5 – Flowchart Design Principles

Adapted from (Chaudhuri, 2020; Lodemann et al., 2022). Where possible, principles and design shapes in Figure 3 *are in accordance with the ISO standard 5807:1985 Information processing — Documentation symbols and conventions for data, program and system flowcharts, program network charts and system resources charts*

1. Agree in advance a minimal set of design shapes, and only use shapes from the set (an example set is provided below)
2. The flowchart should read from top to down and left to right
3. Each shape must only have one entry point and one exit point, with the important exception of the decision symbol
4. All decision branches must be well-labelled
5. Events or decisions that require the availability of information must clearly show this
6. Events or decisions that require a staff member (or other resource) must clearly show this
7. Consider the use of re-entrant processes when events are stochastic and apply to multiple parts of the flowchart

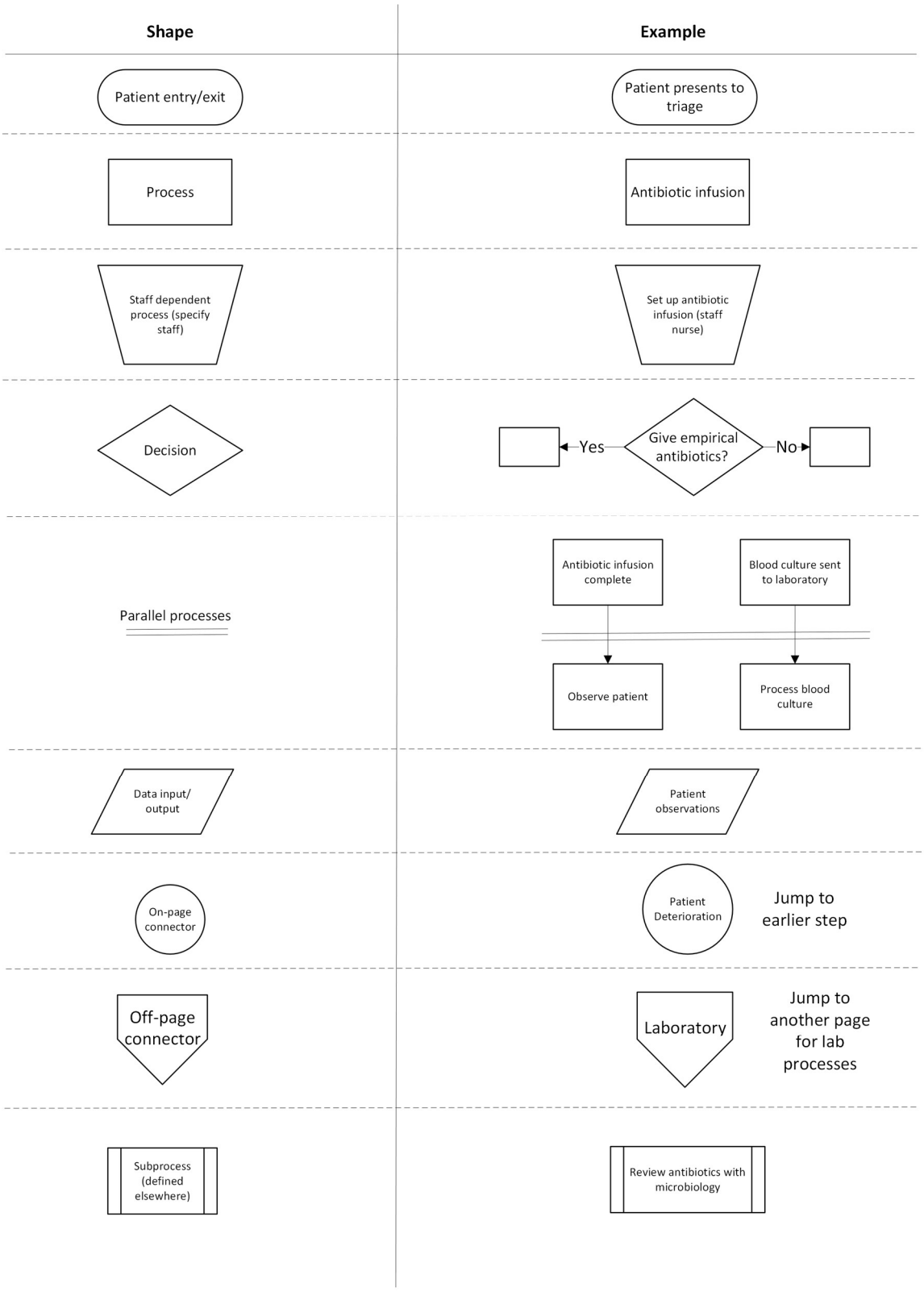


Figure 3 – Proposed flowchart shapes (based on ISO 5807)

Appendix 6 – Staff consent form

To facilitate printing, the request form is appended in the next page.

Staff Participant Consent Form

Title of Study: Computer Simulation Model of patients with bloodstream infection, sepsis, and systemic Infection (CSM-BSI)

Principal Investigator: Alessandro Gerada

Thank you for replying to the invitation to participate in CSM-BSI as an expert staff member. This study aims to explore your experiences, perceptions, and understanding related to the management of patients with bloodstream infection, sepsis, and system infection. Before you decide whether to participate, it is important for you to understand the purpose, procedures, and risks of the study. Please read this form carefully and ask any questions you may have before agreeing to participate.

Purpose of the study:

The purpose of this study is to create a computer software that simulates patients with bloodstream infection, sepsis, and other systemic infections. This software will be used to inform our understanding of these conditions, and optimise the management and treatment of patients.

Procedures:

If you agree to participate in this study, you will be asked to attend structured interviews and focus group sessions. The interviews and focus groups will be conducted by trained researchers and will be audio recorded to allow qualitative analysis of the conversations. This will be stored alongside a general description of your role (no personal identifying information will be stored). The interviews and focus group sessions will be organised at a time and location that is convenient for you.

Risks:

There are minimal risks associated with participating in this study. We will not specifically ask questions related to how you personally have managed individual patients. Please inform the researcher if there are any questions or topics that you do not wish to discuss.

Benefits:

This study will give you the opportunity to contribute to medical research and the possibility of gaining a deeper understanding of your experiences, perceptions, and attitudes related to patient care.

Confidentiality:

Voice recordings will be stored in a secure platform with access only to researchers authorised by the study team. Voice recordings will generally only be accessed for transcription purposes. Transcriptions of your conversations during the study structured interviews and focus groups will be analysed and used to inform the computer software development. We will not seek to collect any other personal data.

Voluntary participation:

Participation in this study is completely voluntary. You may refuse to participate or withdraw your participation at any time without penalty or loss of benefits to which you are otherwise entitled. Study session attendance must be outside of time that is rostered for delivery of clinical care.

Contact information:

If you have any questions or concerns about the study, you may contact the principal investigator at alessandro.gerada@liverpool.nhs.uk.

Consent:

By signing below, you indicate that you have read and understood the information provided in this consent form and agree to participate in this study.

Participant Signature

Date

Appendix 7 – Data flow template

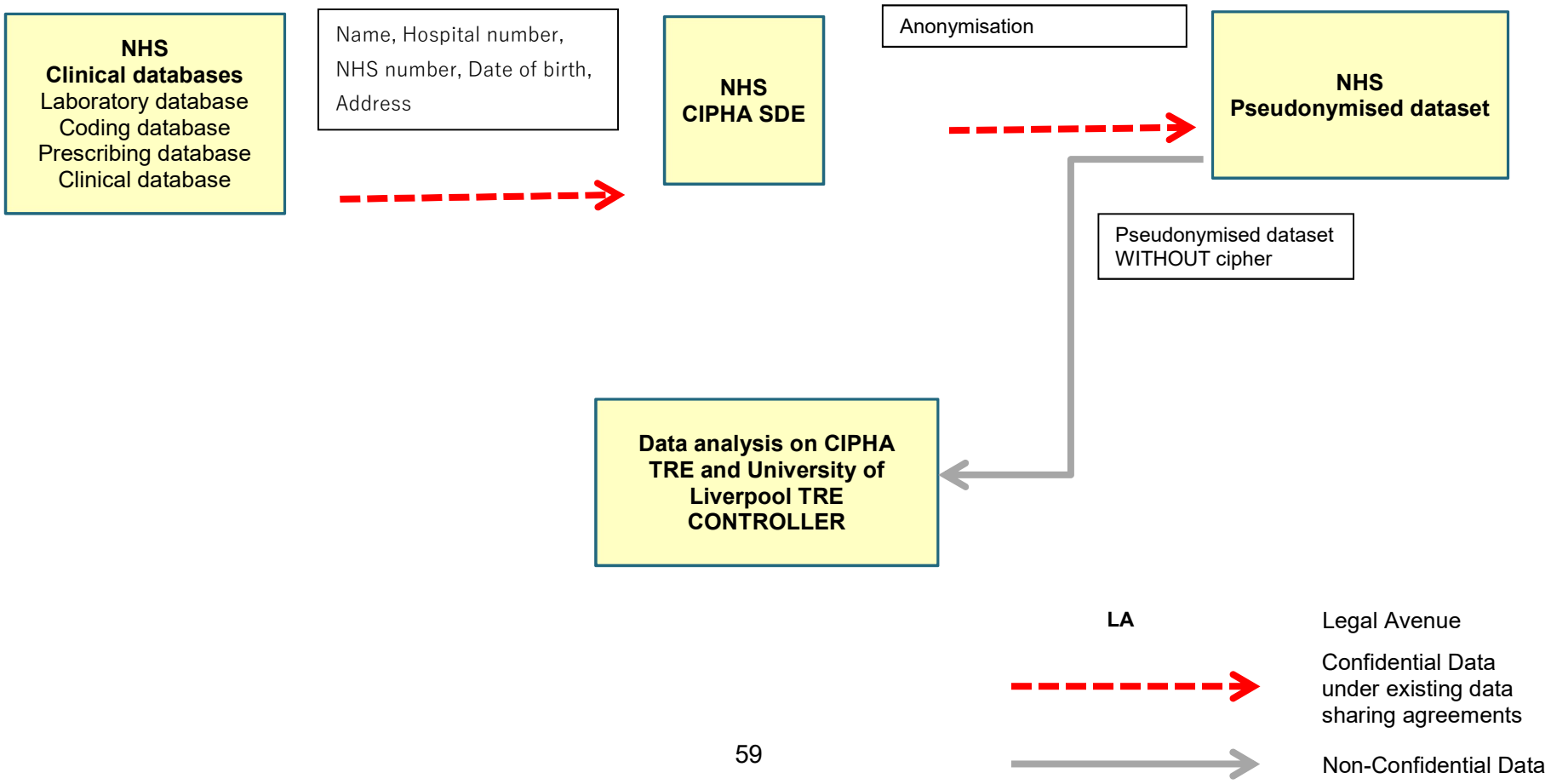
This template is based on the Medical Research Council Health Data Access Tool Kit (Medical Research Council, 2023).

Data Flow Diagram

Project Name:

Computer Simulation Model of patients with Bloodstream infection, Sepsis and systemic Infection (CSM-BSI) Date: 2023-10-18

Chief Investigator: Alessandro Gerada IRAS ID: 334396



Description of dataflow:

Data will have originated from NHS IT systems (LUFT) and transferred to CIPHA (within NHS firewall) for integration and pseudonymisation process.

Data will be requested by the study team from CIPHA via a DAAG request on the basis of the Population Health benefit of this study (under existing population health data sharing agreements).

Study team will be processing fully pseudonymised data on CIPHA TRE or University of Liverpool cloud servers.

Appendix 8 – List of additional files

Name	Filename	Version	Date	Comment
Participant letter of invitation	CSM-BSI Participant letter of invitation_v1.0.docx	1.0	2023-10-18	To be sent through email
Study information – ward poster	CSM-BSI study information ward poster.docx	1.0	2023-10-18	
Staff participant consent form	Staff Participant Consent Form v1.0.docx	1.0	2023-09-18	Content is the same as Appendix 6
SoECAT form	CSM-BSI_SoECAT_Wellcome_150722.xlsx		2022-07-15	
PI's CV	CV_AG.pdf			
Observational data entry database	CSM-BSI-data-entry.accdb	0.5	October 2023	
SPARK sponsorship application	SPARK-CSM-BSI_sponsorship_application.docx		2023-07-12	
Non-technical summary	CSM-BSI_non_technical_summary.docx	1.0	2023-10-20	
Signed DARF form				
Intent to sponsor				

*** END DOCUMENT **