

Developing an Antimicrobial Strategy for Sepsis in Malawi

-

Thesis submitted in accordance with the requirements of the Liverpool School of Tropical Medicine for the degree of Doctor in Philosophy by Joseph Michael Lewis

August 2019

Contents

Preface	9
1 Introduction	11
1.1 Chapter Overview	13
1.2 Sepsis in sub-Saharan Africa	13
1.3 ESBL-E in sub-Saharan Africa	13
1.4 Conclusions	13
1.5 Thesis overview	13
1.6 Appendix	13
1.7 References	13
2 Methods	15
2.1 Chapter Overview	15
2.2 Study site	15
2.3 Clinical Study	18
2.4 Diagnostic Laboratory Procedures	23
2.5 Molecular methods	26
2.6 Bioinformatics	26
2.7 Statistical Analysis	27
2.8 Study Team	28
2.9 Data Collection and Storage	28
2.10 Ethical Approval, Consent and Participant Remuneration	29
3 <i>Mycobacterium tuberculosis</i> BSI: an IPD meta analysis	31
4 Sepsis in Blantyre, Malawi	33
4.1 Chapter overview	34
4.2 Aims and Methods	34
4.3 Results	34

5	Early response to resuscitation in sepsis	35
6	Gut mucosal carriage of ESBL-E in Blantyre, Malawi	37
7	Whole genome sequencing of ESBL <i>E. coli</i> carriage isolates	39
7.1	Chapter overview	41
7.2	Methods	41
7.3	Results	41
7.4	Discussion	41
7.5	Appendix	41
	References	43

List of Tables

2.1 Study inclusion criteria	19
--	----

List of Figures

2.1	Map of study site	16
2.2	Overview of patient sampling schedule	22

Preface

Placeholder

Chapter 1

Introduction

Placeholder

1.1 Chapter Overview

1.2 Sepsis in sub-Saharan Africa

1.2.1 Search strategy

1.2.2 Defining sepsis

1.2.3 Applicability of sepsis-3 definitions in sub-Saharan Africa

1.2.4 Sepsis epidemiology in sub-Saharan Africa

1.2.4.1 Incidence

1.2.4.2 Risk factors: the sepsis population in sub-Saharan Africa

1.2.4.3 Outcomes

1.2.5 Sepsis aetiology in sub-Saharan Africa

1.2.5.1 Bacterial zoonoses, Rickettsioses and arboviruses

1.2.5.2 HIV opportunistic infections: PCP, histoplasmosis and cryptococcal disease

1.2.6 Sepsis management

1.2.6.1 Early goal directed therapy

1.2.6.2 Evidence to guide antimicrobial therapy in sSA

1.2.6.3 Evidence to guide intravenous fluid therapy in sub-Saharan Africa

1.3 ESBL-E in sub-Saharan Africa

1.3.1 Search strategy

1.3.2 Introduction: definition and classification of ESBL-E

1.3.3 Global molecular epidemiology of ESBL-E: an overview

1.3.3.1 1980s-1990s: First identification of ESBL in nosocomial pathogens

1.3.3.2 1990s-2010s: Emergence and globalisation of CTX-M

Chapter 2

Methods

2.1 Chapter Overview

This chapter gives an overview of the clinical study which underpins this thesis, and the laboratory and computational procedures used in analysis. Further details are given in the individual chapters, where necessary.

2.2 Study site

2.2.1 Malawi

Malawi is a country of 17.5 million people in South-Eastern Africa[1]. It is one of the poorest countries in the world: it is a low income country under the World Bank classification, with a 2017 Gross National Income (GNI) per capita of \$320 in US dollars[2], and was ranked 171st of 189 countries in 2017 by the human development index (HDI), a composite statistic of life expectancy, education and income per capita indicators[(UNDP human development reports)][3]. In 2010, 71% of the population was estimated to survive on less than \$1.90 per day. Life expectancy at birth in 2017 was 63 years, and though significant progress is being made, neonatal and under-5 mortality remains high at 23 and 55 per 1000 live births, respectively. The population is largely rural (83% in 2017), with a young population (44% under the age of 15, 2017) and high fertility rate[2]. Malaria is endemic, and there is an ongoing generalised HIV epidemic: adult HIV prevalence (age 15-49) was estimated to be 9.6% in 2017 (UNAIDS), though falling from a peak of 16.6% in 1999[4]. HIV antiretroviral therapy (ART) national scale up began in 2004 and in 2017 71% of eligible adults and children were estimated to be receiving ART[4]. It is classed by the WHO as a high-TB/high-HIV



Figure 2.1: Malawi, showing administrative boundaries (North, Central, and South regions), Lilongwe, the capital city and Blantyre, the study location. Source: openstreetmap.org, used under Creative Commons Attribution ShareAlike 2.0 licence CC-BY-SA

burden country, with an estimated TB incidence rate of 131 (95% CI 70-210) cases per 100000 population per year[WorldHealthOrganisation2018].

It has a subtropical climate, with three main seasons: a warm wet season from November to April, a cooler dry winter period from May to August and a hot dry period from September to October. Blantyre city, the location of the study in this thesis, is the second city of Malawi with a population of 585000. It is located in Blantyre district, population 995000 in 2018[1], in the Shire highlands at an altitude of 1000m (Figure 2.1).

2.2.2 Queen Elizabeth Central Hospital

Queen Elizabeth Central Hospital (QECH), located in Blantyre city, is the tertiary referral hospital for the Southern Region of Malawi. It has 1300 beds but often operates above capacity, and is the only site providing free inpatient healthcare to the adult population of

Blantyre district. Since 2011 it has had a dedicated emergency department for adults, the Adult Emergency and Trauma Centre, staffed 24 hours a day. Since 2015 (and for the whole of the study period), attendees to the AETC must be referred from a primary health clinic. Adults attending the AETC are triaged by a nurse and then reviewed by a doctor or clinical officer; if admission under a specialty team (including medicine) is deemed appropriate then a patient will be reviewed by an intern or registrar from the admitting specialty and usually by a consultant within 24 hours. There is a 6-bed AETC resuscitation area in which oxygen concentrators, cardiac monitors and a defibrillator are available; none of these items are available in the rest of the AETC.

There are two dedicated single-sex medical wards, each of approximately 60 beds, and one mixed-sex TB ward. Male and female high-dependency units (HDUs), each with a capacity of six beds, have oxygen concentrators (or, if available, oxygen cylinders) to deliver supplemental oxygen. The medical wards are staffed by two or three trained nurses and a variable number of nursing students. Basic nursing care is usually provided by a patient's relative or friend, called a 'guardian.' Patients on the medical wards are reviewed twice-weekly by a consultant physician and then variably at other times by junior doctors, clinical officers or medical students depending on the availability of staff. Malawi national treatment guidelines suggest ceftriaxone for the treatment of sepsis requiring hospitalisation.

2.2.3 Participating Laboratories

2.2.3.1 Malawi-Liverpool-Wellcome Clinical Research Programme

The Malawi-Liverpool Wellcome Trust Clinical Research Programme was established in 1995 and since then has been active in researching priority health issues in Malawi. It is an affiliate of the Malawi College of Medicine, and is based in the grounds of QECH in Blantyre. It runs an on-site microbiology laboratory which has provided an aerobic blood culture service to QECH since 1998, and also provides CSF microscopy and culture. Bacterial culture is carried out as per British Society of Antimicrobial Chemotherapy (BSAC) guidelines[5] and the laboratory adheres to UK NEQAS external quality control. It is core funded by the Wellcome Trust.

2.2.3.2 Malawi College of Medicine Tuberculosis Laboratory

2.2.3.3 Wellcome Trust Sanger Institute

The Wellcome Sanger Institute is a research institute based in Hinxton, UK, which was established in 1993, and undertakes research in all aspects of genomics including bacterial genomics as part of the parasite and microbes programme. It has one of the largest DNA sequencing facilities in the world as well as high performance computing clusters. It is funded by the Wellcome Trust.

2.3 Clinical Study

The DASSIM (Developing an Antimicrobial Strategy for Sepsis in Malawi) study was an observational cohort study, recruiting from the AETC at QECH with two broad aims: firstly, to describe the presentation, aetiology and determinants of outcome in sepsis in Malawi; and secondly to determinants of carriage of ESBL-E in sepsis survivors.

2.3.1 Entry Criteria

The study was open for recruitment between 19 February 2017 and 2 October 2018; there were three arms. Firstly, in order to define the aetiology of sepsis, adults with sepsis attending AETC were recruited as early as possible in their attendance to QECH following triage. Secondly, in order to clearly define the relative effects of antimicrobial exposure versus hospital admission in ESBL-E acquisition and carriage, antibiotic unexposed adults attending AETC with no plan for antimicrobial administration were recruited. Finally, to define baseline flux of ESBL-E antibiotic-unexposed community members were recruited. Detailed inclusion and exclusion criteria for each arm are given shown in Table 2.1. For logistic reasons, recruitment occurred 7am - 5pm Monday to Friday.

Exclusion criteria were the same for all arms of the study: Participants were not eligible for enrollment if they were unable to give informed consent and no guardian available to provide proxy consent; they spoke neither English nor Chichewa; or they lived > 30km from Blantyre city. The antibiotic-unexposed inpatients and community members were matched on age (+/- 5yr) and sex to sepsis survivors (defined as patients surviving to 28 days). In addition community members were matched on location to sepsis survivors; putative households for recruitment were identified by random walk from the houses of surviving sepsis participants, with initial direction established by spinning a bottle on the floor.

Table 2.1: Study inclusion criteria

Study Arm	Inclusion Criteria
Arm 1 - Sepsis	Adults (16 years or over) attending AETC AND Axillary temperature $> 37.5^{\circ}\text{C}$ or history of fever within 72 hours AND Life threatening organ dysfunction defined by any one of: Oxygen saturations < 90 percent on air, respiratory rate > 30 breaths/minute, systolic blood pressure $< 90\text{mmHg}$, glasgow coma scale < 15
Arm 2 - Inpatient	Adults (16 years or over) attending AETC AND No suspicion of infection or plan for antimicrobial administration AND No antimicrobial therapy within last 4 weeks
Arm 3 - Community	Adults (16 years or over) AND No antimicrobial therapy within last 4 weeks

Note:

AETC = Adult Emergency and Trauma Centre

2.3.2 Study Visits and Patient Sampling

2.3.2.1 Enrollment assessment and first six hours

An overview of the study visit schedule is shown in Figure 2.2. At enrollment, following consent a baseline questionnaire was completed (by the guardian if the participant was obtunded) to capture background demographic and clinical data, a Chichewa-language EQ-5D health related quality of life questionnaire completed, and sample collection undertaken (see below). For hospitalised patients, enrollment aimed to be as soon after triage as possible, and arm 1 (sepsis) patients underwent a more intensive assessment over the first six hours of enrollment. Data on therapies administered over this time period were captured by hourly review by a member of the study team, including the time of administration of antimicrobial therapy

and volumes of intravenous fluid administered, the latter visually confirmed by the study team member. Hourly vital signs and standardised bedside assessment of intravascular status were performed: central capillary refill, limb temperature and passive leg raise. For this latter test, blood pressure and pulse rate were measured with the participant sitting at 45 degrees, and one minute after lying the participant supine and lifting the legs to 45 degrees. All blood pressure measurements were made noninvasively with an automated cuff (Omron M2, Omron, Japan), oxygen saturations measured with a dedicated study pulse oximeter (Contec CM50, Contec Medical Systems, China), and temperature measured in the axilla with a digital thermometer (Omron FWH000, Omron, Japan).

All treatment decisions were at the discretion of the participant's attending healthcare worker.

2.3.2.2 Subsequent visits

Hospitalised participants were briefly reviewed daily Monday to Friday following admission whilst on the wards to capture details of therapies administered. For the 72 hours following admission, vital signs were repeated and an assessment made of intravenous fluid administered; if records of fluid balance were incomplete (as was anticipated to be the case in the majority of cases) then fluid administered was estimated from participant- or guardian-reported number of bags of fluid administered.

The visit schedule for the study is shown in Figure 2.2. Arm 1 (sepsis) and 2 (hospital inpatients) were reviewed by a study team member 7 and 28 days following enrollment, and at 3 and 6 months; community members were reviewed at 28 days and 6 months. At each visit, details of any antimicrobials or other therapies received and any health care facility contact were collected, and the Chichewa language EQ-5D was repeated. Samples were collected as below.

These study visits occurred preferentially at QECH, but at the patient's home if it was difficult for them to attend the hospital. If patients missed a scheduled visit, then attempts were made to contact them by telephone. If these attempts were unsuccessful then attempts to visit them at home were undertaken by a member of the study team.

2.3.2.3 Blood, urine, and stool, sputum and CSF collection

Blood was collected from arm 1 (sepsis) participants only on enrollment and at the 4-week visit. At baseline, blood was collected aseptically directly into collection tubes with a vacutainer device with the following order of draw and volumes: one BacT/Alert (BioMerieux, France) aerobic blood culture bottle (7-10ml), one BD BACTEC Myco/F Lytic mycobacterial culture

bottle (3-5ml) (Beckton Dickinson, United States), one serum (10ml) and two EDTA (4ml, both Grenier Bio-One, Austria) samples. Finger prick for capillary blood was used for point of care diagnostics as described below. At the 4-week visit, 10mls of convalescent serum was collected. At baseline, urine was collected from all arm 1 (sepsis) participants into a sterile polypropylene universal container (Alpha Laboratories, UK) either by the participant themselves (if this was possible) or with the aid of a disposable bedpan.

At the enrollment visit and all other visits for all arms of the study stool was collected into a sterile polypropylene universal container (Alpha Laboratories, UK) with aseptic technique. If it was not possible for a patient to provide stool then a rectal swab was taken using a sterile rayon-tipped swab (Technical Service Consultants Ltd, UK) inserted into 2-3cm into the rectum, rotated for approximately 10 seconds and placed into Amies gel media for transport to the laboratory.

The decision to collect sputum for Xpert testing for tuberculosis or to perform lumbar puncture (LP) rested with the participant's attending healthcare worker. Lumbar puncture, where done, was carried out by QECH AETC or medicine department doctors or clinical officers and 7-10ml of cerebrospinal fluid (CSF) was aseptically collected. Sputum samples were collected in sterile polypropylene universal container and transported to the Malawi College of Medicine (CoM) tuberculosis laboratory. Myco/F lytic bottles for mycobacterial blood culture were also transported to the CoM tuberculosis laboratory; all other samples were transported to the MLW laboratories in the first instance.

2.3.2.4 Imaging: chest x-ray and ultrasound scanning

Ultrasound examination following the FASH (Focussed ASsessment for HIV-associated tuberculosis) protocol[6] was undertaken for all sepsis participants by me, following training with a UK-registered consultant radiologist. The protocol aims to ascertain presence or absence of six features likely to be associated with extrapulmonary TB: pericardial effusion, abdominal lymph nodes, pleural effusion, ascites, hypoechoic focal liver lesions or hypoechoic focal splenic lesions. Scanning was undertaken with a xxxxxx device with a curvilinear probe; images were saved and 10% of scans were read by a UK-registered consultant radiologist for quality control.

Decision to undertake chest x-ray was at the discretion of the participant's attending healthcare worker. If chest x-ray was done then digital pictures were taken and stored in the study database.

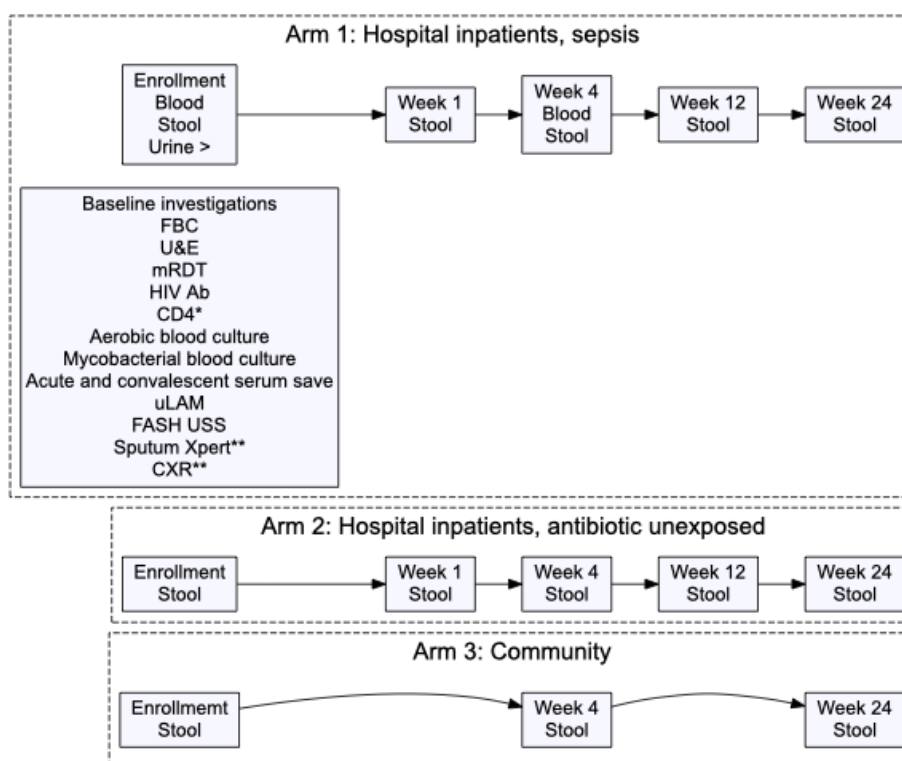


Figure 2.2: Overview of patient sampling schedule. FBC = Full blood count, U&E = Urea and electrolytes, mRDT = Malaria rapid diagnostic test, uLAM = Urinary LAM, FASH USS = Focussed assessment for HIV-associated tuberculosis ultrasound scan

2.3.3 Outcomes and sample size calculations

The two co-primary outcomes for the study are:

- 1) 28 day mortality.
- 2) Acquisition of ESBL-E detected in stool by aerobic culture by 28 days.

Two power calculations were undertaken: to inform the size of the sepsis cohort, and one for the non-sepsis participants. The sepsis component was powered to detect clinically relevant risk factors for death; we made the a priori assumption that a risk ratio of 2 or more is clinically relevant in this setting and patient population. The initial calculation suggested 250 patients with severe sepsis with 50% mortality (as seen in previous studies in Malawi) would have 80% power to detect risk factors for death with prevalence of 20-50% that confer a risk ratio (RR) of 1.5-3; therefore, a RR of 2 is likely to be detected. Logistic considerations resulted in a reduction of a feasible sample size to 225; post-hoc power calculation suggested 80% power to detect risk factors conferring a RR of 2 with 25-50% prevalence, assuming 50% mortality.

For the second co-primary outcome, assuming 30% of participants with sepsis (and hence antimicrobial exposure) acquire ESBL-E by 28 days, and that 50% of them die by day 28, 125 antibiotic unexposed hospitalised participants would give 80% power to detect a 50% relative difference in acquisition. Logistic considerations resulted in the target sample size being reduced to 100 participants; revised post-hoc power calculation suggested that the 80% power to detect a 50% relative difference was maintained.

In order to define the baseline flux of ESBL-E organisms and ESBL genes, one community member for each antibiotic unexposed hospitalised participant, giving a target sample size of 100.

2.4 Diagnostic Laboratory Procedures

Results of all laboratory diagnostic procedures were fed back to participant's attending healthcare worker in real time. If any investigation result became available after the participant was discharged (e.g. Mycobacteraemia was identified) then they were contacted and called back to the hospital to be reviewed by me, and referred for care to the relevant QECH department.

2.4.1 Point of care diagnostics

Point of care tests were carried out on capillary blood for capillary lactate (Lactate Pro 2, Arkray, Japan), for *P. falciparum* HRP-2 antigen (Paracheck Pf, Orchid Biomedical, India) and, if HIV status was unknown, to test for HIV antibodies using Determine HIV 1/2 kit (Abbott Diagnostic Division) and Unigold HIV 1/2 kit (Trinity Biotech Inc.) following Malwian national guidelines[7].

2.4.2 Laboratory diagnostics

2.4.2.1 Haematology and biochemistry

Automated full blood counts (Beckman Coulter HmX Haematology Analyser, Beckman Coulter, USA) were undertaken in the MLW laboratories on EDTA whole-blood samples, as were CD4 cell counts (Becton Dickinson FACSCount, Becton Dickinson, USA) for HIV-infected participants. Serum samples were allowed to settle upright for 30-60 minutes and then centrifuged (at 1300 g for 10 min). Biochemistry testing for urea and electrolytes (Beckman Coulter AU480 Chemistry Analyser, Beckman Coulter, USA) was undertaken immediately and a maximum of three 1.8ml aliquots of serum stored at -80C for later analysis.

2.4.2.2 Aerobic blood and CSF culture

Blood in aerobic culture bottles was incubated in an automated system (BacT/Alert BioMerieux) and identified to species level using the API system (Biomérieux, France) and standard techniques (ref). Anaerobic culture was not available. Coagulase-negative *Staphylococci*, *bacillus spp.*, *diphtheroids* and alpha-haemolytic *Streptococci* other than *S. pneumoniae* were considered as contaminants. Antimicrobial sensitivity testing was undertaken using the disc diffusion method following BSAC guidelines to amoxicillin, chloramphenicol, co-trimoxazole, gentamicin, ciprofloxacin and ceftriaxone.

CSF, where available, was cultured onto blood, chocolate and Sabouraud agar and then identification of any growth undertaken as above.

2.4.2.3 Mycobacterial blood culture

Blood in Myco/F Lytic bottles was cultured at 37C in and inspected daily for the first 14 days with a handheld UK Wood's lamp, and once every two days thereafter. Contents of the bottles were centrifuged (3000xg for 20 minutes) within 48hr of detection of fluorescence,

examined by ZN and gram stain to exclude contamination, then inoculated into mycobacterial growth indicator tubes (MGIT, Becton Dickinson Diagnostic Systems, United States) for up to 6 weeks. Isolates were classified as *M. tuberculosis* or nontuberculous Mycobacteria by microscopic cording and MPT-64 lateral flow assays (TAUNS Laboratories, Japan).

After 8 weeks of culture of Myco/F Lytic bottles if no fluorescence was seen, then centrifugation and microscopy was carried out. If no organisms were seen, the culture was reported as “no growth.”

2.4.2.4 Sputum Xpert

Sputum Xpert testing (Cepheid, United States) was carried out as per the manufacturers instructions: sputum specimens were mixed with sample reagent and incubated at room temperature for 15 minutes. The liquefied specimen was then loaded into the Xpert MTB/RIF test cartridge for processing.

2.4.2.5 Urinary LAM

Urinary LAM testing was carried out on all available urine samples from HIV-infected participants, using the Alere Deterine TB LAM Ag lateral flow assay (Alere, United States) following the manufacturer’s instructions. Frozen urine samples were used: they were allowed to come to room temperature and then briefly centrifuged to remove debris. 60 microliters of urine was applied to the sample pad and the result read after 25 minutes by comparing to the provided reference scale card. The results were read independently by two readers, with a tie-break read by a third reader in the even of disagreement.

2.4.2.6 Selective stool culture for ESBL-E

Stool and rectal swabs received in the laboratory were stored at 4C pending processing, before being plated onto commercially available ESBL selective media (CHROMagar ESBL media, CHROMagar, France) and cultured aerobically at 37C overnight. Rectal swabs were streaked directly onto the plate, or a cotton tipped applicator used for solid stool. CHROMagar is also chromogenic: *E. Coli* colonies are pink, *Klebsiella spp.*, *Enterobacter spp.* and *Citrobacter spp.* blue and other species white. Blue or white colonies were speciated using the API 20E system. Morphologically distinct colonies were confirmed to be ESBL producers using combination disc methods: the putative ESBL producer was cultured overnight on ISO-sensitest agar with discs of cefotaxime and ceftazidime (30 micrograms) with and without clavulanic acid (10 micrograms), and ESBL production confirmed if there was a difference of at least 5mm

between discs with and without clavulanic acid . For organisms likely to carry a chromosomal AmpC and hence be capable of hydrolysing cefotaxime and ceftazadime, (*Enterobacter spp.*, *Citrobacter freundii*, *Morganella morganii*, *Providencia stuartii*, *Serratia spp.*, *Hafnia alvei*) cefipime (30 micrograms), an AmpC-stable cephalosporin was used with and without clavulanic acid (10 micrograms). Antimicrobial sensitivity testing to meropenam, amikacin, gentamicin, co-trimoxazole, chloramphenicol and ciprofloxacin following BSAC guidelines[5] was undertaken for a subset of confirmed ESBL-producers.

2.4.2.7 Acute and convalescent serologies

PHE stuff

2.5 Molecular methods

One of each morphologically distinct colony from ESBL-E selective culture was taken forward for DNA extraction and whole genome sequencing. DNA was extracted from overnight nutrient broth culture using the Qiagen DNA mini kit (Qiagen, Germany) following the manufacturers instructions. Extracted DNA was shipped to the Wellcome Sanger Institute where it underwent library preparation according to the Illumina protocol and paired-end sequencing on Illumina HiSeqX (*check this*) by the WSI DNA pipelines team.

2.6 Bioinformatics

All analyses were undertaken on the WSI computing cluster running Ubuntu v12.04.2, Precise Pangolin. Quality control of reads for each sample was carried out using metrics provided the WSI QC pipeline including total number of reads, and statistics obtained by mapping a random 100 Mbases from each sample to a reference and calculating depth of coverage, number of heterogeneous SNPs, GC content and insert size. Kraken[8] v0.10.6, a kmer-based taxonomic assignment tool, was used to assign reads to a species and to identify suspected sample contamination or species misclassification. Reads were *de novo* assembled with SPAdes[9] v3.11.0, with default settings. Assembly metrics of quality (e.g. number of contigs, largest contig, N50 [the contig length on which at least half of the bases in an assembly have been assembled]) were calculated using QUAST[10] v4.6.0 and completeness and contamination of the assemblies assessed by checkM[11] v1.0.7, a tool that uses co-located sets of lineage associated genes to assess genome completeness and contamination. Contaminated samples

(with > 25% contamination as defined by CheckM) or poor assemblies (with N50 < xxxxx) were discarded.

Prokka[12] v1.5 with a genus specific database from RefSeq was used to annotate the retained assemblies. The Roary v1.007 pan-genome pipeline[13] was used to construct a core genome with a blastp percentage identity of 95%; genes were considered core if they were present in 99% or more of isolates. A core gene multiple sequence alignment was generated using maaft[14] v7.205, SNP-sites identified using SNP-sites[15] v2.4.1 and the resultant SNP alignment used to build a maximum likelihood phylogenetic tree using IQ-TREE[16] v1.6.3. Trees were visualised in the ggtree v1.14.4 package[17] in R.

Ariba[18] v2.12.1 was used to identify AMR-associated genes using the SRST2[19] database, to identify plasmid replicons using the PlasmidFinder database[20] and to perform *in silico* multi-locus sequence typing (MLST) using the database from <http://mlst.warwick.ac.uk/mlst/dbs/Ecoli> accessed via www.pubmlst.org. *E. coli* phylogrouping was performed with a quadruplex *in silico* PCR using the Clermont scheme[21]. To track carriage of similar bacteria within and between individuals, clustering of the core gene alignment was performed using hierarchical BAPS (bayesian analysis of population structure) algorithm[22], implemented in the rhierbaps package v1.1.0 in R, with 2 levels. To track putative mobile genetic elements (MGE), a difficult task with short-read sequencing, isolates ESBL-gene containing contigs were identified using BLASTn[23] v2.7.0 with and then contigs clustered using cd-hit[24] v4.6 with sequence identity threshold 0.9, word size 5 and accurate mode (command line flag -g) to produce mutually exclusive ESBL-gene contig groupings.

2.7 Statistical Analysis

Details of analysis methods are given in the relevant chapters. All analyses were undertaken using R v3.5.1 and any Bayesian modelling using Markov Chain Monte Carlo (MCMC) methods using Stan v2.18 via the Rstan interface with R. This thesis was written using the bookdown package in R and the code to generate it as well as all analysis scripts are available at <https://github.com/joelewis101/thesis>. Unless otherwise states, all boxplots show median and IQR as boxes, 1.5 times IQR as whiskers and outliers (any points outside the whiskers) as points.

2.8 Study Team

I am the principal investigator of the study, and led a study team which consisted of: Emma Smith, a medical doctor who managed the running of the study from March 2017 - May 2017; study nurses Lucy Keyala, Grace Mwaminawa and Tusekile Phiri who recruited patients, collected samples and completed follow up visits; field workers Witness Mtambo and Gladys Namacha who recruited patients, collected samples and completed follow up visits; and laboratory technicians Madlitso Mphasa and Rachel Banda who processed laboratory specimens, did stool culture and DNA extractions, and the uLAM testing. The MLW core laboratory staff undertook haematology and biochemistry testing and blood and CSF culture, and the Malawi CoM tuberculosis laboratory staff carried out tuberculosis culture and Xpert testing.

2.9 Data Collection and Storage

Data collection used both the Open Data Kit software (ODK) and structured TeleForm paper forms. Electronic ODK forms were loaded onto Asos ZenTouch tablets running Android using the ODK Collect Android app. Completed forms were pushed daily to the dedicated secure study SQL (structured query language) database built and administered by the MLW data team. Teleform forms were checked by me, batched and scanned by the MLW data team, and variable values automatically extracted by the TeleForm system; validation was undertaken by the MLW data team and data queries generated for missing or invalid values. Queries were resolved either by manually reconciling with the forms or by discussion with the clinical team. Once the data passed validation it was pushed to the SQL database. Completed paper TeleForm records were stored securely in the MLW data department. All data on the study database is stored securely with access restricted to the study PI and the database administrators in the MLW data department. Other paper records including consent forms were stored in a locked cabinet with access restricted to members of the study team. Results of laboratory investigations in the MLW laboratory were stored in the MLW PreLink laboratory information management system (LIMS), anonymised and linked only to the participant unique study ID number. For analysis, anonymised data were extracted from the study SQL and LIMS databases as comma delimited files.

2.10 Ethical Approval, Consent and Participant Remuneration

The study protocol was granted ethical approval by the Malawi College of Medicine Research Ethics Committee (COMREC), protocol number P.11/16/2063 and the Liverpool School of Tropical Medicine (LSTM) Research Ethics Committee, protocol number 16-062. LSTM acted as the study sponsor. All study team member were trained in NIHR Good Clinical Practice (GCP).

All participants in the study provided informed, written consent, as follows: if a patient lacked capacity to provide informed consent themselves then informed consent was sought from the patient's representative, usually their guardian. Patients with no representative and lacking capacity to provide informed consent were not recruited. Initially, eligible patients (or their representative) were approached by a member of the study team and the study was explained to them including study procedures, risks and benefits, financial and confidentiality considerations and how to obtain more information. Written patient information leaflets were provided in English and Chichewa. If the patient/representative was willing to enter the study then they were asked to sign and date two copies of the consent form, and provided with a copy of the form to keep. If either the patient or representative was unable to read then the consent form was read to them by the study team, and witnessed by an additional staff member who is not a member of the study team. If the patient (or their representative) agreed to enter the study, then the witness signed and dated the form. Any patient who was enrolled to the study after consent from a representative and who then regained the capacity for informed consent was then approached independently for informed consent. A patient could withdraw at any time without giving a reason.

Hospitalised patients were not financially compensated for their time and/or transport costs, but all other participants were, following standard MLW guidelines: 500MWK (approximately USD\$0.7 at February 2019 exchange rates) was provided to participants if they were visited at home and 2000MWK (approximately USD\$2.8 at February 2019 exchange rates) or their reported transport costs (whichever was higher) for participants who attended QECH for study purposes.

Chapter 3

Mycobacterium tuberculosis BSI: an IPD meta analysis

Chapter 4

Sepsis in Blantyre, Malawi

Placeholder

4.1 Chapter overview

4.2 Aims and Methods

4.3 Results

4.3.1 Study population

4.3.2 Symptoms and health-seeking behaviour

4.3.3 Admission physiology and laboratory investigations

4.3.4 Aetiology

4.3.5 Treatment

4.3.6 Outcome

4.3.7 Determinants of mortality

4.3.8 Discussion

4.3.9 Further work

4.3.10 Appendix

Chapter 5

Early response to resuscitation in sepsis

Chapter 6

Gut mucosal carriage of ESBL-E in Blantyre, Malawi

Chapter 7

Whole genome sequencing of ESBL *E. coli* carriage isolates

Placeholder

7.1 Chapter overview

7.2 Methods

7.2.1 Bioinformatic pipeline

7.2.2 Global *E. coli* collection

7.2.3 Statistical analysis

7.3 Results

7.3.1 Samples and quality control

7.3.2 Phylogroup, MLST and core genome phylogeny of study isolates

7.3.3 Study isolates in a global context

7.3.4 Antimicrobial resistance determinants

7.3.4.1 β -lactam resistance

7.3.4.2 Quinolone resistance

7.3.4.3 Aminoglycoside resistance

7.3.4.4 Chloramphenicol, co-trimoxazole, tetracycline and other resistance determinants

7.3.4.5 Clustering and lineage association of AMR determinants

7.3.5 Plasmid replicons

7.3.6 Testing metadata associations: SNP distance, hierBAPS sequence clusters and ESBL-clusters

7.3.6.1 Hierarchical BAPS clustering of core gene pseudosequences

7.3.6.2 ESBL-clusters

7.3.6.3 Assessing for healthcare-associated lineages

7.3.6.4 Assessing for within-patient conservation of lineage or MGE

7.4 Discussion

References

- 1 National Statistics Office of Malawi. Malawi population projections. www.nsomalawi.mw (accessed 10 Feb 2019).
- 2 World Bank. World Bank Country and Lending Groups. <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups> (accessed 28 Feb 2019).
- 3 United Nations Development Programme. Human Development Indices and Indicators: 2018 Statistical Update. 2018.
- 4 UNAIDS. Malawi Country Profile. <http://aidsinfo.unaids.org/> (accessed 28 Feb 2019).
- 5 British Society for Antimicrobial Chemotherapy. BSAC Methods for Antimicrobial Susceptibility Testing. 2013.
- 6 Heller T, Wallrauch C, Goblirsch S *et al.* Focused assessment with sonography for HIV-associated tuberculosis (FASH): a short protocol and a pictorial review. *Critical ultrasound journal* 2012;**4**:21. doi:10.1186/2036-7902-4-21
- 7 Ministry of Health. Government of Malawi. Malawi HIV Testing Service Guidelines. 2015.
- 8 Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biology* 2014;**15**:R46. doi:10.1186/gb-2014-15-3-r46
- 9 Bankevich A, Nurk S, Antipov D *et al.* SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology* 2012;**19**:455–77. doi:10.1089/cmb.2012.0021
- 10 Gurevich A, Saveliev V, Vyahhi N *et al.* QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 2013;**29**:1072–5. doi:10.1093/bioinformatics/btt086
- 11 Parks DH, Imelfort M, Skennerton CT *et al.* CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome research* 2015;**25**:1043–55. doi:10.1101/gr.186072.114

- 12 Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;**30**:2068–9. doi:10.1093/bioinformatics/btu153
- 13 Page AJ, Cummins CA, Hunt M *et al.* Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 2015;**31**:3691–3. doi:10.1093/bioinformatics/btv421
- 14 Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution* 2013;**30**:772–80. doi:10.1093/molbev/mst010
- 15 Page AJ, Taylor B, Delaney AJ *et al.* SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microbial Genomics* 2016;**2**. doi:10.1099/mgen.0.000056
- 16 Nguyen L-T, Schmidt HA, Haeseler A von *et al.* IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution* 2015;**32**:268–74. doi:10.1093/molbev/msu300
- 17 Yu G, Smith DK, Zhu H *et al.* Ggtree : An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* 2017;**8**:28–36. doi:10.1111/2041-210X.12628
- 18 Hunt M, Mather AE, Sánchez-Busó L *et al.* ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. *Microbial genomics* 2017;**3**:e000131. doi:10.1099/mgen.0.000131
- 19 Inouye M, Dashnow H, Raven L-A *et al.* SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Medicine* 2014;**6**:90. doi:10.1186/s13073-014-0090-6
- 20 Carattoli A, Zankari E, García-Fernández A *et al.* In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrobial agents and chemotherapy* 2014;**58**:3895–903. doi:10.1128/AAC.02412-14
- 21 Clermont O, Christenson JK, Denamur E *et al.* The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environmental Microbiology Reports* 2013;**5**:58–65. doi:10.1111/1758-2229.12019
- 22 Cheng L, Connor TR, Siren J *et al.* Hierarchical and Spatially Explicit Clustering of DNA Sequences with BAPS Software. *Molecular Biology and Evolution* 2013;**30**:1224–8. doi:10.1093/molbev/mst028
- 23 Altschul SF, Gish W, Miller W *et al.* Basic local alignment search tool. *Journal of Molecular Biology* 1990;**215**:403–10. doi:10.1016/S0022-2836(05)80360-2

24 Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 2006;**22**:1658–9. doi:10.1093/bioinformatics/btl158