University of Southampton

Faculty of Clinical and Experimental Sciences

School of Medicine

Acute Undifferentiated Febrile Illness in the UK and Investigating the Diagnostic Utility of Metagenomic Next Generation Sequencing

Rebecca Louise Houghton

ORCID ID 0000-0002-7139-4866

Thesis for the degree of DM Infection and Immunity

January 2022

University of Southampton

Abstract

Faculty of Clinical and Experimental Sciences

School of Medicine

Thesis for the degree of DM Infection and Immunity

Acute Undifferentiated Febrile Illness in the UK and Investigating the Diagnostic Utility of Metagenomic Next Generation Sequencing

Rebecca Louise Houghton

Aims

Acute undifferentiated febrile illness (AUFI) has not been well studied in the UK. The global literature suggests that the majority of AUFI is likely to be attributable to infection, however a high proportion of individuals remain without a microbiologically confirmed diagnosis. Indiscriminate broad-spectrum antimicrobial use is common and mortality can be high. This proof of concept study aimed to characterise the clinical features, aetiology, antimicrobial use and clinical outcomes of adults hospitalised with AUFI at a large teaching hospital in the UK. In addition, metagenomic next generation sequencing (mNGS) was retrospectively used as an experimental diagnostic tool in an attempt to further elucidate the microbiological causes of AUFI.

Methods

One hundred adults with AUFI and fifty healthy volunteers were recruited in this prospective controlled cohort study. AUFI participants were recruited in hospital and followed up at 4 to 6 weeks following onset of fever. A standardised set of diagnostic tests were performed in addition to those requested by the treating physician. EDTA whole blood, serum, nasopharyngeal swabs and urine were taken for research purposes. Final diagnoses were recorded once follow up was complete. Where there was uncertainty in the final diagnosis this was resolved by independent adjudication between the principal investigator and the lead researcher; clinicians both trained in infectious diseases. The same research samples were taken from healthy volunteers and telephone follow up occurred one week after recruitment. Those who developed infective symptoms at follow up were excluded. Retrospective mNGS of EDTA whole blood and serum samples from twenty-five AUFI participants and five healthy volunteers was performed.

Results

Over half of participants with AUFI did not receive a clinically credible diagnosis, 52% (52/100). Of those with a confirmed diagnosis, infectious causes predominated (viral 27% (27/100), bacterial 18%; (18/100)), with few non-infectious diagnoses made (3%; 3/100). Extensive microbiological and radiological investigations were performed, but only a small proportion (5.0%; 70/1401) contributed to the final diagnosis. Empirical antimicrobial use was common (81%; 81/100). Antimicrobials were prescribed in 84.6% (44/52) of those without a diagnosis supported by standard diagnostic testing and 77.1% (37/48) of those with a diagnosis, (OR 0.61; 95% CI 0.21-1.77; *p*=0.45). The median [IQR] antimicrobial duration was 6.0 [1.0-12.0] days across the entire AUFI cohort, 5.6 [1.6 to 9.9] days in those without a diagnosis and 5.0 [0.1 to 12.1] days in those with a diagnosis (difference 0.63, 95% CI -2.46-1.96; *p*=0.99). The median [IQR] length of stay for the whole cohort was 2.0 [1.0-4.1] days. Undiagnosed participants had significantly longer hospital stays than those with a diagnosis (median 2.9 IQR [1.6-4.9] days versus 1.7 [0.8-1.6] days; difference of 1.2 days (95%CI 0.04 to 1.66); *p*=0.036). There was a trend for undiagnosed participants having on-going symptoms at follow up (undiagnosed 59.6% (31/52) versus diagnosed 39.6% (19/48)(OR 0.44, 95% CI 0.2-4.9; *p*=0.07) although this was not statistically significant. In this study, mNGS was not able to confirm microbiological diagnoses made by standard of care diagnostics or provide additional diagnostic information for undiagnosed participants.

Conclusions

This is the first proof of concept prospective study of adults presenting to a UK hospital with AUFI. Accepting of its small size, this study highlights the high proportion of participants who remain undiagnosed despite extensive investigation, that the majority of participants received antimicrobials and the use of broad-spectrum agents was common. Infection was the predominant cause of AUFI with a wide range of pathogens detected. Despite a growing body of evidence demonstrating mNGS can contribute to the diagnosis of infection, mNGS was not proven to be a useful diagnostic tool in this study, however a number of study limitation may have contributed to this finding. The wide range of pathogens identified and limited contribution of standard diagnostic testing suggests a broad, untargeted approach to infection diagnostics should be sought. Large, multi-centre studies should be considered to further characterise the burden of AUFI in the UK and to further explore the use of untargeted diagnostic technologies such as mNGS to improve the diagnosis.

Table of Contents

Table of Tables

Table of Figures

Research Thesis: Declaration of Authorship

Print name: Rebecca Louise Houghton

Title of thesis: Acute Undifferentiated Febrile Illness in the UK and Investigating the Diagnostic Utility of Metagenomic Next Generation Sequencing

I declare that this thesis and the work presented in it are my own and have been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before

Signature: Date: 12/11/2022

Acknowledgements

I would like to express my gratitude to my primary supervisors Dr Tristan Clark and Dr David Cleary for their consistent support, scientific insight and guidance throughout this piece of work. To Dr Gemma Kay and Professor Justin O'Grady for collaborating and sharing their expertise for the mNGS analysis in this study. To the wider clinical, laboratory and research teams for your hard work, support and guidance throughout this project. To my family, friends and colleagues- particularly my good colleagues at Hampshire Hospitals, I would like to extend my sincere thanks for your prolonged support, flexibility and understanding.

Definitions and Abbreviations

Ab	Antibody
AIDs	Acquired immune deficiency syndrome
AE	Adverse Event
Ag	Antigen
AMU	Acute Medical Unit
ASO	Anti-streptolysin titre
AST	Antimicrobial sensitivity testing
AUFI	Acute undifferentiated febrile illness
AUROC	Area under receiver operating curve
Bp	base pairs
BHI	Brain Heart Infusion Broth
CFU	Colony Forming Unit
CI	Confidence Interval
CLED	Cystine-Lactose-Electrolyte Deficient agar
CRP	C reactive protein
CRF	Case report form
CT	Computed tomography
DNA	Deoxyribonucleic Acid
CSF	Cerebral spinal fluid
ddNTP	Di-deoxynucleartriphosphates
dsDNA	Double Stranded Deoxyribonucleic acid

DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
EBV	Epstein Barr virus
E. coli	*Escherichia coli*
ED	Emergency Department
EDTA	Ethylenediaminetetraacetic acid
FUO	Fever of unknown origin
FBC	Full blood count
FWF	Fever without focus
gDNA	Genomic Deoxyribonucleic acid
HIV	Human immunodeficiency virus
HSV	Herpes simplex virus
ICU	Intensive care unit
Kbp	Kilobase pairs
LFD	Lateral flow devices
LMIC	Low and middle income countries
MALDI-TOF	Matrix assisted laser desorption/ionisation time of flight
MERS	Middle Eastern Respiratory Syndrome
MHRA	Medicines and Healthcare Products Regulatory Agency
MLST	Multi-locus sequence typing
mNGS	Metagenomic next generation sequencing
MRI	Magnetic resonance imaging
NAATs	Nucleic Acid Amplification Tests
NHS	National Health Service

NICE	National Institute of Clinical Excellence
NGS	Next Generation Sequencing
OD	Optical Density
PBS	Phosphate buffered saline
PDC	Potential diagnostic clue
PCR	Polymerase chain reaction
PPD	Purified protein derivative
rDNA	Ribosomal deoxynucleic acid
qPCR	Quantitative Polymerase chain reaction
REC	Regional Ethics Committee
RNA	Ribonucleic acid
RPM	Revolutions per minute
SAB	Sabouraud dextrose
SAE	Serious Adverse Event
SARS	Severe Acute Respiratory Syndrome
SBI	Significant bacterial infection
SFI	Serious focal infection
SMR	Standardised mortality rates
S.pyogenes	*Streptococcus pyogenes*
ssDNA	Single stranded DNA
S.aureus	*Staphylococcus aureus*
SePSI	Sequencing in Patients with Suspected Infection Study
TB	*Mycobacterium tuberculosis*
TE	Tris- Ethylenediaminetetraacetic acid buffer

UHS	University Hospital Southampton
UK	United Kingdom
USII	Underlying serious infection illness
USS	Ultrasound scan
VZV	Varicella zoster virus
WCC	White Cell Count
WGS	Whole Genome Sequencing

Chapter 1 Introduction

1.1 Background

1.1.1 What is acute undifferentiated febrile illness?

The aetiology of fever has been a continual source of interest and research for philosophers, physicians and scientists, despite developments in diagnosis and treatment; infection remains the leading cause of fever [1-3]. Our understanding of fever has evolved over the centuries. From Galen's second century description of fever as an independent disease state [4] to Koch and Pasteur recognition that microorganisms can cause disease [5] and Semmelweis' theory that 'cadaverous particles' transferred on the hands of medical students were linked to maternal mortality introducing chlorine hand wash as a key infection control measure [6], infection continues to contribute to much global morbidity and mortality [7-9].

The Global Burden of Disease Study in 2019 demonstrated an overall decline in the global burden of communicable diseases since 1990; however, infection driven illness still accounts for five of the twenty-five most prevalent diseases across all age groups [10]. Lower respiratory tract infections are the 4th largest contributor, followed by diarrhoeal disease (5th), HIV/AIDS (11th), Tuberculosis (12th), and malaria (14th). The greatest burden of infectious disease is in 0-9 year olds, where twelve of the twenty-five most prevalent diseases were attributable to infection. In the age range 10-24 years, infectious causes are less predominant; and this trend continues as age advances with three communicable diseases in the top 25 causes for each age group of 25-49 years, 50-74 years and over 75 years. It is notable however that diarrhoeal diseases and tuberculosis contribute to disease burden in all ages groups and that the overall second largest contributor to disease burden in the 25-49 year age group is HIV/AIDs. Lower respiratory tract infections feature in all but ages 25-49 years.

This study provides valuable insight into the global impact of infectious disease, however, potential inaccuracies and accessibility issues with the primary data may have led to errors in the calculated disease rates reported. Furthermore, 'lower respiratory tract infection'; 'diarrhoeal diseases' and the 'other unspecified infections' do not list a microbiologically proven diagnosis. This may lead to misclassification, for example, an acute presentation of inflammatory bowel disease can be indistinguishable from infective colitis without appropriate diagnostic testing. Equally, a diverse range of pathogens, each with varying management and severity, causes

respiratory tract infections. Despite these limitations, this study clearly demonstrates the significant contribution of infection to the global burden of disease.

Despite having an armoury of diagnostic tests and therapeutics to manage infection, identification of the underlying pathogen often still eludes treating physicians [11-13]. Increasing antimicrobial resistance (AMR) makes a microbiologically proven diagnosis even more important as with increasing AMR there is a decreasing chance an empirical antimicrobial will be effective against a suspected infection [14]. Furthermore, failure to diagnose the cause of an acute febrile illness may risk transmission of an infectious disease to others and a lost opportunity to apply timely public health control measures [15] as exemplified by the rapid transmission seen in the current SARS-CoV-2 pandemic. Rapid, accurate diagnostic testing for infectious disease facilitates appropriate management of patients, avoids unnecessary antimicrobial use and may identify novel and re-emerging pathogens in time to interrupt the chain of transmission.

1.1.1.1 Literature review of acute undifferentiated febrile illness

A literature search of EMBASE and MEDLINE databases was performed from the year 1974 (beginning of available records) until 30th October 2015 and set to provide automated updates via email whenever a new study was available which met the search criteria. The broad search terms 'infection' or 'fever' and 'unexplained' or 'undiagnosed' were used and limited to studies in humans published in the English language. A total of 3294 articles were returned from the initial search and 3149 articles excluded as were not relevant to the search criteria. After de-duplication a total of 132 articles were included in this narrative summary of the literature review. The reference list of these 132 articles were reviewed for suitable articles missed by the literature search and included if meeting the original search criteria.

Acute undifferentiated febrile illness (AUFI) is described in the literature as an acute febrile illness without localising signs or symptoms suggestive of the underlying cause [13, 16-25]. 'Non-malarial fever' [26-29], 'severe febrile illness' [30] and 'unexplained fever' [31, 32] are also terms used to describe this cohort of patients and will be included in this narrative summary. Unlike a fever of unknown origin (FUO) which was originally defined by Petersdorf and Beeson as a fever duration exceeding 21 days [33], patients with AUFI are typically febrile for less than 21 days at the time of presentation. However, in the wider AUFI literature where fever duration is stated, there is much variation in the definition used; 5-21 days [34], < 15 days [21, 35], 5-15 days [22], <14 days [25, 36-38], 2-14 days [39], 3-14 days [40, 41], < 10 days [12], < 8 days , 2-14 days [39], 3-14 days [40, 41], < 10 days [12], < 8 days [42], < 7 days [17], > 3 days [43], within 2 days [44]. For the purposes of this review, a fever duration of 21 days or less incorporates the wider AUFI

literature and provides some resilience against the broad geographic spread of studies where presentation may be delayed due to issues accessing or affording healthcare [9, 38, 45].

By definition, both FUO and AUFI remain undiagnosed despite initial investigation and physician review [32, 46]. Some of those who present with AUFI will continue to be febrile for over 21 days and subsequently fit the case definition for FUO. The need to distinguish between the two lies in the differing aetiology of fever in each group. In AUFI where a cause is identified, this is most likely to be attributable to infection [1, 11, 19], whereas the majority of diagnosed FUO is attributed to systemic inflammatory conditions or malignancy [47, 48]. AUFI presents with an initial febrile illness, which may be accompanied by non-specific signs and symptom. The most frequently reported symptoms associated with AUFI are illustrated in table 1.1.

Table 1.1 AUFI symptoms described in the literature (n=9)

Symptom Reported	Median	Range	References
Headache	69.0%	6.4-88.5%	[12, 13, 17, 20, 24, 25, 35, 37, 38]
Disappetite	67.5%	54.5-80.4%	[13, 38]
Myalgia	64.6%	44.3-79.3%	[13, 17, 20, 25, 37, 38]
Chills	48.1%	35.6-65.6%	[12, 20, 35, 37]
Malaise	46.5%	-	[12]
Sore throat	46.3%	17.2-89.5%	[12, 13, 17, 24, 35, 38]
Cough	42.4%	26.3-58%	[12, 13, 17, 24, 35, 38]
Abdominal pain	27.1%	10.4-61.0%	[13, 17, 20, 38]
Rhinorrhoea	25.8%	21.7-35.8%	[24, 38]
Backache	24.6%	21.2-28%	[13, 38]
Nausea	23.6%	18.1-52.0%	[13, 17, 20, 37, 38]
Vomiting	22.3%	9.1-53.0%	[13, 17, 20, 24, 37, 38]
Arthralgia	17.8%	15.8-58%	[13, 20, 25, 38]
Eye pain/ conjunctivitis	17.4%	6.7-28.0%	[13, 38]
Diarrhoea	15.6%	12.3-18.9%	[13, 24]
Earache	13.6%	-	[24]
Coryza	11.2%	-	[17]
Breathlessness	8.0%	1.1-14.8%	[13, 37]
Rash	4.5%	2.9-23.1%	[13, 17, 25, 38]
Chest pain	1.9%	-	[13]

Only four studies report on clinical signs detected in patients with AUFI and these are detailed in table 1.2.

Table 1.2 Clinical signs reported in AUFI studies (n=4)

Signs Reported	Median	Range	References
Dehydration	7.1%	-	[38]
Lymphadenopathy	4.0%	2.1-5.3%	[13, 17, 38]
Jaundice	3.3%	1.6-19.0%	[17, 35, 38]
Eschar	3.2%	-	[13]
Haemorrhage	2.4%	1.1-3.3%	[13, 17, 38]
Hepatomegaly	1.0%	0.8-4.3%	[13, 17, 38]
Splenomegaly	0.5%	0.5-0.5%	[13, 38]

Screening blood tests may provide clues to a diagnosis but are very rarely conclusive. For example, a fever, high white cell count, raised C-reactive protein (CRP) and deranged liver function tests may be seen in a patient with a pyogenic liver abscess [49]. Equally, this picture is also observed in leptospirosis [50] but the two conditions require very different management and follow up. Whilst a pyogenic liver abscess is most likely to be secondary to endogenous spread from the gastrointestinal tract (GI tract), leptospirosis may warrant investigation into the environmental exposure, which led to infection, highlighting the need for a microbiologically proven diagnosis. Where available, details on blood test findings in AUFI are presented in table 1.3.

Table 1.3 Blood test abnormalities in AUFI reported in the literature (n=6)

Blood test abnormalities	Median	Range	References
Raised CRP	50.0%	50.0-89.8%	[25, 51]
Thrombocytopenia	44.1%	27.3- 0%	[13, 17, 35, 51]
Anaemia	35.0%	-	[20]
Elevated liver transaminases	34.6%	25.0-53.9%	[13, 25, 35, 51]
Elevated WCC	28.0%	9.3-35.0%	[13, 17, 35, 51]
Low WCC	20.0%	13.2-55.0%	[13, 20, 25]
Deranged renal function	18.0%	10.0%-26.0%	[35, 51]

No published reports of a United Kingdom (UK) population based study evaluating undifferentiated febrile illness were identified (up until January 1st 2022). Of the 46 available relevant studies specifically of AUFI identified outside the UK, the most studied region was India (15), followed by Cambodia (4, including 1 modelling study based on data from both countries), Thailand (4), Nepal (3), Laos (2, including 1 modelling study), Tanzania (2), Vietnam (2), tropical/subtropical (2), Southern Spain (1), Colombia (1), Antilles (1), Brazil (1), Afghanistan UK military (1), South East Asia (1), Korea (1), Northern Australia (1), China (1), Vietnam (1), Pakistan

(1), Ecuador (1), Indonesia (1).

Generalising the findings from these studies to UK based AUFI is problematic for several reasons. The geographical variance in the aetiology of fever or 'spatial heterogeneity' between the studied sites and the UK is an important factor. Prasad *et al* noted marked geographic variation in the range of pathogens responsible for fever in their review of 45 fever studies in low and middle-income countries (LMIC) [30]. However, no studies were included from Europe, Southern and Middle Africa, Eastern Asia, Oceania, Latin American and Caribbean regions and so these results cannot be widely generalized. The World Health Organisation (WHO) report on fever management in peripheral settings acknowledges spatial heterogeneity of infectious disease as an area requiring further research. This research is needed to inform the appropriate provision of diagnostic testing and treatment which is relevant to the infections an individual may have encountered [2].

Many factors contribute to the spatial heterogeneity of infectious disease. Climate and seasonal changes are more notable in tropical and sub-tropical regions and may influence the prevalence of a given infection at a given time, for example, infections with Dengue virus increase during monsoon season due to an abundance of fresh water breeding sites for its mosquito vector [34, 52, 53]. Additionally, insect vectors and animals reservoirs of infection will differ across different regions, as will the occupations of the population studied. For example, the percentage of agricultural workers in India was 43% in 2019 in contrast to 1.5% of the UK populations in the same year [54, 55]. This difference may contribute to the high number of zoonotic infections such as leptospirosis (range 0.14%-39.1%) and rickettsial infection (range 2.9-40.8%) observed in AUFI studies in countries where agricultural work is common [12, 13, 17, 20-22, 34]. Unlike the UK, many countries do not offer healthcare services which are free at the point of service [45]. Lack of a free health service may reduce the likelihood those with fever will access healthcare, be able to afford diagnostic testing or return for follow up. The impact of self-funded healthcare has been noted as a limitation in a number of AUFI studies [9, 22, 38]. Differences in vaccination availability and uptake will also have an impact on vaccine preventable illness [45].

No UK based AUFI studies were identified from the literature search, however, there are a number of large European paediatric infection studies, which include UK hospitals and a UK based ICU study of 'undiagnosed serious infectious illness' USII in adults. These studies provide some insight into the cause of fever in European children and severe infectious disease in the UK adult ICU setting.

The EUCLIDS (European childhood life threatening infectious disease study) took place over four years in nine European countries and one hospital in The Gambia and included 194 hospitals [56].

The data presented here is restricted to that collected at European sites. Children aged one month to 18 years old with sepsis or severe focal infection were recruited. A total of 2844 participants were recruited and had a complete dataset for analysis. Of these 53.1% were male and the median age was 39.1 months. Standard diagnostic testing was performed at the discretion of the treating team. Over half (56.8%) of participants presented with serious focal infection (SFI) and the remaining with sepsis (43.2%). Of the SFI group the main diagnoses were pneumonia (18%), central nervous system (CNS) infection (16.5%) and skin and soft tissue (8.7%). Causative organisms were identified in 47.8%, leaving 52.2% without a confirmed pathogen. The most commonly identified pathogens were *Neisseria meningitidis* (9.1%), *Staphylococcus aureus* (7.8%), *Streptococcus pneumoniae* (7.7%), *Streptococcus pyogenes* (5.7%) and Gram-negative bacilli (11%). Viral infection accounted for 6.5% with Enterovirus, Rhinovirus and Respiratory syncytial virus (RSV) being the most commonly identified.

Admission to paediatric intensive care (PICU) was required for 37.6% (1070/2844) and of these, 41.6% had no causative organism identified. The mortality in this study was 2.2% (57/2844) and higher in the sepsis than the SFI group (4.9% and 0.5% respectively). Of those who died 33.3% had no causative organism identified. Interestingly despite 40% of individuals receiving antibiotics prior to blood cultures being obtained, no correlation was found between antibiotics given prior to blood cultures being taken and organism identification.

The main limitation of this study is that patients with *Neisseria meningitidis*, *Staphylococcus aureus, Streptococcus pneumoniae* and *Streptococcus pyogenes* could be recruited once the organism had been identified if the participant had not already been recruited. Therefore these infections may be overrepresented as causes of sepsis and SFI. No data was provided on type and duration of antimicrobials or the total proportion of patients who received antimicrobials and the study focused on sepsis and SFI rather than AUFI. Diagnostic tests varied between sites and there was no minimal diagnostic set of investigations performed which may have limited the opportunity for a diagnosis to be made. Additionally, follow up did not extend beyond discharge so post discharge mortality and morbidity may have been missed.

This large, prospective, multicentre study is the first of its kind to examine life threatening infectious diseases in paediatrics. Although not a specifically focussing on AUFI it demonstrates the wide range of pathogens causing sepsis and SFI in children in Europe and provides detailed outcome data demonstrating the high reliance on PICU and high mortality in those needing critical care support. It is notably that even in this well-defined group of sepsis or SFI still over half of participants remained undiagnosed [56].

The PERFORM study (Personalised risk assessment in febrile illness to optimise real-life

management across the European Union) is a large on-going paediatric study of febrile illness. Data from this study has contributed to a number of sub studies. One of these studies is the MOFICHE study (Management and outcome of febrile children in Europe) which the next two studies are part of.

In 2021 Hagedoorn *et al* developed and validated a reliable clinical prediction tool to determine children with a high risk of bacterial infection across five different hospital sites. This study included 16,268 children aged 0-18 years presenting to the emergency department (ED) with a fever greater or equal to 38^{o}C in the preceding 72 hours. Twelve hospitals were included in eight different countries including three sites in the UK. Standard investigations were performed as part of clinical care.

Upper respiratory tract infection (URTI) was the most common diagnosis (45%), followed by lower respiratory tract infection (LRT) 18%, gastrointestinal infection (14%), AUFI (9%), and invasive bacterial infection (IBI) (0.8%). Of IBI 119 were bacteraemias, 15 meningitis and 9 bone and joint. The organisms detected in this study included *Streptococcus pneumoniae* (21%), *Staphylococcus aureus* (19%), *E.coli* (10%), *Neisseria meningitidis* (7%), *Coagulase negative staphylococci* (7%), *Kingella kingae* (5%), *Group B Streptococcus* (4%), *Group A Streptococcus* (4%), *Salmonella species* (4 %) and other (4%). Of those with IBI, 93.3% received antimicrobials at their first ED attendance and 100% during their admission. ICU admission was required in 7.4% of those with IBI and 0.8% of non-IBI participants. This study once again highlights the broad range of pathogens responsible for fever and although the proportion of participants with AUFI was fairly low (9%), it is unclear what proportion of participants were diagnosed on clinical presentation without a causative pathogen being identified. Antibiotic and outcome data was not presented for the whole cohort and there was no standardised set of diagnostic investigations. Coagulase negative Staphylococci were identified in 7%, however no reference is made to the clinical significance. This group of organisms commonly contaminate blood cultures and so the clinical significance of this finding should be questioned. This study was not designed to explore AUFI specifically. However it is a large, multicentre study, which provides insight into the causative pathogens and clinical syndromes responsible for paediatric fevers presenting to ED across Europe [57].

Borensztajn *et al* also used data from the MOFICHE study to examine the cause of fever in adolescents presenting to ED across Europe. The study included 12 hospitals across eight different countries and compared the presentation of adolescents (12-18 years) with younger children. Of 38,120 participants, 25.4% were admitted to hospital, 9.9% received intravenous antimicrobials and 0.4% were admitted to ICU. The most common diagnoses were URTI 52.3%, LRTI 14.7%, gastrointestinal infection (10.4%), fever without focus (7.7%). Of these diagnoses

made, 56.3% were presumed viral infections, 22.3% were presumed bacterial infection, 15.3% were unknown and 6% were 'other'. Details of type and duration of antimicrobial, outcome and duration of admission data were not included. This study found that adolescents were more often triaged as highly urgent than younger children (29.6% versus 24.5%, p<0.001) were more likely to present with focal neurological (1% versus 0%) and meningeal signs (1% versus 0%) and had a higher rate of serious bacterial infection SBI, (15.8% versus 8.4%) despite accounting for a low proportion of ED attendances overall [58].

These large multi-centre paediatric studies demonstrate the wide range of causative pathogens responsible for paediatric fever, that ICU admission rates and mortality can be high in sepsis and SFI, that not all individuals receive a diagnosis and importantly, the aetiology of fever changes as a child gets older. These studies further support the need for large, well-designed, multi-centre, prospective studies of the aetiology of fever in the UK and Europe as this data is not currently available.

The UK study was a survey of 19 adult and paediatric intensive care units (ICU) in the UK and took place over the 18-month period surrounding the 2012 Olympic Games [59]. The aim was to identify undiagnosed serious infectious illness (USII) in patients requiring ICU support and evaluate a surveillance model designed to detect a deliberate release of an infectious agent. Although it did not directly examine patients with AUFI it gives some insight into severe infectious illnesses presenting to ICUs in the absence of UK based AUFI studies.

Twenty-two USII's were identified during the study period, 19 of which were in adults and 10 of these were associated with travel outside the UK in the preceding six months. The predominant presenting syndromes consisted of "respiratory", "presumed bacterial sepsis", "neurological" and "cardiac syndromes". A 45.5% (10/22) mortality was noted in patients with USII. A diagnosis was achieved in 35.3% (12/ 34), and the majority of diagnoses were attributable to infection 83.3% (10/12). Two cases were due to invasive fungal infection (histoplasmosis and aspergillosis), and the remaining cases were attributable to disseminated mycobacterium tuberculosis (TB), leptospirosis, *Haemophilus specie*s respiratory infection, pneumococcal sepsis and blood stream infections with drug resistant Amp C β-lactamase producing *Escherichia coli* (*E. coli*), *Pseudomonas*, *Staphylococcus* and *Enterococcus* species.

The key limitations of this study were that it was restricted to the ICU setting, outcome data was not broken down by diagnosis, no detail was provided on fever duration and insufficient detail of the laboratory confirmed diagnoses was provided. For example, '*Staphylococcus species'* does not distinguished between *Staphylococcus aureus* (*S.aureus*), which is far more likely to explain a severe clinical illness and *Staphylococcus epidermidis* which is may be a result of contamination of

the blood culture with skin flora, or associated with a line infection. In addition, all units were selected for their geographical proximity to the 2012 East London Olympic games site with 13 of 19 sites being in London, which may have skewed the range and severity of disease encountered. In 2000-2008 period, East London was found to have high levels of socioeconomic deprivation and an increased standardised mortality ratio (SMR) for non-infectious disease, this is in contrast to South West London, which had low levels of socioeconomic deprivation, and lower SMR for non-infectious disease [60]. With such variation within the city, it is difficult to extrapolate this data to the wider UK population. Despite the limitations, this study does give some insight into the burden of severe infection presenting in a range of UK based intensive care units. It highlights the wide range of causative pathogens identified from patients with USII, that the majority (63.6%; (14/22) remain undiagnosed and that the combined mortality of diagnosed and undiagnosed cases is high (45.5%; 10/22).

1.1.2 Why is acute undifferentiated febrile illness important?

1.1.2.1 Burden and impact of acute undifferentiated febrile illness

The global burden of AUFI of infectious aetiology is considerable [17, 46, 59, 61]. In a large proportion of AUFI cases, a laboratory confirmed diagnosis cannot be made (median 39.6%, IQR 33.9-47.2%) although there is variability in the reported literature (range 3.2- 65.0%) [13, 17, 20-22, 24-26, 32, 34-37, 39-41, 44, 52, 62-67]. The clinical course of patients with acute fever may range from a mild, self-limiting illness to rapid deterioration and death [13, 20, 30, 46, 59] and reliably predicting those likely to deteriorate is extremely challenging [46]. The wider impact on the population and the socioeconomic burden of AUFI has not been well studied.

No studies were identified which directly examine the socioeconomic burden of AUFI. However, a study by Radhi *et al* demonstrated paediatric patients admitted with fever had significantly longer hospital stays (mean 2.87 days versus 2.3 days respectively, 95%CI –0.76 to -0.19 p=0.001) and febrile children had more blood tests performed than afebrile children (125 tests per 206 patients versus 99 tests per 562 patients respectively) [68]. D'Acremont *et al* demonstrated in a study of Tanzanian children with AUFI that 25,743 diagnostic tests were performed, yielding only 1,232 diagnoses - an average of 20.9 diagnostic tests per diagnosis, or 4.7% of tests contributing to the diagnosis, highlighting the inefficiency and expense of the current diagnostic approaches [65].

Outcome data is infrequently reported in AUFI studies but where available, mortality ranges from 0-26% (median 6.2%, IQR 17.3-10.8%) [8, 9, 11, 13, 20, 34, 40, 43, 44, 52, 64] and admission to ICU is required in 4.5-24.5% (median 8.8%, IQR 5.3-11.2%) of patients [11, 34, 41, 64]. One large Cambodian study examined 1,225 febrile episodes in children and found 1.6% of patients were

discharged against medical advice, 0.2% required palliative care and transfer for further care was needed in 0.7% of patients [44].

Non- AUFI studies of sepsis and infection highlight the impact of infection without a microbiologically confirmed diagnosis. For example, Garnacho-Montero *et al* demonstrated that receiving inappropriate antimicrobials was associated with a higher 28 day, 60 day and in-hospital mortality (OR, 8.1; 95% CI, 2.0-33.5) in 460 severely septic non-surgical ICU patients [69]. Similarly, Nishiguchi *et al* found patients with infective endocarditis (IE) who received empiric antimicrobials prior to blood cultures had a delayed diagnosis (median 24; [IQR 13.0-37.3] versus median 20; [IQR 9.0-30.5]), longer hospital stays (median 47 [IQR 36.0-53.3] versus median 40; [IQR 35.3-46] days) and a higher mortality rate (30.0%; (6/20) versus 11.0%; (2/11)) than those who did not have antibiotics in the community. This study was limited by its small size, single centre and that confounding factors limit the assessment of true causality between a delayed diagnosis and increased mortality [70]. Overall, these studies highlight the possible patient impact of a delayed microbiologically confirmed diagnosis and how diagnostic tests, which rely on live organism being cultured, may impact on patient outcomes.

HIV can present as an AUFI either in the initial seroconversion illness or, in advanced HIV due to an opportunistic infection. All patients presenting with AUFI should be offered a HIV test. If positive, opportunistic infections must be considered and patients supported to access specialist HIV care and treatment. High rates of HIV are seen in the AUFI literature [10]. A large systematic review of severe febrile illness studies in LMIC reported 21.2% (1,988/9,365) of patients tested for HIV, had a positive test. This study also found HIV positive patients had a significantly increased risk of blood stream infection 59.4% (1,667/2,805) versus 52.8% (1,357/2,566)(OR 1.3, 95% CI = 1.2–1.5, p<0.0001)[30]. Roberts *et al* found 7.6% (23/301) of febrile travellers returning to the UK were HIV positive [71]. Abrahamsen *et al* reported HIV rates of 1% (1/100) in Northern Indian study of AUFI [8] and Crump *et al* found 12.2% (57/467) HIV positivity in infants and children and 39% (157/403) in adults and adolescents in Northern Tanzania [29]. These studies demonstrate the burden of HIV and the importance of HIV testing in AUFI.

The studies described in this section share some similar findings, demonstrating the diverse range of causes or AUFI, the potential for high mortality and need for intensive care support and the high proportion of patients who remain without a laboratory confirmed diagnosis. The impact of undiagnosed AUFI on hospital admission rates, antimicrobial usage, long-term morbidity and transmission of infectious disease to others is not well understood. What is apparent is the current AUFI literature is less relevant to a UK population. There are marked knowledge gaps in understanding the key causative pathogens, antimicrobial use and clinical outcomes including

mortality rates in AUFI in the UK. There is a need for well-designed studies of AUFI in the UK to better understand these important factors [3, 7, 30].

1.1.2.2 Diagnostic challenges of acute undifferentiated febrile Illness

There are no widely accepted guidelines for the investigation and diagnosis of AUFI however, there is some overlap in the approach described in the literature. A thorough history and examination should be performed. Specific risk factors for infection should be identified, including travel abroad, any unwell contacts, vaccination history, exposure to animals, rural areas, fresh water, insect or animal bites and new or unusual foods including unpasteurised dairy products [11]. Baseline investigations should include a full blood count (FBC) [11, 17, 19], renal function [11, 19], liver function, [11, 19], blood cultures [11, 26, 30] and urine analysis [19].

Pathogen directed investigations (such as blood films for malaria), should be guided by risk factors identified in the history, clinical signs and the results of baseline investigations [30]. The AUFI literature supports diagnostic testing for malaria [17, 19, 30], mycobacterial cultures [30], nasopharyngeal swabs [26, 30], pathogen specific serology [11, 17, 26, 30], cerebral spinal fluid analysis [11], tissue biopsies where indicated [30], chest radiography [11, 19], abdominal ultrasonography [11, 19], PCR for specific pathogens such as Dengue, Chikungunya and HIV testing [30]. Although with the exception of the returning traveller study [71], none of this data is from a UK setting and therefore, cannot be widely extrapolated.

Co-detection of potential pathogens is reported in 15.6% (range 1.9-33.8%) [12, 13, 24, 25, 65, 66, 72] of individuals studied, presenting the challenge to understand which organisms are disease causing and which are not [39]. For example, in a study of severe febrile illness in Northern Tanzania Crump *et al* identified malarial parasites in 60.7% (528/870) of individuals with AUFI but recognised malaria was only responsible for 1.4% (14/870) of clinically significant fever, suggesting a large proportion of malaria parasites identified in this study were associated with subclinical infection in patients with immunity [29].

Access to diagnostic testing can be challenging for patients with AUFI; Mittal *et al* noted limited access to viral diagnostics in Northern India [22], Manock *et al* had a 'limited diagnostic panel' in their study at the Amazon basin [17], Nhiem Le-Viet *et al* had access to Rickettsial serology only in Vietnam [13] and Mørch *et al* noted the 'low accuracy' of diagnostic testing in AUFI in India prompting caution in their interpretation [39]. In the UK, there are clear standards for the quality and interpretation of diagnostic tests performed in clinical laboratories. However, there are no standard set of investigations which must be provided by each laboratory and so there is much variability in the tests offered between laboratories across the UK even within the National Health

Service (NHS).

Thangarasu *et al* attempted to protocolise the diagnosis and management of fever in a pilot study of 342 adults presenting with AUFI to the emergency department (ED) in Chennai. Those with sepsis 1.8% (6/342) or localising signs 37.1% (127/342) were excluded leaving 209 patients with AUFI included in the study. The main objective was to reduce unnecessary antimicrobials and reduce the number of diagnostic tests performed where fever was self-limiting, however, there are some limitations of the study and incomplete follow up data is presented in the published manuscript [43]. In this study, patients who presented on day 1 or 2 of fever had no investigations or antimicrobials but returned on day 3, patients presenting on day 3 or 4 of fever had a FBC, malaria screen and urinalysis, and those presenting on day 5 had a blood cultures in addition. Thirty-day telephone follow up was completed. Of 342 patients, the majority (54.1%; 113/209) presented on the 1st or 2nd day of fever, of which, 57.5% (65/113) were treated according to protocol. In the analysis of participants who followed the protocol, fever resolved without investigation or intervention in 75.4% (49/65). Of the remaining 42.3% (48/113) who did not follow the protocol, investigations contributed to management in 25% (12/48) of cases. This study was performed at a single site and no data was included for the 46% (96/209) of patients who presented after 3 or more days of fever and 30 day follow up including mortality and hospital admission rate was not presented. There was insufficient diagnostic data presented, 17.7% (20/113) being labelled 'other specific diagnosis'. Of note, 31% (35/113) are reported to have 'UTI' or 'malaria' as their final diagnosis, conditions which may deteriorate without prompt treatment and the lack of follow up data means unintended consequences or harm associated with this protocolised approach to AUFI cannot be ruled out [43].

Symptomatic crossover between infective and non-infective aetiologies further complicates the diagnostic challenges of AUFI. In a study of UK students with acute Epstein barr virus (EBV) infection, cervical lymphadenopathy was identified in 88% (50/57) of cases [73]. However, cervical lymphadenopathy is also a well-documented feature of lymphoproliferative disorders such as lymphoma and of *Mycobacterium tuberculosis* (TB) infection [74-76] and all may share common features such as fever, night sweats and deranged inflammatory markers. TB and lymphoma have also been reported to co-occur in the same individual at the same time resulting in rapid deterioration and death, this highlights the need for unbiased diagnostic approaches capable of identifying multiple disease processes [77].

1.1.2.3 Potential impact of undiagnosed infection on antimicrobial resistance

The WHO recognise AMR as one of the top ten threats to global health and development [78]. Antibiotic consumption in healthcare has now exceeded agricultural use in the UK [79] and the

WHO has developed a global action plan to target this problem [78, 79]. The 2019 report on AMR states that drug-resistant infection causes at least 700,000 global deaths a year and predicted this number could increase to 10 million by 2050 if no action is taken. If the rise in AMR is not halted then the economic impact is predicted to be similar to the 2008-2009 global financial crisis which could drive those in low and middle income countries further into poverty and worsen global inequality [80].

Antimicrobial resistance impacts on our ability to safely treat patients [81], provide appropriate prophylaxis for surgical procedures [82], and is associated with increased mortality [83], and longer intensive care stays [84]. In addition, antimicrobials have numerous unwanted complications such as drug interactions [85], hypersensitivity reactions [86], intravascular line infections, *Clostridiodes difficile* infection [87, 88] and development of within host antimicrobial resistance [89] which can then be transmitted to others and the environment. Schwarz *et al* identified vancomycin resistant enterococci and *amp*C beta-lactamase producing *Enterobacteriacia*e in hospital waste water, also finding the *van*A vancomycin resistance gene in biofilms found in drinking water suggesting the possible contamination of drinking water with organisms found in hospital waste water [90].

Current studies suggest between 3.8-57.2% (median 22%, IQR 17.4-49.3) of acute febrile illnesses in low and middle-income countries may be attributable to a viral infection [30, 65] and this may be an underestimate due to limited access to diagnostic testing. Suttinont *et al* studied AUFI in five rural hospitals in Thailand and found overall 10.7% (90/845) of fever was attributable to viral infection [21]. Viral infection is most commonly self-limiting and requires supportive treatment only. Despite this, a high proportion of patients with AUFI are treated with antimicrobials (median 51.5% [IQR 46.9-67.6%] range 19.2-100%) [13, 24, 25, 42]. The most commonly used agents include amoxicillin, doxycycline, cephalosporins, quinolones, macrolides and anti-malarials [13, 19, 63]. One large Vietnamese AUFI study of 762 participants identified that approximately half (49.7%; 379/762) of patients received antimicrobials and 19.2% (73/379) of these were treated with 2-3 different agents [13]. There is very little data on antimicrobial resistance in the AUFI literature; two studies report significant rates of resistance in *Salmonella* species, a group of organisms responsible for typhoid fever. Murdoch *et al* report 52.6% (61/116) resistance to nalidixic acid, an indicator for quinolone antibiotic resistance with 2.6% (3/116) of organisms displaying multi-drug resistance [67]. Chheng *et al* report 90% (20/22) of *Salmonella* isolates were resistant to ciprofloxacin and 86% (19/22) were multi-drug resistant [44].

Clinical and diagnostic uncertainty can result in hospitalised patients receiving antimicrobials simply to cover the possibility of bacterial infection [19, 91], resulting in inappropriate and

unnecessary prescriptions. A large meta-analysis of 154 studies (of which 18 were European studies), found 74.6% (22845/30623) of patients with COVID-19 were prescribed antimicrobials despite only 8.6% being judged to have bacterial co-infection, indicating a large proportion of these patients are inappropriately prescribed antimicrobials [92]. Similar rates of bacterial co-infection in COVID-19 have been reported in another study [93]. Of note, this study took place early in the pandemic when there was little data on the incidence of bacterial co-infection, which may have affected prescribing rates.

The five-year UK plan for tackling AMR clearly states the development and access to diagnostic tests in infection has a critical role in reducing unnecessary antimicrobial use. Targets are to be set by 2025 for the number of diagnoses supported by a diagnostic test [79]. Improving infectious disease diagnostics has the potential to prevent unnecessary antimicrobial prescriptions, facilitate the use of narrow spectrum antimicrobials and critically, conserve antimicrobials for when they are truly indicated, reducing some of the pressure on AMR.

1.1.2.4 Acute undifferentiated febrile illness, emerging infection and pandemics

Murchison deemed obtaining a microbiological diagnosis essential for appropriate patient and outbreak management as early as 1858. Murchison was a British physician who travelled widely and worked as an associate physician, then physician at the London Fever Hospital for ten years. There he studied fever, published 'A Treatise on the Continued Fevers of Great Britain" [94] and in 1862, was appointed Editor of the Fever Hospital's annual report for 8 years. Murchison described yearly variations of typhus, relapsing and pythogenic (typhoid) fever in England, Wales and Scotland, noting geographical and year on year differences in the prevalence of the differing pathogens. He concluded that understanding separate disease entities is essential to understand the "causes and modes of prevention" [95]. In the 511 years prior to Murchison's report in the Lancet there were two notable pandemics recorded; the Bubonic plague caused by *Yersinia pestis* in 1347 and the Smallpox pandemic in 1500. However, in the 137 years since Murchison's report, there have been ten further notable pandemics and the significant 2013 Ebola epidemic (1881 Cholera, 1918 Spanish flu influenza, 1957 Asian flu influenza, 1968 Hong Kong flu influenza, 1981 HIV/AIDS pandemic, 2003 Severe Acute Respiratory Syndrome (SARS), 2009 Swine flu influenza, 2012 Middle Eastern Respiratory Syndrome (MERS), 2013 Ebola, 2015 Zika, 2019 SARS-CoV-2) [96].

Studies predict the next emerging pathogens with the potential for pandemic spread are likely to be zoonotic, viral and originate in wildlife as we have seen with SARS-CoV-2, which is thought likely to have originated in a species of bat or pangolin [97]. Risk factors for emerging infection include densely populated areas, changes in land use and agricultural and human activities

occurring in areas with high wildlife biodiversity [15]. Some studies have identified areas of the UK as a moderate to high-risk hot spots for the emergence of a novel zoonotic infection [15, 98].

Fever is a key symptom of all infections responsible for the previous notable pandemics, with the exception of HIV where many will be asymptomatic at the time of infection [99]. Fever is prominent in SARS-CoV-2 98% [100], Ebola 72% [100] and Zika virus infection 70% [101] and the range of clinical presentations is broad [100-104]. Once a pandemic is established, all patients presenting with AUFI will be considered infected with the pandemic organism until proven otherwise in an attempt to interrupt onwards transmission. However, before a pandemic is recognised, patients remain undiagnosed and can contribute to on-going transmission. This was illustrated by Li *et al* who described the first 425 confirmed cases of SARS-CoV-2 in Wuhan, estimating an epidemic doubling time of 7.4 days and presenting clear evidence to suggest human-to-human transmission, which had previously been disputed [105]. Undiagnosed cases can facilitate healthcare associated outbreaks with associated mortality and morbidity [106]. Infected individuals, not unwell enough to seek medical attention may continue to travel, carrying the infection across geographical borders, as seen with the initial spread of SARS-CoV-2 through Europe from multiple introductions from Italy and Hubei [107].

In order to provide resilience against emerging infections, as Murchison alluded to 511 years ago, improvements in infection diagnostic testing must be made to reliably identify an infectious pathogen whether it be previously known, novel or re-emerging.

1.2 Limitations of current infectious disease diagnostics

Gaining a clinically credible microbiological diagnosis in patients with AUFI is essential for three key reasons; to guide patient management, prevent unnecessary antimicrobial use and prevent the onwards transmission of infectious diseases. Additionally, where a diagnosis of a chronic infection such as HIV is made, appropriate follow up and contact tracing can occur preventing future morbidity and mortality. In this chapter, the commonly used microbiological techniques will be discussed with regards to diagnosis in the setting of AUFI.

1.2.1 Requirement for *a priori* clinical suspicion of the likely causative pathogen

Current clinical diagnostic tests require a practitioner to suspect a particular set of pathogens before appropriate samples and diagnostic tests can be performed [108]. This approach is not robust against novel or unexpected pathogens and may lead to the true cause of infection being missed. With the exception of obtaining blood culture samples for readily cultured pathogens such as *E. coli* and *S.aureus,* which should be performed in all patients presenting with suspected sepsis [109] most diagnostic techniques are specific for a small number of pathogens guided by the clinical request and samples sent. For example, a physician may request Hepatitis A, B and C serology for a patient with an acute transaminitis but this would miss other common causes of this illness such as Hepatitis E, Epstein Barr virus (EBV) and Cytomegalovirus (CMV) unless the clinical laboratory team intervenes and broadens the range of tests performed.

Common infections acquired in the UK such as influenza or occult bacteraemia may present as an acute undifferentiated febrile illness without initially localising features [21, 26, 110]. For example, influenza accounted for 20.9% and 8.8% of AUFI respectively in studies performed in Vietnam and Cambodia [13, 24]. However, if a nasopharyngeal swab in viral transport media or a respiratory sample is not collected and influenza testing requested, the diagnosis cannot be confirmed and the patient would remain undiagnosed, may not receive the appropriate treatment and may transmit the infection to others.

1.2.2 Standard UK infection diagnostics

Commonly used diagnosed techniques for infection include microscopy, culture, molecular techniques such as polymerase chain reaction (PCR) and serological testing for an antibody response. New technology, including matrix assisted laser ionisation/desorption time of flight (MALDI-TOF) has dramatically reduced the time taken to identify organisms cultured from standard techniques and can now offer antimicrobial sensitivity testing (AST) as well as identification. Whole genome sequencing is increasingly used to confirm transmission events in

outbreaks and untargeted mNGS has a growing evidence base for identifying infections where standard diagnostics have failed.

1.2.2.1 Direct microscopy

Microscopy is widely used for the rapid diagnosis of infections including TB, malaria, enteric parasites, and genital infections such as bacterial vaginosis and trichomonas. Microscopy is cheap, uses few resources and can be performed without electricity. Microscopy can give a rapid result within minutes of the sample receipt. Dark field microscopy of syphilitic chancres for treponemes was found to be more sensitive for primary syphilis than initial EIA antibody assays [111].

Microscopy has some important drawbacks. Usually only a small proportion of the sample is examined, and so organisms in low abundance may be missed. It is operator dependent and requires appropriate training. Additionally, inappropriate storage or delay in examination of some samples may cause organisms to degrade. Microscopic techniques are usually followed with culture to improve sensitivity and provide AST. Additionally, some samples such as CSF and joint fluid are now complimented with molecular testing to increase the diagnostic yield [112].

1.2.2.2 Culture

Enriched culture is regarded as the 'gold standard' for the identification of bacterial and fungal pathogens of sterile site such as blood, CSF, urine and joint fluid [113-115]. Culture requires few resources and is cheap and widely available. If a pathogen is identified from a sterile site such as blood cultures [116], CSF [117], joint fluid [118] or ascitic fluid [119], not deemed to be due to contamination, it is likely to be clinically significant. Culture facilitates the right conditions for rapid division of organisms allowing further identification and AST. In addition to guiding patient management, antimicrobial sensitivity patterns can be used in outbreak investigation. Organisms with the same antibiogram identified in two patients linked in space and time may reflect transmission of this organism from one patient to another, however, WGS has now been demonstrated to be more accurate at determining transmission events [120].

The yield of bacterial culture is affected by the quality of the sample, the culture medium to which the organism is inoculated and atmosphere and temperature in which it is incubated. This requires experienced laboratory staff and appropriate clinical information to ensure the right conditions are provided to promote the growth of the suspected pathogen. Additionally, blood cultures may not detect organisms in low abundance unless an adequate volume of blood is taken [121]. Delay from obtaining a blood culture to incubation has been shown to significantly decrease the yield [122], and some organisms may perish before they are identified, particularly if

the sample is taken after antibiotics. These variables limit the sensitivity of culture-based methods for clinical diagnostics and the yield is generally poor [123]. Culture has been superseded by molecular testing for routine viral diagnostics and is now being replaced or enhanced by molecular techniques such as PCR and MALDI-TOF for bacterial and fungal pathogens [124] to improve the accuracy and turnaround time of results.

1.2.2.3 PCR

PCR uses short oligonucleotide primers, which have a complementary sequence to a specific DNA region of the target organism. DNA polymerases bind to the primers and replicate the region of interest which, when repeated over multiple cycles, amplifies the region of interest. The DNA is then labelled with a fluorescent dye and the quantity of DNA can be determined by the strength of the signal detected. Multiplex PCR can be used to amplify DNA sequences from multiple pathogens/genomic sites in a single sample, direct PCR is performed straight from a sample with no need for prior extraction of DNA/RNA. PCR can be used to look for specific targets such as 16S or 18S. Bacterial (16S) and fungal (18S) ribosomal structural RNA (rRNA) is more abundant within cells than genomic DNA and this approach identifies whether fungal or bacterial organisms are present within a sample. If 16S or 18S rRNA is detected, more specific PCR primers can be used to target a specific species based on the rRNA result. It is worth noting that PCR specific to a particular species is more specific and sensitive than 16S or 18S PCR but requires selection of the right primers to detect the organism(s) of interest.

A number of studies have found PCR to be comparable or even superior to culture for pathogen detection in clinical samples. Jorden analysed Ethylenediaminetetraacetic acid (EDTA) samples taken for full blood count from neonates with suspected sepsis using 16S PCR and found that 96.2% (51/53) of isolates from babies with positive blood cultures were detected by 16S PCR. Of the two missed cases (*Haemophilus influenza* and *Enterococcus species*) one EDTA sample was taken at a different time to the blood cultures so may be explained by a transient bacteraemia, and one sample had a high WCC likely resulting in high levels of human DNA, which may impact on the ability of the primers to bind to the proportionally smaller quantity of bacterial DNA [125]. Amar *et al* analysed 4,627 faecal samples from the English case-control Infectious Intestinal Disease Study and found PCR detected many more enteric pathogens than standard testing (75%; 1816/2,422 versus 53%; 1291/2,422). Interestingly, 19% (410/2,205) of controls had a pathogen identified by standard diagnostics, and 42% (926/2,205) by PCR, possibly indicating an underlying level of asymptomatic carriage [126] and demonstrated the need for clinical correlation of any diagnostic test.

PCR has a rapid turn-around time, high sensitivity and specificity and the ability to detect

organisms, which may have been killed or damaged by concurrent antimicrobials or an activated immune response. The main limitations of PCR are cost, availability of laboratory expertise, access to an appropriate laboratory space where contamination can be minimised and adequate freezer and fridge space for reagents. In addition, there must be a clinical suspicion of which organism(s) may be present in the sample and an appropriate history provided to enable the correct assay to be used. Low level/equivocal results and laboratory contamination can cause confusion and AST is restricted to gene targets coding for known resistance mechanisms. It is not possible to determine strain relatedness using commercially available PCR and further assessment such as whole genome sequencing or pulse-field gel electrophoresis must be performed in the outbreak setting [127, 128]. Microarray and multiplex PCR methods designed to diagnose bacteraemia often rely on small volume input samples running the risk of missing low grade infection or are run on positive blood culture samples which can delay diagnosis by 24-48 hours, similar to that of standard diagnostic methods [129].

1.2.2.4 Serology

Serological tests detect pathogen specific IgG and IgM antibodies and require an acute and convalescent sample to look for IgM to IgG antibody switch, confirming the acute phase IgM titre has fallen and IgG titres have risen. Serological testing is used to confirm infection from a wide range of pathogens and remains the 'gold standard' diagnostic technique in the UK Standards for Microbiology Investigation (UKSMI), including Hepatitis B [130], Hepatitis C [131], HIV [132], and syphilis [133]. Serum is obtained at presentation and 4-6 weeks following the acute illness. Serological examination is most useful for non-culturable or slow growing organisms such as *Coxiella burnetii,* Leptospira species, and histoplasma, viruses including HIV, EBV, and CMV, and parasites including schistosomiasis and strongyloides.

The most commonly used laboratory methodology is enzyme linked immunosorbent assay (ELISA), which can be fully automated. Simply, ELISA is an enzyme immunoassay technique in which micro titre plate wells (which serves as the solid-phase) are coated with an antibody (typically monoclonal) or recombinant antigen specific to the analyte of interest. The key element of the technique is a highly specific antibody-antigen interaction. A patient sample is added to the wells and if the target pathogen antigen or pathogen directed antibody is present, this will bind and be immobilised on the plate surface. Following removal of unbound molecules, this immobilised material is then complexed with a second antibody, which is linked to a reporter enzyme. The bound material is then incubated with an appropriate substrate and quantification is achieved by measuring the activity of the reporter enzyme. Typically ELISAs involve a

colorimetric reaction in which the absorbance of the coloured product is proportionate to the amount of bound antigen (Ag) or antibody (Ab) present.

Laboratory based serum analysis is relatively rapid, does not require a live organism and can be performed on historical samples providing they have been properly stored. However, it provides no data on strain relatedness or AST, will only detect pathogens specifically looked for, requires expert training and expensive equipment with stable power and temperature controls. Additionally, IgM cross reactivity with other organisms is common and can lead to diagnostic confusion [134], severely immunocompromised individuals may not mount an appropriate antibody response so may have falsely negative serology [135] and where an acute serum is non-diagnostic, the results of a convalescent sample may be too late to influence patient management.

1.2.2.1 Point of care testing

Point-of-care tests (POCT) go some way to address the need for rapid infection diagnostics. The POCT technology available ranges from highly sensitive direct from sample analysers using PCR and microarray technology, to the much cheaper and simple to use lateral flow devices. The common trade-off between the two being lower sensitivity and specificity of LFDs and the higher cost and low throughput of highly sensitive POCT analytical platforms.

There are commercially available kits for molecular platform POCT's for respiratory [136], gastrointestinal [137] and neurological infections [138] which have been found to be comparable to standard diagnostic testing and require minimal training of non-laboratory staff. POCT in respiratory virus testing has been shown to be safe, reduce the hospital length of stay and impact on antimicrobial prescribing [139]. However, they are expensive, do not provide full AST or strain typing and lack the ability to detect novel or unexpected pathogens. Sample throughput is another limitation as only a small number of samples can be run simultaneously with a run time of between 30-90 minutes per sample preventing high throughput testing.

Lateral flow devices (LFD), are now widely available for asymptomatic population screening for SARS-CoV-2 antigen [140]. Rivett *et al* demonstrated asymptomatic screening of healthcare workers for SARS-CoV-2 by PCR testing identified a 3% (31/1,032) positivity rate, although when where further investigated only 9.7% (6/31) reported no symptoms prior to or in the 7 days following the test [141]. LFDs are most sensitive at higher viral loads (lower *Ct* values on PCR, *Ct* <25) [140] and recent work by Lee *et al* has demonstrated those with lower *Ct* values (and higher viral load) are more likely to transmit infection and so identifying these cases is essential for population outbreak management [142]. It is notable that the majority of infections in this study

were due to the B.1.1.7 (α) variant of SARS-CoV-2 which was first displaced by the B.1.617.2 (δ) in 2020 and is now being displaced by the B.1.1.529 (O) variant in the UK, although it is biologically plausible that a higher viral load in B.1.617.2 (δ) and B.1.1.529 (O) infection will also result in higher infectivity.

LFDs are based on the same scientific principles as serological testing but can be used in non-laboratory environments. They can detect either antigen [143] or antibody [144]. The immunochromatography assays employed by LFDs are based on sandwich ELISA technology (where an antigen is sandwiched between two specifically designed monoclonal antibodies) and can be run on chromatography paper by capillary action. Monoclonal antibodies can be fixed in different zones across the paper. Anti-human IgG is used as a 'control zone' to ensure the test has worked appropriately then a second 'test zone' contains the target Ag or Ab of interest [145].

LFDs have been used for significant infectious diseases such as HIV [146] and Malaria [147] are low cost and easy to use. LFDs generally have a lower sensitivity and specificity than laboratory-based testing methods such as PCR, are only able to detect a single target and performance can vary between manufacturers. Peto *et al* found only 9% (6/64) of LFD's from different manufacturers had suitable characteristics to be considered for mass population testing for SARS-CoV-2 antigen and that kit failure rates were high (5.6%) [143]. Similar findings to that of Moshe *et al* in their study of LFD's for SARS-CoV-2 seroprevelance [144].

1.2.2.2 Interferon-gamma release assays and Tuberculin skin test

Both the Tuberculin skin test (TST) and Interferon-gamma release assays (IGRA) can contribute to the diagnosis of latent *Mycobacterium tuberculosis* (MTB) infection. Both methods measure the immune response driven by interferon-gamma which is released by T cells previously exposed to MTB antigen. The TST is performed by intradermal injection of purified protein derivative (PPD) and the cutaneous skin reaction measured 48-72 hours later [148]. TSTs are relatively easy and inexpensive but require a second clinic visit for the skin reaction to be read. PPD is not specific for MTB antigens and positive TSTs can be seen following Bacillus Calmette-Guerin (BCG) vaccination and infection with some non-tuberculous mycobacteria (NTMs).

IGRAs require a single blood draw and are unaffected by BCG used for vaccination or for the treatment of bladder carcinoma. Like TSTs, some NTM infections can cause a positive IGRA. The main two antigens included in IGRAs are early secreted antigenic target 6 (ESAT-6) and culture filtrate protein 10 (CFP-10). These antigens are present in MTB in high abundance but also present in *M. kansasii, M. marinum, M. flavescens*, and *M. szulgai* [149] and so infection with

these organisms, may cause a positive IGRA. IGRAs require live lymphocytes and so the timing, storage and transportation of sample is critical to ensure a valid result.

Neither TSTs or IGRAs can distinguish between active or latent infection and it is possible to have a negative test in the context of active MTB infection, particularly in immunocompromised hosts. However, both are useful screening tests for latent MTB provided they are interpreted in the correct clinical context. IGRAs have the benefit of requiring a single appointment to gain a result and are less affected by BCG administration and have been widely adopted by TB screening programmes in the UK.

1.2.3 Summary

Current standard of care infection diagnostics are useful for common, suspected pathogens providing the correct patient samples are taken in a timely manner, appropriate clinical details are provided and the necessary laboratory resource and expertise is available to process them. Depending on methods used, results can be available in a clinically relevant timeframe such that patient management can be optimised. However, current diagnostics are limited by the requirement to suspect the pathogen(s), which may be present in the sample, obtain the right sample types and for some tests, the need for laboratory expertise and specialist equipment, which can be costly.

There is a gap in the current diagnostic repertoire particularly where unexpected or novel pathogens are present. This gap could be addressed by new technology such as mNGS. Utilisation of sequence based approaches to enhance standard diagnostics could provide an untargeted analysis of samples and an opportunity to identify a causative pathogen not only in those with undiagnosed infection but those with where the underlying infecting agent is novel or unsuspected [150, 151]. If sequencing costs continue to fall and its clinical diagnostic utility is proven, mNGS may become a more accessible diagnostic tool to help address the burden of undiagnosed infection.

Due to the need for a controlled laboratory setting with robust data storage capacity, specific laboratory and analytical expertise and expensive consumables, it is unlikely that mNGS will completely replace standard diagnostic testing or provide a global solution to the burden of undiagnosed infection. However, one centralised laboratory, the Manipal Centre for Virus Research in Southern India has begun providing comprehensive infection diagnostics including mNGS over a wide geographic area. Set up in a research and surveillance capacity in collaboration with the Centres for Disease Control and Prevention (CDC), cases are identified by clinicians at local hospitals before samples of urine, serum, nasopharyngeal swabs blood culture and stool are

collected for analysis. A dedicated logistical team co-ordinate the transport of samples to the laboratory from an area covering 33 hospitals across the 3.2 million square kilometre country. Results are available within 24 hours and so could conceivably impact on patient management [66]. This approach demonstrates that comprehensive and expensive infection diagnostics requiring significant expertise to perform and interpret them can become accessible if the right infrastructure is put in place.

1.3 DNA sequencing

There is a clear need for improved diagnostics in AUFI and growing evidence for the utility of untargeted mNGS to diagnose infection where standard laboratory and clinical approaches have failed. As sequencing methods become more accessible and affordable their role in the real-time diagnosis of patients presenting to hospital with AUFI needs to be further explored. This section will outline current sequencing techniques relevant to infection and explore their use for improved pathogen detection and analysis.

1.3.1 Sanger sequencing

The first genome sequencing methods were developed in 1977 by Frederick Sanger [152]. In Sanger's chain termination method a single stranded DNA fragment is sequenced using radioactively or fluorescently labelled deoxynucleotide triphosphates (ddNTPs). The sample DNA fragments are split into four reactions, each with a DNA polymerase, standard deoxynucleotide triphosphates (A, T, C, G) and one of four labelled ddNTPs (A, T, C or G). The ddNTPs lack a 3'OH group, which leads to sequence termination once incorporated into the DNA complementary strand. The resultant double stranded DNA fragments from each reaction are heat denatured and run though polyacrylamide-urea gel electrophoresis in four lanes. This separates them by fragment size and the DNA sequence can be read across the lanes using ultraviolet light. Sanger and colleagues further developed this method with 'shot-gun' sequencing. Shot-gun sequencing allows sequencing of larger genomes using the chain termination method, it is faster and less expensive however, it requires a large amount of computational power to re-align the reads and is prone to error in repetitive genomes due to misplacement of similar sequence reads. In shotgun sequencing, large genomes are first broken into smaller fragments then individually sequenced. The smaller sequence reads are then computationally reassembled using overlapping regions of the sequence reads to assemble the target genome. Shotgun sequencing is most efficient if there is a known reference genome making it less useful for de-novo assembly for pathogen discovery purposes.

Sanger and colleagues were the first to sequence the first complete genome of bacteriophage ϕX174 in 1977 which has a 5,375 base pair (bp) single stranded DNA genome (ssDNA) [153] and then bacteriophage λ in 1982, a 48.5 kbp double stranded DNA genome (dsDNA) [154]. In 1990 the methodology was used to sequence Cytomegalovirus (CMV) which houses the largest genome of viruses known to infect humans (236 kbp, dsDNA) [155] and Vaccinia virus (190 kbp, dsDNA) [156] the aetiological agent of smallpox. However, it wasn't until 1995 when Fleishmann *et al* were able to overcome the computational limitations associated with assembling hundreds of

thousands of DNA sequences associated with the significantly larger bacterial genome and publish the first bacterial genome of *Haemophilus influenza* (1985.8 kbp) [157]. Sanger sequencing was the original methodology employed for the Human Genome Project but gradually adapted for faster methods allowing more samples to be analysed in parallel. The limitations of Sanger sequencing are that it is slow as a single DNA fragment is analysed at a time and it is less sensitive than newer NGS techniques, which have evolved from it.

1.3.2 Second/next generation sequencing

The original Sanger method is time consuming and laborious but has formed the basis for many currently used commercialised platforms. 'Second generation sequencing' also referred to as 'next generation sequencing' has now been developed where DNA fragments can be amplified in parallel prior to sequence analysis resulting in a massive increase in output and decreased cost as multiple fragments and samples could be analysed simultaneously [158]. Amplification techniques include Illumina's 'bridge amplification' (HiSeq, MiSeq) where two types of flow-cell bound oligonucleotides facilitate the amplification of tagged sample DNA in parallel. Analysis of the DNA sequence is then performed using an approach based on Sanger's 'sequence by synthesis' method. dNTP labelled with a 'terminator', which both emits fluorescence and terminates the DNA sequence is incorporated into the growing sequence. Once the fluorescence is read, the terminator is washed off and the process repeated. The light emitted with each reaction determines the base added to each bound DNA oligonucleotide and the DNA sequence can be determined [159].

1.3.3 Third Generation Sequencing

Third generation sequencing often refers to long read, real-time sequencing such as that performed by Oxford Nanopore Technologies (ONT) platforms such as the MinION where a single DNA molecule is sequenced as it is passed through specifically designed protein nanopores inserted in a synthetic polymer membrane. Enzymes are ligated to double stranded sample DNA which is then drawn into the nanopore via an electrochemical gradient. The enzyme feeds single stranded DNA into the pore, which disrupts the electrical current running across the nanopore. This disruption of the electrical current is characteristic depending on the sequence of the bases travelling through the pore at the time allowing the DNA sequence to be read in real-time. The MinION, the original platform developed by ONT and available on an early access programme since 1994, was the first truly portable sequencer running via a USB cable to a laptop and is smaller than a standard mobile phone. Nanopore sequencing has since been used for a diverse range of applications including real-time outbreak analysis, RNA sequencing and 16S ribosomal

RNA amplicon sequencing as described in the following section. One of the most relevant advantages for clinical applications is that analysis can be performed in real-time as the DNA is passed through the nanopore dramatically shortening the time from sample to analysis [160].

The diverse range of applications, real-time analysis and small size of the MinION sequencer lends itself well to healthcare related applications including outbreak investigation, patient diagnostics and antimicrobial resistance testing [160].

1.4 Clinical application of genome sequencing for pathogen detection and analysis

Since the first bacterial genome was sequenced, genome sequencing has become integral to most university affiliated clinical research settings. The SARS-CoV-2 pandemic has further accelerated this process with 882,400 whole or partial SARS-CoV-2 genomes submitted from around the globe and accessible from the National Centre for Biotechnology information (NCBI) Virus website at the time of writing [161]. Whole genome sequencing can predict antimicrobial susceptibility [150, 162, 163] determine an organism's place within an outbreak [164-169] including identifying novel pathogens [170] and known pathogens that have undergone sufficient mutations such that they are not recognised by standard diagnostic testing [171].

1.4.1 Genomics and antimicrobial resistance

Whole genome sequencing (WGS) can reliably detect antimicrobial resistance genes in *M. tuberculosis* and is now the main method used in the UK National Reference laboratories. WGS is quicker than standard phenotypic sensitivity testing and had comparable sensitivity (92.3%) and specificity (98.4%) [162]. Similarly, WGS predicted antimicrobial resistance of *S. aureus* with a sensitivity and specificity of 99% and 97% respectively when compared to standard culture based techniques [163]. Nanopore based sequencing has been shown to accurately identify bacterial pathogens and resistance genes from urine without prior culture, yielding results within ≈4 hours from sample collection, facilitating early rationalisation of antimicrobials [150].

Metagenomic NGS has also being used to understand the environmental burden of AMR [172]. Fresia *et al* used mNGS to analyse sewage samples from 79 sites in 60 countries estimating the burden of antimicrobial resistance genes. Interesting, the study found AMR gene abundance and diversity to be very different in Europe/North-America and Oceania to Africa/Asia and South America and theorised this may be due to socioeconomic differences between these regions. Only a single sample was analysed from each site and further data is needed to confirm these findings [173]. Understanding the environmental burden of AMR genes, could help to support the global initiative to reduce AMR by identifying where emerging hotspots for AMR are likely to occur allowing targeted investigation and intervention [78].

1.4.2 Epidemiology and outbreak analysis

Strain typing of outbreak organisms is usually performed at a reference laboratory with data being available days to weeks later. This delay means outbreak investigations rely on epidemiological data, which can be inaccurate and cases may be missed or falsely included in the outbreak. An early application of WGS was to attempt to mitigate this delay in outbreak management.

Köser *et al* retrospectively analysed seven outbreak samples and seven non-outbreak samples from a *Methicillin Resistant S. aureus* (MRSA) outbreak on a neonatal unit. WGS demonstrated clustering of the outbreak strains, confirming the outbreak. Furthermore, a missed transmission event was identified in two patients who had shared adjacent bed spaces, both developing an MRSA bacteraemia, 25 days apart. WGS confirmed the bacteraemia isolates were related and transmission had occurred [120]. This study was a retrospective analysis of a small outbreak in a single site, however it illustrates how WGS can be used to confirm outbreaks and identify transmission events that epidemiological investigations have failed to identify.

Quick *et al* developed a novel six-hour protocol for the Illumina MiSeq to allow real-time analysis of isolates from a large outbreak of *Salmonella enterica,* affecting over 30 hospitalised patients, with community, national and European links [174]. Large-scale hospital transmission was excluded by this study, however, the outbreak organism was identified on a food trolley and from eight staff on the outbreak ward. Interestingly, three patients on outbreak wards were found to have asymptomatic carriage highlighting a potential on-going reservoir of transmission. One month after the outbreak, following an update to Oxford Nanopores MinION V7 chemistry Quick *et al* retrospectively sequenced one outbreak and one non-outbreak isolate, the first time this technology had been used in this way. Within 30 minutes both the outbreak and non-outbreak isolate could be identified as *Salmonella enterica* and within 50 minutes the isolates could be distinguished from outbreak and non-outbreak strains.

Quick *et al* subsequently took the MinION to Guinea, West Africa and sequenced 142 Ebola isolates from the largest Ebola epidemic on record [175]. Sample receipt to complete analysis was performed within 1-2 working days despite the bioinformatics analysis being performed remotely via an intermittent Internet connection. The sequence data generated was highly concordant with that generated by Illumina sequencing with no false positive variants called. The noted limitations of this approach include the requirement for a reference genome meaning it is inappropriate for *de-novo* assembly of a novel pathogen which may be the case in a new outbreak, the need for a thermocycler to complete the RT-PCR step plus the required power supply to do this and the reliance on a robust internet connection to provide the bioinformatics

support remotely. This work demonstrates that WGS can be set up rapidly (within 2 days) in a remote setting and generate sequence data good enough to track the ongoing viral transmission of the largest outbreak seen (prior to SARS-CoV-2) since this technology became available.

WGS and mNGS have proven utility in the investigation of outbreaks caused by known pathogens, including TB [169, 176] *S. aureus* [165, 177] and *Neisseria meningitidis* [178]. The introduction of rapid, mobile sequencing platforms such as the MinION, enable real-time sequencing analysis to be performed in non-laboratory settings. Genome sequencing has contributed to real-time investigation of both the recent Ebola [175] and Zika virus outbreak [179] and continues to provide insight into the current SARS-CoV-2 pandemic [97, 107, 180].

1.4.3 Emerging infection

Novel infections that emerge to cause human disease may not be detected by current diagnostics. This may result in inappropriate management of patients, onwards transmission of infection and this may contribute to an outbreak or future pandemic. Untargeted mNGS has been used to identify novel pathogens in patients with undiagnosed infection. For example, Bas-Congo virus was the first Rhabdovirus to be associated with a haemorrhagic presentation when it was identified by mNGS from three cases of haemorrhagic fever in the Democratic Republic of Congo (DRC) two of these cases being fatal [170]. Greninger *et al* used mNGS to demonstrate that enterovirus D68 is associated with acute flaccid myelitis between 2012-2014 in the USA [181], which had not previously been reported. Using mNGS Wise *et el* discovered circulating Oropouche virus in five patients in Ecuador, a country in which it had not previously been identified [171]. The isolated strain from these patients was found to have a mutation in the region of its genomic sequence, which served as the binding site for the reverse primer used in the published qRT-PCR assay for the virus explaining why the virus had not previously been detected by a standard PCR assay [171].

Lassa virus is endemic in Western African countries including Benin, Ghana, Guinea, Liberia, Mali, Sierra Leone, Togo and Nigeria. It can cause a severe viral haemorrhagic sydrome in humans with a high case fatality rate (15-20%) in those who are hospitalised [182].

Lassa virus (LASV), from the Arenaviridae family has a highly variable RNA genome, with up to 25% variability in the S and 32% variability in the L segment. Intra-host variability has been documented with several viral lineages detected in the same host [183]. This variability makes RT-PCR based methods challenging as primer binding targets may differ between strains [184] presenting a challenge to detect all circulating strains. In addition, recent data has shown Lassa virus outbreaks can be associated multiple introductions from the reservoir rodent (Mastomys

natalensis) to humans, rather than human-to-human transmission and so outbreaks may be non-clonal [183, 185].

Kafetzopoulou et al utilised Sequence independent single primer amplification (SISPA) and nanopore sequencing to provide real-time seqeuncing of the 2018 Lassa virus outbreak in Nigeria. Thirty-six genomes were selected across the outbreak at a range of viral loads. SISPA was selected in preference to amplicon based sequencing to overcome the highly variable genome and an individualised reference genome had to be selected for each sample. This study directly supported the public health effort to curtail the outbreak by demonstrating it was due to multiple introductions from the reservoir host as opposed to a more virulent strain causing increased human to human transmission [185]. This complexity of data would not be possible without mNGS technology and demonstrates the power of mNGS in non-clonal outbreaks and pathogens with highly variable genomes.

1.5 Metagenomic next generation sequencing

1.5.1 Background

The last 15 years has seen sequencing technology becoming more affordable and accessible, mNGS has been proposed as a potential diagnostic tool for infectious disease diagnostics and has shown favourable sensitivity against standard diagnostic investigations. Metagenomic NGS aims to comprehensively sequence all RNA and DNA within a clinical sample. Methods can be used to sequence DNA, RNA or both depending on the protocol used. Host genomic material present in the sample is also sequenced but this can be computationally removed prior to analysis if required. mNGS is being explored as a tool for infectious disease diagnostics and where AUFI is concerned has the notable benefit of being 'untargeted', i.e. has the ability to detect a wide range of pathogens without specifically determining what these may be before analysis/sampling.

1.5.2 Application of mNGS to clinical settings

Investigating the diagnostic use of mNGS, Duan *et al* demonstrated an increased sensitivity of mNGS compared with standard diagnostic testing in a study of 109 adults admitted to a single hospital in Shanghai over a 14 month period [186]. Of the cohort analysed, 84.4% (92/109) of patients were deemed to have an infectious disease by the treating physician, 14.7% (16/109) a non-infectious illness and in 0.9% (1/109) the diagnosis was unknown. This study found mNGS had an overall higher sensitivity than culture with 67.4% (60/89) of infection diagnoses identified by mNGS versus 23.6% (21/89) for culture against a wide range of pathogens and sample types including common Gram negative bacteria, nutritionally variant streptococci, *Pneumocystis carinii* (PJP), *Mycobacterium tuberculosis* and CMV and a wide variety of sample types. This difference was significant for tissue p=0.025, blood p <0.001, sputum p =0.03 and bronchial alveolar lavage fluid p =0.002 [186]. This study demonstrates mNGS can be used in a clinical setting and may be superior to current infection diagnostics. However, limitations include the retrospective nature, single centre and small sample size. There is no detail on standard diagnostic methodology and 'culture' is listed as the comparator which would not typically be used for PJP as it is notoriously difficult [187]. Additionally, *Sacromyces cerevisiae* was reported as a potential pathogen in 2.9% (2/69) patients. *Sacromyces species* have been identified as an emerging pathogen associated with probiotic use but significant infection is thought to be rare and limited to immunocompromised individuals, careful clinical correlation is required to determine its significance, particularly as immunocompromised individuals were excluded from this study [188].

Parize et al support the findings of Duan *et al* in a multi-centre, blinded prospective study comparing untargeted mNGS and conventional microbiology in immunocompromised adults. Of 101 adults recruited, clinically relevant pathogens were detected by mNGS in a significantly higher proportion of patients 36% (36/101) versus 11% (11/101) for conventional diagnostics. This difference was upheld even when conventional diagnostics were continued for thirty days only improving the detection rate to 19% (19/101). mNGS missed five diagnoses, three of which were localised infection so may not have been detectable in the blood samples (HSV 1, candida vulvitis, Nocardia subcutaneous abscess) and two arose after the mNGS sample was taken although the details of these two cases are not provided. This study demonstrated a high sensitivity and high negative predictive value of mNGS on blood samples of immunocompromised adults for pathogen detection. The main limitation of this study was that the NGS samples were only taken at one time point whereas standard diagnostics were continued for thirty days. However, this in itself is very useful as it demonstrated the yield from standard diagnostics improved from 11% (11/101) to 19% (19/101) when repeat sampling was performed but this still did not approach the sensitivity of mNGS. Additionally, fungal targets were not included in the NGS analysis and so could not be compared. It is also conceivable that immunocompromised individuals may have a higher pathogen burden due to lack of host immune response and therefore these findings may not be reproducible in individuals who have competent immune systems [189].

1.5.3 Role of metagenomic next generation sequencing in acute undifferentiated febrile illness

mNGS has been used to diagnose infection where extensive investigations have failed, for example, Wilson *et al* describe a case of a 14 year old immunocompromised boy presenting three times over 4 months with a headache and hydrocephalus, with prior extensive diagnostic work up. A diagnosis of Leptospirosis infection was finally made with unbiased mNGS leading to appropriate treatment and recovery [190]. Naccache *et al* identified a subsequently fatal Astrovirus infection in CSF and brain tissue of a patient with chronic lymphocytic leukaemia (CLL) using mNGS and their sequence-based ultra-rapid pathogen identification pipeline (SURPI) [191]. In addition, Wang *et al* recently report the use of NGS to diagnose Leishmaniasis on peripheral blood and bone marrow aspirate on day five of inpatient admission leading to the initiation of appropriate therapy. These findings were corroborated by species specific PCR [192]. Broad range PCR and mNGS have been used to retrospectively diagnose rabies [193] and mNGS also helped diagnose *Balamuthia mandrillaris* encephalitis in a 13 year old girl who was suspected to have TB meningitis [194]. However, it must be noted, in most cases described above, clinical samples were analysed by experts in the field outside the scope of routine healthcare laboratory.

Work has been done to examine the role of mNGS as a diagnostic tool in AUFI and undiagnosed fever. Susilawati *et al* identified a pathogen with mNGS sequencing in 80% of individuals presenting with AUFI who remained undiagnosed after standard investigations. Although the overall numbers were very small (8/10) the pathogens were significant and included *E. coli*, *Orientia tsutsugamushi*, Dengue virus and Jingmen tick virus. Difficulty was noted in obtaining high quality DNA and also in the large number of contaminating (and human) reads identified [16]. Yozwiak *et al* used Virochip microarray and deep RNA and DNA sequencing to identify a further 37% (45/123) of undiagnosed tropical febrile illness in Nicaragua, the major pathogens being Dengue virus, HHV6 and hepatitis A virus. However, members of the Anelloviridae virus family, Tenotorque virus (TTV) accounted for 21/45 17% of these. The role of TTV's in human infection is yet to be fully understood and TTV is often identified in healthy individuals [195].

Untargeted mNGS has a shown a great deal of promise to identify infectious pathogens in clinical specimens where other methods have failed and there is a clear unmet need for this in AUFI. However, clinically focused, large, prospected studies are needed to fully explore its potential and limitations within the routine clinical setting. Understanding how mNGS could integrate into the existing health service [196] and standardising methodologies and provide a robust analytical pathway suitable for the clinical diagnostic laboratory, accessible to low resource settings should be a focal point of further research.

1.5.4 Limitations and further developments

Metagenomic NGS has the potential to improve our ability to rapidly diagnose and treat infection and identify and manage outbreaks of common and novel pathogens. However, there are a number of significant barriers, which need to be overcome before mNGS can be widely used as a diagnostic test in a clinical setting. Firstly, clinical diagnostic tests must be standardised and regulated and meet an acceptable pre-determined sensitivity, specificity and error rate. An example of the high standards for molecular based diagnostics is illustrated by the standards for detection of SARS-CoV-2 RNA in clinical samples. The UK Medicines and Healthcare products Regulatory Agency (MHRA) state that RNA extraction based assays for SARS-CoV-2, have a desirable sensitivity and specificity greater than 99% with a 95% two-sided confidence interval above 97% with an acceptable standard being a sensitivity and specificity above 95% with a 95% two-sided confidence interval above 90% [197]. Other significant issues with analysis including sample contamination with non-pathogenic organisms and the high human DNA content of clinical samples compared with the much less abundance pathogen genomic material of interest which can overwhelm the sequence output.

Metagenomic NGS studies to date have been experimental and used a diverse range of sample preparations, extraction methods and sequencing protocols to generate sequence data on different sample types, none of which has been standardised. This is well described by Govender *et al* who performed a meta-analysis of 2023 samples included in 13/21 eligible studies examining the diagnostic test accuracy of mNGS in multiple sample types including blood, CSF and orthopaedic samples using a range of methodologies [198]. This study found 76% (16/21) studies reported contamination of samples and 24% (5/21) did not comment on contamination at all. Not all studies included negative, positive and internal controls 61% (13/21) essential for quality control. In 19% (4/21) of studies where antimicrobial resistance profiles were deduced from sequence data, phenotypic correlation was found in 88% of cases, however high major and very major error rates were observed at 5% (0%-12%) and 24% (8%-40%) respectively. The authors note that the US Food and Drug Administration require an error rate ≤ 3% with a lower limit confidence interval ≤1.5% and an upper limit of ≤7.5% far below what is currently seen in the literature illustrating the gap between current methods and a robust clinical diagnostic test. The high cost of $130-685 per sample was noted and the high percentage of human reads with a median value of 91%, demonstrating the need for improved human DNA depletion methods.

Novel approaches to human DNA depletion have been developed to try and address this problem. Charalampous *et al* report a novel saponin based human DNA depletion step in a study of Nanopore metagenomic sequencing of lower respiratory tract samples [199]. With this methodology, 99.99% of host DNA was depleted from respiratory samples and nanopore sequencing was found to be 96.6% sensitive and 41.7% specific as compared to culture on the 40 samples analysed.

For accurate mNGS results, sample selection and storage, data storage and complex bioinformatics analysis is needed for robust reproducible results. RNA is prone to degrade if not stored appropriately and therefore how a sample is obtained, transported and stored could have a significant impact on the subsequent yield. Additionally, some samples have a low burden of organisms such as CSF, which can be problematic for mNGS. Perlejewski *et al*, compared mNGS to RT PCR/PCR in 21 CSF samples from patients with viral encephalitis [200]. The group first assessed their metagenomic protocol for sensitivity using serial dilutions of HIV and HBV as RNA and DNA targets respectively. MNGS correctly identified serial dilutions of HIV and HBV, which had been added to CSF samples of an uninfected patient down to 10^2 copies of HIV and 10 copies of HBV virus. The same mNGS protocol was used to analyse CSF samples via the RNA metagenomic pathway for 28.6% (6/21) samples with proven Enterovirus or the DNA metagenomic pathway for 62% (13/21) samples containing Herpes Simplex Virus (HSV), 4.8% (1/21) with CMV and 4.8% (1/21) with Varicella Zoster Virus. Only 28.6% (7/21) of all isolates were correctly identified by

mNGS, including 16.7% (1/6) Enterovirus and 46.2% (6/13) HSV isolates. This study highlights the importance appropriate patient and sample selection for mNGS as in this example standard diagnostic testing performed significantly better than mNGS. However, this was a retrospective study and there is no detail provided on sample storage or processing prior to mNGS, which may have influenced the yield. In addition, if a novel pathogen had been present, RT-PCR would not have identified this and mNGS would be a more suitable methodology.

Analytical bioinformatics software and sample protocols are now more freely available to support interpretation of mNGS data, however, there is no agreed approach to best tackle a particularly sample type, clinical scenario or clinical question which may lead to variation in study findings. There is a clear need for standardised studies of the clinical diagnostic utility of mNGS and best analytic approaches. The STROBE metagenomics guideline published in 2020 attempts to highlight the common errors in the current metagenomic literature and provide a framework to inform better study design and reporting [201].

Despite the challenges and limitations of current available literature, mNGS remains a plausible tool for infection diagnosis where standard diagnostics have failed and can perform as well or better than standard techniques depending on the sample type and methodology employed. To further explore the diagnostic utility of mNGS for AUFI well designed, controlled clinical studies of well characterised patients with infection, with appropriate and clearly documented sampling, storage and analysis are needed.

1.6 Knowledge gap and summary

At the time of writing there are no published UK-based studies of AUFI in adults. The global literature demonstrates a high burden of disease from AUFI, with high levels of mortality and morbidity demonstrated in some studies. No microbiologically confirmed diagnosis is made in a high proportion of cases, however, there is great variation reported in the studies reviewed here. Empirical broad-spectrum antimicrobial use is very common and although there is limited data available, levels of AMR in this group are concerning where reported. Data from LMIC, largely in tropical and sub-tropical regions, cannot be extrapolated to predict UK based disease. This is due to differing disease prevalence, presence of different vectors and the differing occupational and environmental exposures highlighted previously.

The lack of data on AUFI in the UK adult population mean that key questions remain unanswered. These include, but are not limited to; presenting clinical features, aetiology of the illness, how many cases remain undiagnosed, details of antimicrobial use, how many cases are attributable to imported and vector borne diseases, and whether there is evidence of onwards transmission of infection. SARS-CoV-2 has highlighted and reinforced Murchison's view from 137 years ago that identifying the causative agent of an infectious process is essential for patient management and containing its spread. The risk of emerging infection originating and/or being transported into the UK is not insignificant, particularly as the earth continues to warm allowing insect vectors to travel northwards altering the geographic spread of infectious disease. This highlights the need for diagnostic tests able and available to identify novel or unexpected pathogens.

The clinical study described here is believed to be the first, proof of concept clinical study examining adults presenting to a UK hospital with AUFI. Firstly, it aims to provide a detailed description of one hundred adults presenting to hospital with AUFI, seeking to identify a clinically credible microbiologically confirmed diagnosis in all. Fifty healthy volunteers are included as a control group. Details of presenting symptoms and signs, significant exposures such as travel and antimicrobial use will be presented in addition to the results of standard diagnostic testing. Secondly, the diagnostic utility of mNGS will be explored in participants with AUFI.

1.6.1 Aims and objectives

1.6.1.1 Aims

- To describe the baseline clinical features, laboratory, radiological, and microbiological findings and clinical outcome of adults presenting to a UK hospital with an acute undifferentiated febrile illness
- To assess the diagnostic utility of mNGS in adult patients presenting with acute undifferentiated febrile illness via comparison to the current standard routine diagnostic testing

1.6.1.2 Objectives

- Describe the clinical features, microbiological diagnosis and outcome of adults presenting to a UK hospital with an acute undifferentiated febrile illness using a prospectively recruited case-controlled study
- Compare the proportion of patients given a final credible microbiological diagnosis with mNGS and standard diagnostic testing in participants tested with both methods
- Assess the concordance of pathogen detection with mNGS and conventional diagnostic test by comparing the individual diagnoses made with the two modalities in those patients with a diagnosis achieved by standard of care diagnostic tests
- Assess the clinical significance of potential pathogens detected by mNGS by comparing the organisms detected by mNGS in patients with AUFI and those in healthy controls
- Assess the potential clinical impact of mNGS as a real time diagnostic test in patients with AUFI with emphasis on additional 'actionable diagnoses' made using mNGS
- Inform the feasibility and design of a planned randomised controlled trial evaluating the clinical and health economic benefits of using mNGS as a 'front door' real time diagnostic platform in hospitalised adults with suspected infection

Chapter 2 Methods

2.1 Acute undifferentiated febrile illness clinical study

2.1.1 Study design and patient population sampled

This prospectively recruited case controlled observational study (The use of unbiased mNGS for pathogen detection in adults hospitalised with acute undifferentiated febrile illness and suspected infection: an observational pilot study: SePSI, ISRCTN11747901) was carried out at a single UK site with some mNGS sample analysis carried out at a second site.

University Hospital Southampton NHS Foundation Trust (UHS) is a tertiary referral hospital serving a 1.9 million local population and provides specialist services to a total of 3.7 million people in central southern England and the Channel Isles [202]. UHS also serves Southampton's busy passenger and cargo port which received ~2 million passengers a year prior to the COVID-19 pandemic [203]. Consequently, unwell travellers and crew contribute to the range of imported infections seen at the UHS site. This study was designed as a small proof of concept study of acute undifferentiated fever in the UK to explore the aetiology, outcome and antimicrobial use within this group and determine the clinical utility of unbiased mNGS as a diagnostic tool in AUFI.

2.1.2 Sample size

The study aimed to include approximately 50 participants who remained undiagnosed at follow up, 50 with a clinically credible diagnosis and 50 age-matched healthy volunteers making a total sample size of 150 participants. The estimated sample size of 100 and the proportion of participants who would be diagnosed (50%) were based on a retrospective analysis of the UHS inpatient infectious diseases database (2013-2014), which listed all patients referred to the infectious disease service (unpublished data). This data suggested that 50 patients with AUFI could be recruited at UHS per year and that approximately half of these would achieve a clinically credible diagnosis by standard diagnostic testing.

2.1.3 Ethical approval

The study protocol and all study documents were approved by the local UHS research and development team and by the York and Humber Regional Ethical Committee (REC) on the 15th October 2015, REC number: 15/YH/0429, IRAS number: 174619 prior to commencing the study. Two substantial amendments were submitted. The first to allow Qiagen PAXgene™ blood RNA

tubes to be taken from half (25) healthy volunteers and the second to add a research collaborator. These were both approved by the REC.

2.1.4 Informed consent

All recruited patients (including healthy volunteers) provided informed written consent prior to participation (see Appendix A1 and A2). At least thirty minutes were given for patients to read the participant information leaflet and the opportunity to discuss with their next of kin and the research nurse or doctor prior to recruitment. Participants were able to withdraw their consent at any point in time and had a direct email address and telephone contact for the lead investigator.

2.1.5 Participant selection

2.1.5.1 Febrile participants

Due to the small study size and single site design, all medically trained researchers on the study delegation log were also members of the clinical infectious diseases team and/or acute medical unit (AMU) team at the time of the study. The clinical AMU team reviewed electronic patient admission lists for the AMU and identified potential participants and in the Emergency Department (ED) Monday to Friday until 4pm as this was the latest time research samples could be processed by the laboratory. Those highlighted to have a fever with no clear cause were considered for participation in the study. Patients were then asked whether they would consider taking part in the research study, if they agreed, a research nurse or doctor approached them. A patient information leaflet (Appendix B) was given to or read out to the individual and if they agreed to take part informed consent was sought as described above.

The inclusion and exclusion criteria were as follows:

2.1.5.1.1 Inclusion criteria

- Aged 18 years or over.
- Deemed by an appropriately trained medical professional to have capacity to give informed, written consent and is able and willing to adhere to the study procedures
- Is a patient in University Hospital Southampton Acute Medical Unit or Emergency Department OR is under the care or advice of the inpatient infectious diseases service
- Recruited to the study
    - within a 72-hour period of first triage by ED staff OR
    - within a 72-hour period of arrival on AMU (if admitted directly to AMU)
- Has an acute febrile illness with a documented fever ≥38°C OR a history of fever in the preceding 72 hours
- Has a duration of illness less than or equal to 21 days
- Has an illness lacking localizable or clear organ-specific clinical features (as determined by the investigators), including but not limited to:
    - Pneumonia (as defined by new radiological consolidation)
    - Urinary tract infection
    - Cellulitis or other skin and soft tissue infections
    - Septic arthritis
    - Infected prosthetic material
    - Pyogenic spondylodiscitis
    - Meningitis
- Has an illness lacking a clear non-infectious aetiology

2.1.5.1.2 Exclusion criteria

- Patients not fulfilling inclusion criteria
- A decision to palliate the patient's symptoms taken by the treating clinicians
- Declines collection of clinical specimens
- Immune compromised as defined by:
  - HIV infection with a CD4 count of less than 200 cells/µl
  - Any primary immunodeficiency
  - Current or recent (within six months) chemotherapy or radiotherapy for malignancy
  - Solid organ transplant recipients on immunosuppressive therapy
  - Bone marrow transplant recipients currently receiving immunosuppressive treatment, or who received it within the last 12 months
  - Patients with current graft versus host disease
  - Patients currently receiving high dose systemic corticosteroids (equivalent to ≥ 40 mg prednisolone per day for ≥3 week in an adult), and for at least three months after treatment has stopped [204]
  - Patients currently or recently (within three months) on other types of immunosuppressive therapy.
- The investigator feels that patient should not be enrolled (i.e. investigator discretion)

2.1.5.2 Healthy volunteers

Healthy age-matched controls were invited to take part in the study by means of local advertising within the hospital and medical school. This local recruitment approach was taken as both medical students and hospital staff are required to have HIV, Hepatitis B virus (HBV) and Hepatitis C virus testing prior to enrolment. Therefore, it was not necessary to test all healthy volunteers for blood borne viruses at recruitment (as was done for febrile participants) as the chance of having undiagnosed serious infection which could be later detected by unbiased mNGS was low. Healthy volunteers were counselled about the small possibility that mNGS performed sometime in the future could identify an infectious agent such as HIV if they had been infected after their occupational health screening and that involvement in the study did not take the place of standard monitoring and healthcare they may otherwise seek for blood borne and sexually transmitted infections.

Healthy volunteers were able to directly contact the research team. If all study inclusion and exclusion criteria (see below) were satisfied following an initial telephone consultation, they were then invited in for a face-to-face appointment. A research nurse or doctor would outline the study and answer and questions and provide them with a participant information leaflet to read

(Appendix B). The participant could then take the information away and make an appointment if they wished to participate or most commonly would choose to participate in the study on the same day. Written informed consent was sought prior to recruitment as previously discussed.

Inclusion and exclusion criteria were as follows:

2.1.5.2.1 Inclusion criteria

- Aged 18 years of age or over
- Deemed by an appropriately trained medical professional to have the capacity to give informed, written consent and is able and willing to adhere to the study procedures.
- A negative pregnancy test for women of childbearing potential
- Has been well with no symptoms of significant illness (including; fever, chills, sweats, myalgia, arthralgia, malaise, weight loss, cough, chest pain, rhinorrhoea, sore throat, abdominal pain, diarrhoea, dysuria, urinary frequency, haematuria, severe headache, collapse or seizure) in the past 14 days.
- Normally fit and well with no significant medical co-morbidity (including chronic cardiovascular, respiratory, renal, hepatic or neurological illness, diabetes mellitus, malignancy)
- Not immune compromised (as defined above).

2.1.5.2.2 Exclusion Criteria

- Failure to meet inclusion criteria
- Receiving antibiotics or antiviral treatment in the last 2 weeks
- Developed a febrile illness within 7 days of recruitment

All healthy volunteers received a telephone call or email follow up at 7 days post recruitment by a research nurse to ensure they did not develop a febrile illness following recruitment. This was to account for a short pre-symptomatic incubation period of an infectious illness at the time of recruitment. In the event an infectious illness had developed, those participants were withdrawn from the study and their samples not tested.

2.1.6 Data collection and sampling

2.1.6.1 Safety

The risks associated with collection of nose and throat swabs and blood are minimal and, where they occur, are likely to be mild. These were the only direct interventions performed as part of the study so the overall risk to study participants was low. Any adverse events (AE) and serious

adverse events (SAE), temporarily relating to study participation were recorded. In the event of a SAE, the principal investigator (PI) was involved in deciding whether this was a study-related event. SAE reporting was limited to 30 days post hospital discharge in this a non-interventional trial.

2.1.6.2 AUFI participants

2.1.6.2.1 Clinical data collection

Patients admitted with AUFI had a thorough clinical history taken and physical examination by the admitting team as standard of care. In addition, a research doctor or nurse completed a detailed case report form (CRF) (See Appendix C). This data included the participant's age, sex, ethnicity, travel and occupational history, exposure to animals or unwell contacts, active symptoms, physical signs, nursing observations as well as the results of laboratory and radiological investigations. All participants with AUFI were followed up at 28 +/- 14 days in the outpatient setting by a research doctor; the follow up section of the CRF was completed at this appointment. If there were on-going medical conditions or symptoms which required review, this was highlighted to the medical consultant responsible for the patient so appropriate follow up could be arranged once participation in the study had ended. Outcome data was collected retrospectively including duration of hospitalisation, use of antibiotics and antivirals, final diagnosis, admission to critical care, 30-day mortality and readmission to hospital.

2.1.6.2.2 Biological samples

Standard clinical investigations are defined as investigations performed to ascertain the likely diagnosis in the context of the clinical presentation. As part of standard care for patients presenting with fever, investigations included: blood cultures, urine cultures, respiratory samples for viral PCR and/or bacterial culture, stool samples for microscopy and culture if diarrhoea was present, serum samples depending on the differential diagnosis and cerebrospinal fluid (CSF) and tissue specimens where appropriate. This full sample set was not obtained for all participants recruited as it was not clinically appropriate in all, the sample set received was recorded on the participant case report form. At the time of this study, testing for pregnancy in women of childbearing potential was performed as standard of care at UHS by urine β-HCG testing and results were communicated to the patient by the responsible clinical team.

Following informed consent, research staff obtained the following research samples: 10 mL EDTA whole blood in plastic K_2EDTA BD collection tubes (BD LifeScience, New Jersey), 2.5 mL whole blood into Qiagen PAXgene™ blood RNA tubes (Qiagen, Hilden, Germany), 10 mL of blood into plastic serum separator tubes (BD LifeScience, New Jersey), nasopharyngeal Sigma Viracult™

swabs in viral transport media (Medical Wire, Corsham) and a mid-stream urine plus faeces if diarrhoea or a history of diarrhoea was present. EDTA whole blood samples were transported to the research laboratory on ice to maximally preserve genomic material. On arrival they were split into 1 mL aliquots and frozen at −20°C within one hour of collection, all other samples were transported at room temperature within one hour to the laboratory. Urine and nasopharyngeal samples were frozen at −20°C and serum samples were separated and the serum aliquoted in 1 mL tubes and frozen at −20°C. If deemed clinically necessary as part of standard care, CSF and untreated tissue biopsy material (such as lymph node) were also collected. Surplus CSF or biopsy material which remained following standard diagnostic testing was stored for later analysis initially at −20°C and then subsequently at−80°C for long-term storage.

Further research samples were obtained at the 28 +/- 14 days follow up appointment: 10 mL EDTA whole blood (BD LifeScience, New Jersey), 10 mL of blood in serum separator gel tubes (BD LifeScience, New Jersey), nasopharyngeal Sigma Viracult™ swabs in viral transport media (Medical Wire, Corsham) and mid-stream urine samples in sterile containers. Both recruitment and follow up samples were processed and stored as per the study laboratory protocol (See Appendix D).

2.1.6.3 Healthy volunteers

2.1.6.3.1 Clinical data collection

Healthy volunteers were approximately age and sex matched to participants in the AUFI group. Following informed written consent, all healthy volunteers had a face-to-face consultation with a research nurse or doctor where the Healthy volunteer CRF was completed (See Appendix E). Basic demographic data was collected including age, sex, ethnic origin and past medical history. Details were elicited of any infectious illness in the preceding two weeks because of the potential for any residual genomic material from a prior infection being later detected by mNGS.

2.1.6.3.2 Biological samples

Prior to recruitment the small possibility of detecting undiagnosed blood borne virus infection (HIV, Hepatitis B and Hepatitis C) by later metagenomic NGS was discussed in the patient information leaflet and the written consent form. It was also discussed in the face-to-face meeting prior to recruitment to check understanding. All female healthy volunteers of childbearing potential were required to have a negative urinary pregnancy test at enrolment. Research staff then obtained samples of 10 mL whole blood into plastic K_2EDTA containing tubes (BD LifeScience, New Jersey), 10 mL of blood in plastic serum separator tubes (BD LifeScience, New Jersey), nasopharyngeal Sigma Viracult swabs in viral transport media (Medical Wire, Corsham) and a mid-stream urine sample in a sterile container. Samples were transported,

aliquoted and stored using the same protocol as the samples from AUFI participants. In 50% 25/50 of healthy volunteers an additional 2.5 mL of EDTA whole blood into Qiagen PAXgene™ blood RNA tubes (Qiagen, Hilden, Germany) was obtained during the same venepuncture.

2.1.7 Analysis

2.1.7.1 Primary outcome measure

The primary outcome of this study was to measure the proportion of patients with a clinically credible diagnosis made with standard diagnostic testing.

2.1.7.2 Secondary outcome measures

The secondary outcome measures were: a) to determine the sensitivity and specificity of mNGS in patients where a clinically credible diagnosis was made using standard current investigation and b) to determine the proportion of 'actionable diagnoses' made by the mNGS results (i.e. a potential change in management was indicated if mNGS result was known by the clinical team in real time).

2.1.8 Data Storage

All participant data was collected on paper case report forms and stored in study files either within the research department or in the principal investigators locked office. Data could only be accessed by members of the research team listed on the delegation log stored in the site file. Data was collated onto an anonymised spreadsheet for analysis, which was stored on the secure hospital trust network drive.

2.1.9 Statistical analysis

Statistical analysis was performed using Microsoft Excel™ for Mac (Version 14.7.2, year 2011) and GraphPad Prism™ Software for macOS (Version 9.2.0, year 2021 Software Inc; La Jolla, California). The entire cohort was described, then subgroup analysis was then performed based on whether patients had a diagnosis made by conventional diagnostics 'diagnosed' or not 'undiagnosed'. Non-parametric continuous variables such as age were summarised using median [interquartile range] and compared between groups using the Mann-Whitney U test. Categorical non-parametric data was described a number (percentage) and equality of proportions between groups was compared using Fishers exact test.

This study was prospectively registered with the ISRCTN trial database, ISRCTN1174790

2.2 DNA extraction and limit of detection analysis

The section describes the methods used to extract *E.coli* DNA from healthy volunteer EDTA whole blood. First, a growth curve for *E.coli* was generated and mid-logarithmic growth phase calculated (2.2.1.1 to 2.2.1.3). Then a dilution series of known quantities of *E.coli* cfu/mL was generated (2.2.1.3). Healthy volunteer EDTA whole blood was then 'spiked' with known quantities of *E.coli* (2.2.2.1). Extraction of microbial DNA was then performed. These experiments used live cultures of *E.coli* and healthy volunteer whole blood in EDTA to mimic human blood stream infection with *E.coli*. The abundance of bacteria present in blood during blood stream infection has been shown to vary with concentrations of bacteria as low as 1 cfu/mL [205]. A range of concentrations of *E.coli* was used to reflect a range of different clinical scenarios and understand the limit of detection of the extraction and enrichment techniques.

2.2.1 Creating a dilution series of *Escherichia coli* ATCC 2955 stock cultures

2.2.1.1 Culture of *Escherichia coli* ATCC 2955

Powdered Brain Heart Infusion Broth (BHI) (Oxiod, Basingstoke) was made up according to manufacturer's instructions (37 g of powdered BHI broth in 1 L distilled water) and autoclaved at 121°C for 15 minutes. *E. coli* ATCC 2955 was then sub-cultured with a loop onto Cystine-Lactose-Electrolyte-Deficient (CLED) solid agar (BD, Heidelberg) and incubated at 37°C in 5% CO_2 overnight.

Broth culture of *E. coli* ATCC 2955 was performed in 25 mL aliquots of sterile BHI liquid media in 50 mL Falcon tubes. Inoculation was performed using either single colonies from the overnight CLED plates checked for monomorphic pure growth, or for subsequent cultures 20 ml of an overnight broth culture. Inoculated broth was then incubated with secured loosened lids at 37°C in 5% CO_2.

2.2.1.2 Determining growth curve for *Escherichia coli* ATCC 2955

Replicate Falcon tubes were labelled as follows: Growth Curve 1 (GC1), Growth Curve 2 (GC2), Growth Curve 3 (GC3), control and blank. The blank sample was used as a baseline to zero the spectrophotometer prior to each analysis. The control sample was uninnoculated broth subject to the same conditions as GC1-3 and used to control against any bacterial contamination of the sample set. Overnight liquid cultures of *E. coli* ATCC 2955 were re-suspended until uniformly cloudy prior to removing 20 μl for inoculation of GC1, GC2 and GC3. The time was noted down to represent 'time zero' in the resultant growth curve. The lids of GC1-3 and the control were

loosely secured with tape and the lid of the blank was tightly secured. All five Falcon tubes were placed in a rack and moved to the incubator at 37°C in 5% CO_2. Every 90 min 1 mL of culture was removed and the OD_{600nm} measured using a spectrophotometer. The time-point which represented mid-log phase was then calculated using the mean values of samples GC1-3 providing the growth curve demonstrated the four phases of bacterial growth: lag, logarithmic or exponential, stationary and death phases using the following steps. Time was plotted on the x-axis (non-logarithmic) versus optical density (OD_{600mm}) on the y-axis in logarithmic scale. The start and end points of logarithmic growth were identified and used to calculate the midpoint with the following calculation.

Equation 2.1 Calculation of mid-logarithmic growth phase: example

Convert y-axis into Log scale

Mid logarithmic growth point time:

= ((Time of Log growth phase end) – (Time of Log growth phase start)/2) + (Time of Log growth phase start)

= XX:XX (hrs:mins from culture set up)

Mid logarithmic growth point OD

= ((End OD of Log growth phase) – (Start OD of Log growth phase))/2) + (Start OD of Log growth phase)

= X.XX ppm

2.2.1.3 Generating a stock dilution series of *Escherichia coli* ATCC 2955

Steps as described in 2.2.1.1 to 2.2.1.2 were repeated until the optical density in cultures GC1-3 reached that at which mid-log phase has been previously demonstrated. Eighty-one 2 mL microcentrifuge tubes were labelled with a corresponding CLED plate for each dilution. Nine 1 mL samples were made at each dilution step between neat and a 1 in 10^8 dilution in BHI as illustrated in table 2.1. A 10 ul aliquot of each sample was added to the corresponding CLED plate, spread evenly with a plate spreader and incubated at 37°C in 5% CO_2. The microdilution series was clearly labelled and placed in the -20 °C freezer. After 24 hours incubation, plate counts were performed and colony forming units per mL (CFU/mL) were determined for each frozen aliquot using the following calculation: CFU/mL = (CFU on plate) × 100.

Table 2.1 Dilution series set up *E. coli*

	Growth curve cultures (GC1-3)								
Dilution series	**1A**	**1B**	**1C**	**2A**	**2B**	**2C**	**3A**	**3B**	**3C**
A- neat	1AA	1BA	1CA	2AA	2BA	2CA	3AA	3BA	3CA
B- 10^1	1AB	1BB	1CB	2AB	2BB	2CB	3AB	3BB	3CB
C- 10^2	1AC	1BC	1CC	2AC	2BC	2CC	3AC	3BC	3CC
D-10^3	1AD	1BD	1CD	2AD	2BD	2CD	3AD	3BD	3CD
E-10^4	1AE	1BE	1CE	2AE	2BE	2CE	3AE	3BE	3CE
F-10^5	1AF	1BF	1CF	2AF	2BF	2CF	3AF	3BF	3CF
G-10^6	1AG	1BG	1CG	2AG	2BG	2CG	3AG	3BG	3CG
H-10^7	1AH	1BH	1CH	2AH	2BH	2CH	3AH	3BH	3CH
J-10^8	1AJ	1BJ	1CJ	2AJ	2BJ	2CJ	3AJ	3BJ	3CJ

2.2.1.4 Generating a stock dilution series of *Candida albicans* ATCC 90028

A stock dilution series of *Candida albicans* was prepared by repeating steps 2.2.1.1 to 2.2.1.3 above with the following exceptions. In section 2.2.1.1 Sabouraud dextrose (SAB) broth (Thermofisher, Waltham, Massachusetts) was used instead of BHI broth and in section 2.2.1.3 SAB agar was used in place of CLED (Thermofisher, Waltham, Massachusetts).

2.2.2 Microbial DNA extraction from blood

The two extraction kits used in comparison were; the Molysis Complete 5 Extraction Kit (Molzym, Bremen, Germany) and the Blood and Cell culture DNA Maxi Kit (Qiagen, Hilden, Germany). Samples of EDTA whole blood obtained from healthy volunteers were spiked with known quantities of *E. coli* ATCC 2955 which had been stored from the prior experiment.

The Molysis Complete 5 Extraction Kit (Molzym, Bremen, Germany) was used according to manufacturer's instructions following Protocol 2 for a 5mL sample. Briefly, the kit uses a chaotropic buffer to lyse human cells whilst leaving bacterial cells intact after which the host DNA was then degraded by a specific chaotropic resistant DNase (*MolDNase B*). Microbial cell walls were subsequently degraded using *BugLysis* reagent for Gram-positive bacteria and *Proteinase K* treatment for Gram-negative bacterial and fungi. Microbial DNA is then extracted and bound to a filter matrix before washing and eluting. Differential lysis of the human then microbial cells yields a product which is deplete in human gDNA and enriched for target microbial DNA. Eluted DNA was stored in Tris-EDTA (TE) buffer at either 4-12^{o}C if analysed the same day or frozen at -15-25^{o}C if analysis with quantitative PCR was performed later.

The Blood and Cell culture DNA Maxi Kit (Qiagen, Hilden, Germany), was used according to manufacturer's instructions. Briefly, the kit uses gravity assisted Genomic Tips containing a unique Qiagen Anion Exchange Resin able to bind DNA following cell lysis. Both host and microbial DNA can then be eluted. This process was followed by magnetic bead depletion of human gDNA and enrichment of microbial DNA with the NebNext Microbiome DNA Enrichment kit (New England Biolabs, Massachusetts). The NebNext Microbiome DNA Enrichment kit (New England Biolabs, Massachusetts) and Agencourt AMPure XP Magnetic Beads (Beckman Coulter, California) DNA clean up were then used according to manufacturer's recommendation following protocol E2612. The depletion protocol relies on the presence of methylated cysteines in human genomic DNA, which is rarely found in microbial species. Protein A bound magnetic beads, which bind the Fc fragment of MBD2 Fc antibodies, are used to selective bind human gDNA. The MBD2 proteins then bind tightly to CpG regions of human gDNA. Once a magnetic field is applied, tightly bound human gDNA complexed with the magnetic bead/protein A complex is drawn to the side of the tube and microbial DNA can be obtained from the supernatant. Microbial DNA clean up using Agencourt AMPure XP Magnetic Beads (Beckman Coulter, California) then follows to remove any remaining buffer salts or enzymes. Eluted DNA was stored in TE buffer at 4-12^{o}C if analysed the same day or frozen at -15-25^{o}C for analysis later.

2.2.2.1 Spiked EDTA whole blood samples with known quantities of *E. coli* ATCC 2955

Plate counts generated and calculated cfu/mL from the *E. coli* ATCC 2955 dilutions series in section 2.2.1.3 were used to calculate which aliquots of *E. coli* ATCC 2955 (see table 2.2) to add to a 4.1 mL volume of whole blood stored in K_2EDTA in order to generate 1000 cfu/mL, 100 cfu/mL, 10cfu /mL, 1 cfu/mL concentrations of *E. coli* ATCC 2955 once the 0.9 mL frozen aliquot was added (CFU per mL in aliquot / 5). EDTA whole blood was removed from the freezer and left to thaw for at room temperature for 1-2 hours. The appropriate *E. coli* ATCC 2955 aliquots were selected and removed from the freezer and thawed 15-30 minutes prior to set up. EDTA whole blood samples were moved to the Class 1 safety cabinet and 4.1 mL of the EDTA whole blood added to fourteen 50mL falcon tubes and labeled as per Table 2.2.

Table 2.2 EDTA whole blood spiked with *E. coli*

Replicate A	**Replicate B**	**Replicate C**	
A- 1000cfu/ml	B- 1000cfu/ml	C- 1000cfu/ml	Control
A- 100cfu/mL	B- 100cfu/mL	C- 100cfu/mL	Control
A- 10cfu/mL	B- 10cfu/mL	C- 10cfu/mL	
A- 1cfu/mL	B- 1cfu/mL	C- 1cfu/mL	

The corresponding 0.9 mL *E. coli* ATCC 2955 aliquots were added to the labelled 50mL falcon tube based on the previous calculations. Sterile molecular grade water (5 mL) was added to the

control samples. Samples were vortexed and moved to 4-12°C until ready to use to limit further bacterial multiplication.

2.2.2.2 Quantitative PCR

Quantitative PCR was used to determine the quantity of *E. coli* ATCC 2955 and host genomic DNA. The *E. coli* (all strains) uidA (Glucuronidase) Standard Kit (Primerdesign, Camberley) and the Genomic Quantification Assay (Primerdesign, Camberley) were used as per manufacturer's protocols and analysed using the Rotagene-Q platform (Qiagen, Hilden, Germany) Amplification Protocol (Enzyme activation 2 minutes at 95°C, Denaturation 10 seconds at 95 °C, Data collection 60 seconds at 60 °C) for 50 cycles.

2.3 Metagenomic next generation sequencing of targeted samples

Metagenomic sequencing and bioinformatic analysis was performed at a collaborating laboratory site under a material transfer agreement by Dr Gemma Kay, a senior researcher supervised by Professor Justin O'Grady at the Quadram Institute, University of East Anglia.

2.3.1 Sample transfer and storage

Serum and EDTA whole blood samples (stored at -80^{o}C) were couriered directly to the collaborating laboratory on dry ice to be received by a named scientist who performed the analysis described below. No patient identifiable information was transferred to the collaborating team and the collaborating research team were blinded to the diagnoses made by standard of care and which samples were collected from subsequently undiagnosed and diagnosed participants.

2.3.2 Human gDNA depletion

The collaborating research team performed step 2.3.2 to 2.3.4 inclusive. Human gDNA was depleted from both EDTA whole blood and sera using a patented, currently unpublished method devised by the collaborating research group (patent PCT/GB2017/053715). Details of the patented approach can be found at www.patentscope.wipo.int [206] but briefly, cytolysin phospholipase C obtained from *Clostridium perfringens* or a derivative of cytolysin is added to a biological sample followed by a process to physically deplete nucleic acids derived from host cells from the sample so that they are no longer detectable - this method remains under patent and full details are not yet publicly available. This method was compared to other commercially available DNA enrichment methods such as the Looxster Enrichment kit (Analytic Jena); NEBNext Microbiome DNA Enrichment kit, MolYsis Basic 5 kit (Molzym); it was also tested on EDTA whole blood prior to these experiments.

Bacterial pellets were subsequently re-suspended in 350 μl MagNApure bacterial lysis buffer (Roche, Basel, Switzerland) and 20 μl enzyme mix (6 μL mutanolysin (25ku/ml); 3 μL lysostaphin (4 ku/ml); 5 μL lysozyme (10 mg/ml), 4 μL lyticase (10 ku/ml) and; 2 μL chitinase (50 u/ml)) with RNase A and incubated at 37^{o}C for 15 min with shaking at 1000 rpm. Proteinase K (10 μL) was added and the mixture incubated for a further 10 min at 65^{o}C. DNA extraction then followed on 400 μl of sample loaded onto the MagNAPure Compact automated instrument (Roche, Basel, Switzerland) using the MagNAPure Compact DNA isolation kit using protocol version 3.2 (Roche, Basel, Switzerland). DNA was then quantified using the Qubit dsDNA high sensitivity assay

(Thermofisher, Waltham, Massachusetts) and Qubit fluorometer (Thermofisher, Waltham, Massachusetts).

2.3.3 DNA amplification

Resulting DNA samples were then whole genome amplified using the Qiagen REPLI-g single cell kit (Qiagen, Hilden, Germany) as per manufacturer's instructions. DNA was purified with AMPure XP beads (Beckman Coulter, California) at a 1:1 ratio and quantified using the Qubit fluorometer (Thermofisher, Waltham, Massachusetts) and the Qubit dsDNA broad range assay (Thermofisher, Waltham, Massachusetts). Amplified DNA Samples were then diluted to 0.2 ng/µL prior to Illumina Nextera XT library preparation (Illumina, San Diego) as per manufacturer's guidance. A 4 nM library pool was sequenced on the Illumina NextSeq platform with v2 chemistry at 2 ´ 150bp cycles.

2.3.4 Sequence analysis

Sequence reads mapping to the human genome were removed from the raw fastq files generated and the output fastq files run through Kraken to generate microbial profiles using the following data analysis commands:

To count raw reads:

```
cat input.fq | echo $((`wc -l '/4))
```

To remove human reads:

```
minimap2 -a hg38.fa read1.fq read2.fq | samtools view -bS -f 4 - | samtools sort - -o output.bam
```

To convert bam to fastq:

```
bam2fastx –fastq -o output.fq input.bam
```

To identify microbial profile using kraken:

```
Kraken –db kraken/krakendb --threads 12 --fastq-input human_removed.fq | kraken-report – db kraken/krakendb - > kraken_report.txt
```

At the time of analysis, the standard Kraken database used all complete bacterial and viral genomes available on Reference Sequence (RefSeq), including approximately 25,000 genomes.

Refseq is a comprehensive collection of sequences, which is curated by the National Center for Biotechnology Information (NCBI). These profiles were summarized per sample with the top five species, noting that the presence of C*lostridium perfringens* is due to the enzyme used which is isolated from this organism and is therefore amplified during the whole genome amplified process. Other known organisms associated with the sample processing were *Streptomyces griseus, Pseudomonas fluorescens, Alteromonas macleodii, Ralstonia pickettii, and Propionibacterium acnes.*

The anonymised metagenomic NGS data and bioinformatic analysis was provided to the primary research site and matched with individual study participants for interpretation.

2.3.5 Clinical interpretation and statistical analysis

2.3.5.1 Clinical interpretation

Initially analysis of the mNGS data included analysing proportions of human, non-human and unclassified reads in EDTA whole blood, serum and molecular grade water controls. Unclassified reads are those, which cannot be assigned to a human or non-human species present in the database used for bioinformatics analysis, key reasons for this include errors occurring in the sequence, that the organism is not represented in the database or the sequence belongs to an organism which has not previously been sequenced/recognised as a human pathogen. Samples were first analysed by sample type and then by 'febrile', 'healthy' and 'control' categories. Then the non-human reads were analysed first in full and again once known contaminants were removed. Prior to clinical correlation of the non-human reads not associated with methodological contamination, the Health and Safety Executive 'The Approved List of biological agents' [207] was reviewed to correlate the mNGS findings to known pathogens. Reads aligned to organisms not listed in the HSE document were discounted if a literature review failed to identify reports of pathogenicity in humans. The Health and Safety Executive 'The Approved List of biological agents' contains a comprehensive list of organisms responsible for causing human disease generated by the Department of Health and Social Care (DHSE) to ensure the safe handling and processing of organisms of pathogenic potential. In this study, this list was used as a reference to infer an organism's potential role in causing disease. The document covers a broad range of bacterial, viral, fungal and parasitic pathogens.

2.3.5.2 Statistical analysis

All descriptive and statistical analysis was performed using Microsoft Excel for Mac Version 16.56 2021 and GraphPad Prism Version 9.2.0 (283) software. Total sample sequence reads for EDTA

whole blood and serum were compared for statistically significant differences in human, non-human and unclassified reads using a Mann-Whitney test. Combined serum and EDTA whole blood data was compared in participants with AUFI and healthy volunteers using a Mann-Whitney test. Reads associated with methodological contaminants were identified and compared between serum and EDTA whole blood. Once reads associated with contamination were discounted, the remaining non-human reads were analysed.

Chapter 3 Clinical study of AUFI: Results

This chapter is the first of three results chapters and will describe the results of the clinical study examining a cohort of 100 adults presenting to hospital with acute undifferentiated febrile illness.

3.1 Recruitment

3.1.1 Participants with AUFI

Between 18th November 2015 and 15th July 2017, 124 adults admitted to UHS with fever and suspected infection were approached by the study team. Of those approached, 16.9% (21/124) declined participation, three did not meet inclusion criteria (2.4%, 3/124) (lacked capacity, known immunodeficiency, no history of fever) and one hundred eligible participants were recruited (see figure 3.1). Out of the eligible participants, one withdrew from study follow up but consented for their recruitment samples and clinical data to be analysed and 25% (25/100) were lost to follow up. No participant died during the study.

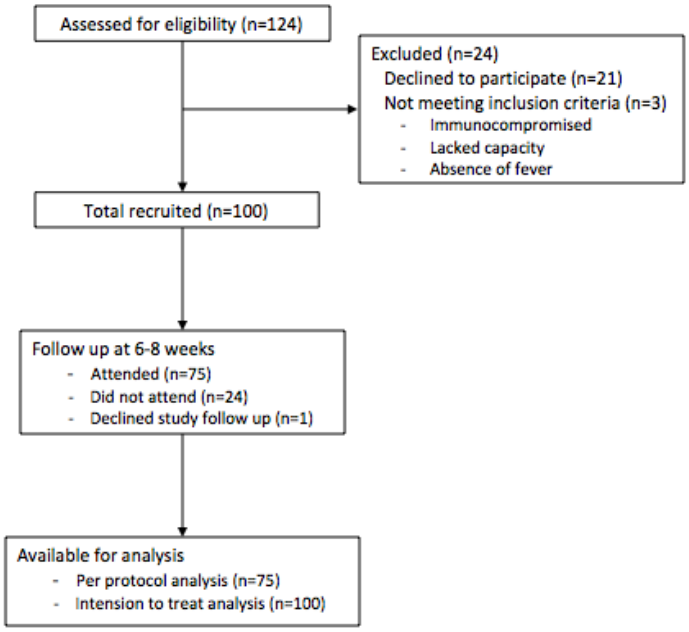


Figure 3.1 Trial profile- participants with AUFI

3.1.2 Healthy volunteers

Fifty-six healthy adult volunteers applied to take part in the study. Four did not meet inclusion criteria (comorbidities n=2, unable to commit to follow up telephone appointment n=2), 92.8% met the inclusion criteria (52/56). However, two developed symptoms of a febrile illness within their one week follow up period and were excluded, leaving 50 participants available for analysis (see figure 3.2).

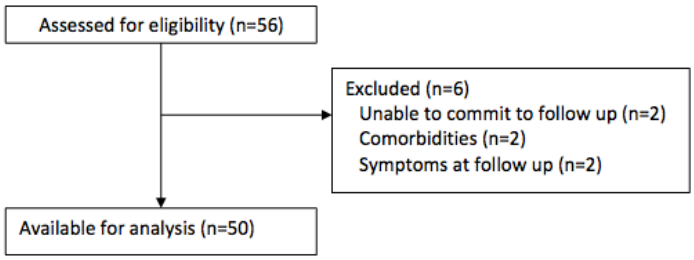


Figure 3.2 Trial profile- healthy volunteers

3.2 Participants with acute undifferentiated febrile illness

3.2.1 Demographics

Of the 100 adults recruited with AUFI, 56% (56/100) were female and the median age was 35 years (IQR 24.0-48.0; range 18.0-87.0) (see table 3.1).

Table 3.1 Demographic information for AUFI participants, n=100

Baseline Data	
Female	44 (44.0)
Male	56 (56.0)
Age	35.0, [24.0-48.0]
Ethnicity	
White British	68 (68.0)
European	9 (9.0)
Indian	8 (8.0)
Black African	4 (4.0)
Pakistani	3 (3.0)
Middle Eastern	2 (2.0)
Other	6 (6.0)
Occupation	
Health and social	13 (13.0)
Retail and motor services	12 (12.0)
Retired	12 (12.0)
Other service (e.g. domestic cleaner, caretaker)	12(12.0)
Professional	10 (10.0)
Homemaker	6 (6.0)
Transport	6 (6.0)
Construction	5 (5.0)
Student	4 (4.0)
Unemployed	4 (4.0)
Finance	4 (4.0)
Administration / Support	3 (3.0)
IT	3 (3.0)
Education	3 (3.0)
Defence	2 (2.0)
Volunteer	1 (1.0)
All data presented as n/n (%) and median [IQR]	

The majority (68%) were of 'white British' ethnic origin, followed by 'European' (9%), 'Indian' (8%) and 'other' (6%) (see table 3.1) The 'other' ethnic group was composed of three British people of Indian (2%) and Pakistani (1%) heritage, one Thai person (1%), one person of Chinese and Kenyan origin (1%) and one where ethnic origin was not specified (1%).

Over half of participants reported the following five occupations, 'health and social care' (13%), 'retail and motor services' (12%), 'retired' (12%) and 'other service' (12%) (which included roles such as caretaker, warehouse worker, pest control, beautician, refuse collector, nursery worker) and 'professional' occupations (10%) (see table 3.1 for details).

3.2.1.1 Comparison AUFI participants and healthy volunteers

There were more females in the healthy volunteer control group 60% (30/50) but the median age was similar to the AUFI group (38 years [28.5-49.5]) (see table 3.2). Similarly, the healthy volunteer group comprised of individuals from white British (74%; 37/50), European (6%, 3/50) and Indian (6%; 3/50) backgrounds (table 3.2). The healthy volunteers were mostly in 'health and social care' occupations (76%; 38/50), which is to be expected as recruitment was focussed on healthcare workers and health science students. Less comorbidity was recorded in the healthy volunteer group (median 0, range 0-3) as expected as this was part of the exclusion criteria.

Table 3.2 Healthy volunteers demographic information, n=50

Baseline Data			**Ethnicity**	
			White British	37 (74.0)
Female	30 (60.0)		European	3 (6.0)
Male	20 (40.0)		Indian	3 (6.0)
Age (years)	38 [28.5-49.5]		Chinese	1 (2.0)
Number of comorbidities			Other	4 (8.0)
Presence of any comorbidity	10 (20.0)		Unknown	2 (4.0)
Gastrointestinal Disease	2 (4.0)		**Occupation**	
Respiratory Disease	3 (6.0)		Health and social	38 (76.0)
Cardiovascular Disease	2 (4.0)		Retired	1 (2.0)
Psychological	1 (2.0)		Professional	2 (4.0)
Obstetric and Gynaecology	1 (2.0)		Student	2 (4.0)
Renal / urological disease	1 (2.0)		Administration / Support	1 (2.0)
Treated malignancy	1 (2.0)		Education	2 (4.0)
Haematological	1 (2.0)		Unknown	4 (8.0)
All data presented as n/n (%) and median [IQR]				

3.2.2 Comorbidities

Comorbidities were recorded from the participant medical history and confirmed by discussion with the participants. Most participants 77% (77/100) had at least one comorbidity (median = 1; IQR 1.0-2.0; range 1-11). The most common comorbidities were gastrointestinal disease 25% (25/100), neurological 21% (21/100) and respiratory 18% (18/100) (see table 3.3). Of the whole cohort, 8% (8/100) reported to be current smokers.

Table 3.3 comorbidities AUFI participants, n=100

Comorbidities	
Number of comorbidities	1.0 [1.0-2.0]
Presence of any comorbidity	77 (77.0)
Gastrointestinal Disease	25 (25.0)
Neurological Disease	21 (21.0)
Respiratory Disease	18 (18.0)
Cardiovascular Disease	17/98 (17.0)
Rheumatological/ Musculoskeletal	10 (10.0)
Psychological	9 (9.0)
Diabetes Mellitus	8 (8.0)
Obstetric and Gynaecology	7 (7.0)
Dermatological	7 (7.0)
Renal / urological disease	5 (5.0)
Ear, Nose, Throat Disease	5 (5.0)
Malignancy	3 (3.0)
Haematological	2 (2.0)
Cerebrovascular Disease	1 (1.0)
Other [1]	7 (7.0)
All data presented as n/n (%)and median [IQR], [1]glaucoma, hay fever, cold sores, pectus excavatum, surgical breast augmentation, vitamin D deficiency and folic acid deficiency	

3.2.3 Travel History

Of the 100 participants with AUFI, 49.0% (49/100) had a history of travel within the three months prior to presentation. Of those 77.6% (38/49) had travelled in the previous four weeks although documentation on the CRF was not clear for 2% (1/49). The median interval between travel (leaving destination) and illness was nine days (IQR 1.0-28.3; range -3.0-61.0). The most popular destination region was Europe (46.9%; 23/49) followed by South Central Asia (24.5%; 2/49) and Sub-Saharan Africa (20.4%; 10/49) (table 3.4). More than one destination of travel was recorded in 16.3% (8/49) of participants with 6.1% (3/49) travelling to four destinations and 10.2% (5/49) travelling to two destinations. The remaining 83.7% (41/49) travelled to a single destination.

Tourism, (55.1%; 27/49), visiting friends and relatives, (44.9%; 22/49) and business, (20.4%; 10/49), were the most common reasons for travel (see table 3.4). Within the 'reason for travel' group the 'other' group comprised one person travelling to a religious retreat and one where the reason for travel was not clearly documented on the CRF (see table 3.4).

Table 3.4 Travel history in participants with AUFI, n = 100

Travel	**median [IQR]**
Interval between travel and illness onset (days)	9.0 [1.0-28.3]
Recent Travel	
Travelled abroad within 3 months of onset of illness	49 (49.0)
Travelled abroad within 4 weeks of onset of illness	38 (38.0)
Region of Travel (refers to 3 month data)	**n= 49 (%)**
Europe	23 (46.9)
South Central Asia	12 (24.5)
Sub-Saharan African	10 (20.4)
South East Asia	8 (16.3)
North Africa /Middle East	3 (6.1)
South and Central America	3 (6.1)
USA/Canada	2 (4.1)
Caribbean	2 (4.1)
Oceania	0 (0.0)
Reason for travel	**n= 49 (%)**
Tourism	27 (55.0)
Visiting friends and relatives	22 (44.9)
Business	10 (20.4)
Volunteer/aid Work	3 (6.1)
Research/Education	1 (2.0)
Other	2 (4.1)
Number of destinations	**n= 49 (%)**
Single destination	41 (83.7)
Two destinations	5 (10.2)
Four destinations	3 (6.1)
All data presented as n/n (%) and median [IQR],	

3.2.4 Presenting clinical features of AUFI Participants

Clinical symptoms and signs were categorised into the following eight groups; systemic symptoms, respiratory, gastrointestinal, musculoskeletal, neurological/ophthalmological, cardiovascular, dermatological and genitourinary. Of the systemic symptoms, fever was reported or documented in all participants being one of the study inclusion criteria. Fever duration data was missing in 1% (1/100) and the median duration of fever at presentation in the remaining 99% (99/100) was 4.0 days (IQR 2.0-5.5, Range 1.0-14.0) (see table 3.5). The most common constitutional symptoms besides fever were lethargy (89%; 89/100), night sweats/chills (80%; 80/100) and loss of appetite (64%; 64/100).

Table 3.5 Systemic symptoms in participants with AUFI, n=100

Systemic symptoms- participants with AUFI	**Median [IQR]**
Fever duration at presentation (days)	4.0 [2.0-5.5]
Any systemic symptom	**100 (100.0%)**
Fever	100 (100.0)
Lethargy	89 (89.0)
Night sweats/ chills	80 (80.0)
Loss of Appetite	64 (64.0)
Lymphadenopathy	35 (35.0)
Weight loss	11 (11.0)
All data presented as n/n (%) and median [IQR]	

Neurological symptoms were reported in 84% (84/100), with the majority of these being headache (96.4%; 81/84) followed by photophobia (36.9%; 31/84) and neck stiffness (25%; 21/84), with only 4.8% (4/84) reporting focal neurology (see table 3.6).

Table 3.6 Neurological symptoms in participants with AUFI, n=100

Any neurological symptom	**84 (84.0%)**
Headache	81 (81.0)
Photophobia	31 (31.0)
Neck Stiffness	21 (21.0)
Conjunctivitis	20 (20.0)
Focal neurology	4 (4.0)
Seizures	0 (0.0)
All data presented as n/n (%)	

Gastrointestinal and respiratory symptoms were reported equally commonly, both being reported in 70% (70/100) of all participants with AUFI. Of the gastrointestinal symptoms and signs reported (see table 3.7), vomiting 54.3% (38/70) was most common followed by abdominal pain 50% (35/70) and diarrhoea 35.7% (25/70). Jaundice, splenomegaly and hepatomegaly were unusual symptoms and were reported equally frequently in 4.3% (3/70) of individuals.

Table 3.7 Gastrointestinal symptoms- participants with AUFI n=100

Any gastrointestinal symptom	**70 (70.0%)**
Vomiting	38 (38.0)
Abdominal pain	35 (35.0)
Diarrhoea	25 (25.0)
Jaundice	3 (3.0)
Hepatomegaly	3 (3.0)
Splenomegaly	3 (3.0)
All data presented as n/n (%)	

The most common respiratory symptom reported was cough 60% (42/70), followed by sore throat 52.9% (37/70) and coryza 35.7% (25/70) (see table 3.8).

Table 3.8 Respiratory symptoms in participants with AUFI, n=100

Any respiratory symptom	**70 (70.0%)**
Cough	42 (42.0)
Sore throat	37 (37.0)
Coryza	25 (25.0)
Shortness of breath	20 (20.0)
Pleuritic chest pain	9 (9.0)
Wheeze	7 (7.0)
All data presented as n/n (%)	

Musculoskeletal symptoms were reported in 53% (53/100) with 75.5% (40/53) of those reporting myalgia and 62.3% (33/53) reporting arthralgia. Rash was present in 31% (31/100), most commonly a macular rash 77.4% (24/31), followed by vesicular 16.1% (5/31) (see table 3.9)

Table 3.9 Musculoskeletal and dermatological symptoms in participants with AUFI, n=100

Any Musculoskeletal	**53 (53.0%)**
Myalgia	40 (40.0)
Arthralgia	33 (33.0)
Any dermatological	**31 (31.0%)**
Macular	24 (24.0)
Vesicular	5 (5.0)
Petechial	2 (2.0)
Other [1]	3 (3.0)
All data presented as n/n (%), [1] urticaria (1), maculopapular (1) and papular rash (1)	

Genitourinary symptoms were infrequently reported 13% (13/100), most commonly dysuria 53.8% (7/13) followed by other 38.5% (5/13) and urinary frequency 7.7% (1/13). Of those with dysuria, 0% (0/7) had a positive urine culture, 57.1% (4/7) had <10 white cells on microscopy, 28.6% (2/7) had 10-20 white cells and no bacterial growth and 14.3% (1/7) had >20 white cells on and bacteria seen on microscopy but no growth and a clinical presentation compatible with urinary tract infection. Cardiac symptoms were unusual 8% (8/100), these were attributable to palpitations 75% (6/8) and chest pain 50% (4/8) (see table 3.10)

Table 3.10 Cardiovascular and genitourinary symptoms in participants with AUFI, n=100

Any Cardiovascular	**8 (8.0%)**
Palpitations	6 (6.0)
Chest pain	4 (4.0)
Any Genitourinary	**13 (13%)**
Dysuria	7 (7.0)
Frequency	1 (1.0)
Genital discharge	0 (0.0)
Genital ulceration	0 (0.0)
Other[1]	5 (5.0)
All data presented as n/n (%),[1] Testicular pain (1), urinary hesitancy (1), bilateral flank pain (1), haematuria (1), painful external genitalia (1).	

3.2.5 Physiological and Laboratory Parameters at Presentation

Admission physiological observations and blood results were collected for all participants. The analysis below was performed and where data was missing this is highlighted.

Table 3.11 Physiological parameters of participants with AUFI, n=100

Physiological parameter	Reference Range	Median [IQR]	Range
Temperature (°C)	36.5-37.5	38.2 [37.1-38.8][1]	35.6-40.9[1]
Systolic blood pressure, (mmHg)	90.0-120.0	120.5 [110.8-135.5][1]	89.0-209.0[1]
Diastolic blood pressure, (mmHg)	60.0-80.0	70.0 [60.0-77.3][1]	40.0-97.0[1]
Heart rate (beats per minute)	60-100.0	100.0 [85.0-110.0][1]	60.0-160.0[1]
Oxygen saturations (%)	95.0-100.0	97.0 [96.0-99.0] [2]	91.0-100.0[2]
Respiratory rate (breaths per minute)	12-16	18.0 [16.0-20.0][3]	11.0-29.0[3]
[1]Evaluated in 96 participants, [2]Evaluated in 95 participants, [3]Evaluated in 92 participants			

Median body temperature was 38.2°C [IQR 37.1-38.8]. Median blood pressure readings were systolic 120.5 mmHg [IQR 110.8-135.5] and diastolic 70.0mmHg [IQR 60.0-77.3]. The median heart rate was slightly elevated at 100 bpm (IQR 85.0-110.0). Median oxygen saturations were normal in 97% (IQR 96.0-99.0, range 91.0-100.0) with only 3% (3/95) of participants on supplementary oxygen (see table 3.11).

The admission blood test data presented below (table 3.12) is complete for all 100 participants except platelets where 1% (1/100) of samples were clumped and could not be analysed and 29% (29/100) participants who did not have a lactate sample taken. The median total WCC was 10.0 $x10^9$/L [IQR 6.0-14.0] with lymphocytes of 1.0 $x10^9$/L [IQR 1.0-1.0] and neutrophils of 8.0 $x10^9$/L [IQR 3.0-12.0]. Median platelet levels were normal at 202.0 $x10^9$/L [IQR 154.0-28.0], median CRP was moderately elevated at 71.0 mg/L [IQR 23.0-159.0]. Median lactate levels were 1.0 mmol/L [IQR 1.0-2.0]. The median alanine transaminase (ALT) was 38.0 U/L, however, the range was very wide (IQR 21.0-78.0, range 8-2251.0), in particular due to one patient with sero-negative autoimmune hepatitis with an ALT of 2251.0 U/L and one with acute hepatitis E infection with and ALT 1924.0 IU/L

Table 3.12 Laboratory parameters at presentation in participants with AUFI, n=100

Laboratory parameter	Reference	Median	Range
Haemoglobin (g/dL)	115.0-180.0	140.0 [131.0-151.0]	101.0-180.0
White cell count (x10^9/L)	3.6-11.0	10.0 [6.0-14.0]	1.4-28.6
Lymphocytes (x10^9/L)	1.0-4.0	1.0 [1.0-1.0]	0.1-8.5
Neutrophils (x10^9/L)	1.8-7.5	8.0 [3.0-12.0]	0.2-26.9
Eosinophils (x10^9/L)	0.1-0.4	0.0 [0.0-0.1]	0.0-0.6
Platelets (x10^9/L)	140.0-400.0	202.0 [154.0-288.0][1]	67.0-584.0[1]
C-reactive protein (mg/L)	0.0-7.4	71 [23.0-159.0]	1.0-346.0
Sodium (mmol/L)	113.0-146.0	136.0 [134.0-137.0]	127.0-143.0
Potassium (mmol/L)	3.5-5.3	3.8 [3.7-4.1]	2.8-4.8
Creatinine (umol/L)	80.0-115.0	73.5 [64.0-92.3]	44.0-181.0
Urea (mmol/L)	2.5-7.8	4.2 [3.3-5.6]	2.3-13.3
Lactate (mmol/L)	0.5-2.2	1.0 [1.0-2.0] [2]	0.6-4.1[2]
Albumin (g/L)	35.5-50.0	38.0 [34.0-41.0]	28.0-58.0
Bilirubin (umol/L)	0.0-20.0	11.0 [8.0-17.0]	3.0-287.0
Alkaline phosphatase (U/L)	30.0-130.0	90.0 [73.0-116.0]	27.0-381.0
Alanine Transaminase (ALT) (IU/L)	10.0-40.0	27.0 [17.0-55.5]	7.0-2251.0
[1]Evaluated in 99 participants, [2]Evaluated in 71 participants			

3.2.6 Investigations of all participants with AUFI

3.2.6.1 Laboratory investigations

Tables 3.13 to 3.19, and figures 3.3 and 3.4 summarise the microbiological investigations of the participants who presented with AUFI. The investigations are grouped together by 'bacteriology' (table 3.13 and 3.14), 'virology and serology' (tables 3.15-3.17 and figures 3.4 and 3.4) and 'CSF, tissue, EDTA whole blood and other microbiological investigations' (table 3.18 and 3.19).

3.2.6.1.1 Bacterial investigations

A total of 242 bacterial investigations were carried out (2.42 bacterial investigations per AUFI participant) but only 6% (15/238) contributed to the final diagnosed.

Table 3.13 Results of bacteriological Investigations of participants with AUFI, n=100

Bacteriological investigations	
Total number of Bacterial investigations performed	242
Contributed to Diagnosis n= (%)	15 (6.0%)
Bacterial investigations per participant	2.4
Sample type	
Blood cultures	**99 (99%)**
Significant positive	2 (2.0)[1]
Contamination with skin flora	4 (4.0)
Negative	93 (93.9)
Midstream Urine White Cell Count (WCC)[2]	**92 (92%)**
>20	3/92 (3.0)
10-20	8/92 (8.7)
<10	81/92 (88.0)
Midstream urine culture	**92 (92%)**
Positive[3]	6/92 (6.5)[4]
Negative	86/92 (93.5)
Stool culture	**16 (16%)**
Positive	5/16 (31.3)[5]
Negative	11/15 (68.8)
Stool *Clostridioides difficile* PCR	**14 (14%)**
Positive	1/14(7.1)
Negative	13/14 (92.9)
Stool Norovirus PCR	**9 (9%)**
Positive	1/9 (11.1)
Negative	8/9 (88.9)
Bacterial throat swab culture	**31 (31%)**
Positive[6]	1/31(3.2)
Negative	30/31 (96.8)
Sputum sample	**4 (4%)**
Positive[7]	1/4 (25.0)
Negative	3/4 (75.0)

All data presented as n/n (%)[1] *E. coli* (1), *Pseudomonas aeruginosa* (1), [2]per high power field, [3]bacterial growth x10^9 CFU per mL)[4] *Enterococcus* species 2% (2/92), *E. coli* 4% (4/92), [5] *Shigella sonnei* (1), *Campylobacter species* (2), *Plesiomonas species* (1), *Salmonella species* (1), [6] (1) *Lancefield Group C Streptococcus*,[7] (1) yeast species

Blood cultures were performed in almost all participants, (99%; 99/100). Of the blood cultures performed, 93.9% (93/99) were negative blood (one or more sets), 4% (4/99) grew skin contaminants in one or more blood culture bottles (four coagulase negative *staphylococci* and one Gram positive bacilli identified as 'diphtheroid') and only 2% (2/99) had significant organisms isolated (*E. coli* and *Pseudomonas aeruginosa*) (see table 3.13).

Bacterial throat swabs for culture were performed in 31% (31/100) of participants. Diagnostic yield was low with only one positive culture 3.2% (1/31) (Lancefield Group C *Streptococcus*) which was not felt to be clinically significant as the clinical presentation was not consistent with this organism, a known coloniser of the oropharynx (see table 3.13).

The positive yield from MSU sample was low with only 7% (6/92) culturing potentially significant isolates. Additionally, most samples did not have significant pyuria (88%; 81/92). Of those participants with culture positive MSU samples, urinary tract infection was clinically deemed to be the cause of the presenting illness in only 33% (2/6), one being diagnosed with urinary tract associated sepsis and the other with pyelonephritis.

Stool samples 16% (16/100) were collected infrequently (16%; 16/100), however, when a stool culture was performed it was positive in almost one-third, 31% (5/16), *Clostridioides difficile* testing was performed in 14% (14/100) and was positive in one participant 7% (1/14). Of the nine stool samples sent for viral PCR, one was positive for Norovirus. All isolates identified from stool samples were deemed to be clinically significant. Sputum samples were rarely obtained (4/100) and none contributed to a confirmed diagnosis. Table 3.14 summarises the contribution of 'bacterial investigations', which contributed to the final clinically credible diagnosis.

Table 3.14

Bacterial investigations contributing to final diagnosis	
E. coli (n= 1 bacteraemia, n=3 urine culture)	4
Campylobacter species (stool culture)	2
Mycoplasma pneumonia (serology and PCR)	2
Plesiomonas shigelloides (stool culture)	1
Salmonella enteritidis (stool culture)	1
Clostridiodes difficile (stool culture)	1
Neisseria meningitidis serogroup B (blood PCR)	1
Shigella sonnei (stool culture)	1
Pseudomonas aeruginosa (blood cultures)	1
UTI (microscopy)	1

3.2.6.1.2 Virological and serological investigations

A total of 918 virological and serological investigations (9.2 investigations per participant) were performed but only 35 (4%) contributed to the final diagnosis (table 3.15) as compared with 6% (12/238) of bacterial investigations.

Table 3.15 Summary of virology, serological and tropical investigations in participants with AUFI n=100

Virological / serological investigations summary table	
Total number of virological/ serological investigations performed	918
Contributed to Diagnosis n (%)	35 (4.0)
Virological / Serological investigations per participant	9.2

Serological testing for Human immunodeficiency virus, Hepatitis B virus (HBV) and Hepatitis C virus was performed for almost all participants (see figure 3.3). One participant (1/97) was incidentally found to have active Hepatitis B virus infection (HBCore +ve, IgM +ve, HBeAg -ve, HBeAb +, DNA 225147 copies/mL) but this was not felt to be the cause of AUFI in this participant.

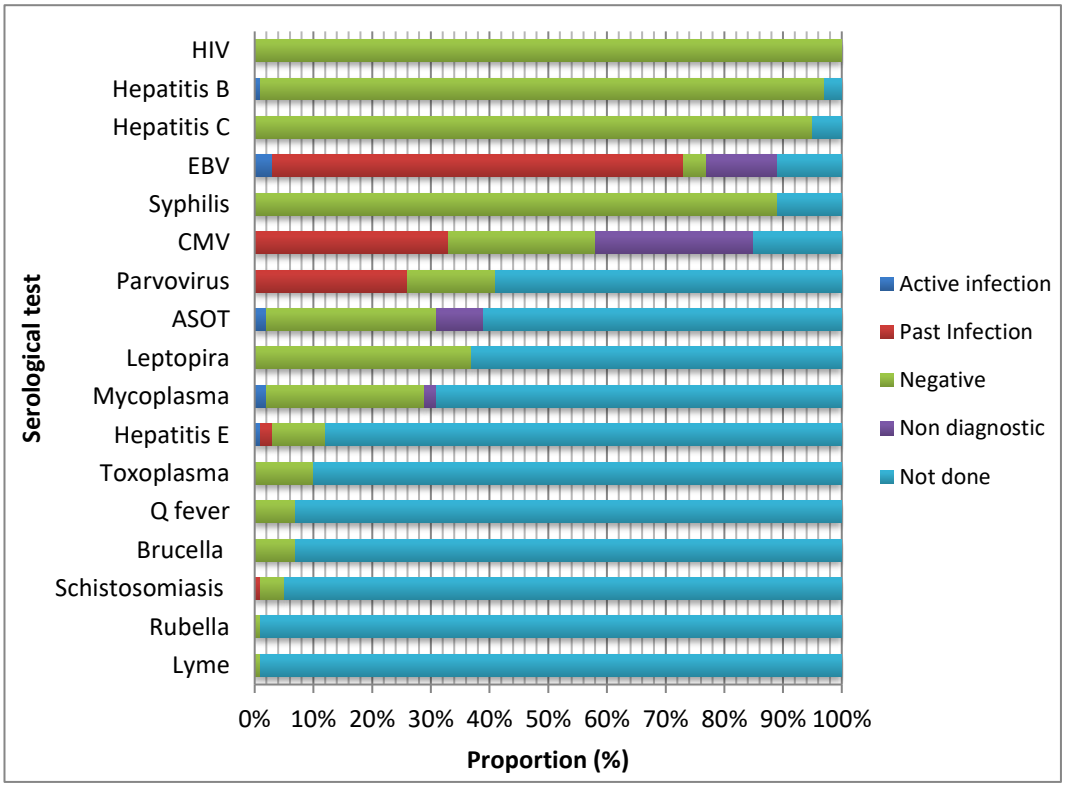


Figure 3.3 Summary of serological tests in participants with AUFI

The majority of EBV serological testing revealed past infection (88.7%; 70/89) as would be expected in an adult population. Three participants (3.4%; 3/89) had serology compatible with

acute EBV infection. One of these participants had EBV DNA detected from an EDTA blood sample in addition. The seroprevalence for past exposure to CMV was lower, (39%; 33/85) with no cases of acute infection detected.

Anti-streptolysin O (ASO) titres were raised consistent with acute infection with Lancefield Group A streptococcus in two participants (5.1%; 2/39), one with acute pharyngitis and a fever and one with acute pericarditis). Eight individuals, (20.5%; 8/39) had non-diagnostic ASOT results, due to a single result of 200 IU/ml (87.5%; 7/8) or a failure of the titre to rise in a paired, convalescent sample (12.5%; 1/8). All Leptospira serology and DNA testing was negative including one low-level positive IgM, which was negative on repeat testing. *Mycoplasma pneumoniae* serology was supportive of acute infection in 6.5% (2/31), one participant also had Mycoplasma DNA detected on throat swab, and the second demonstrated a greater than two-fold rise in titres at follow up and a low-level positive IgM. Mycoplasma serology results were non-diagnostic in 6.5% (2/31), one due to a single sample and the second due to a failure of the titre to rise in a paired, convalescent sample. Serological testing for Hepatitis E was not-often requested, (12%; 12/100) but one case of acute infection was detected, (8.3%; 1/12). Toxoplasma serology was requested in 10% (10/100) and was negative in the majority 90% (9/10) with one participant's sample reported as 'equivocal' which did not confirm on repeat testing, likely reflecting a cross-reactive sample as opposed to a true positive result.

Schistosomiasis serology was performed in 5% (5/100) and was positive in one participant at a low level (20%; 1/5). The clinical history, examination and investigations were not supportive of acute complications of schistosomiasis and the serology was deemed to reflect previously treated infection.

Over a quarter of participants had sufficient travel history or exposures to warrant a specific geographic pathogen panel (see table 3.16). The geographic pathogen panel is determined by the travel and exposure history of the participant. The panel is performed at the Rare and Imported Pathogens Laboratory (RIPL), Porton Down and consists of a collection of PCR based and serological tests of EDTA blood, serum, urine and occasionally CSF as dictated by the pathogens the individual may have encountered. This panel was performed in 27% (27/100) of participants. The majority of the (27/100; 27%) participants tested were negative (21/27; 78%), 19% (5/27) were positive for acute Dengue virus infection and one, (1/27; 4%)), was positive for acute Chikungunya virus infection (with both IgM and RNA detected in blood). The pathogens detected in these six participants were deemed to be the cause of AUFI in all cases. In addition, malaria rapid diagnostic tests (RDT) and thin and thick blood films were performed in those who had travelled to or through malaria endemic countries or that had previously resided in a country with

Plasmodium vivax or Plasmodium ovale species of malaria (33%; 33/100). All malaria diagnostic tests were negative.

Table 3.16 Characteristics of participants with AUFI tested with geographic pathogen panel, n=27

Characteristics of Participants tested with geographic pathogen	**27 (27)**
Female	11 (40.7)
Age (years)	40.0 [30.0-51.0]
Comorbidities	1.0, [1.0-2.0]
Duration of fever at presentation (days)	4.0, [3.0-6.0]
Interval between travel and presentation (days)	3.5 [1.0-9.0]
Region of Travel	
Sub-Saharan African	9 (33.3)
South Central Asia	8 (29.6)
South East Asia	7 (25.9)
Europe	5 (18.5)
South and Central America	2 (7.4)
USA/Canada	2 (7.4)
Caribbean	2 (7.4)
North Africa /Middle East	1 (3.7)
Oceania	0.0 (0.0)
Final diagnosis	
Unknown	14.0 (51.9)
Viral infection[2]	8.0 (29.6)
Bacterial infection[3]	4.0 (14.8)
Non-infection (silicosis from ruptured breast implant)	1.0 (3.7)
Outcome and Length of Stay	
Admission to general intensive care (GICU)	1.0 (3.7)
Died during admission	0.0 (0.0)
Readmission within 30 days	0.0 (0.0)
Died during follow up	0.0 (0.0)
All data presented as n/n (%) and median [IQR], **[2]** Dengue n=5, Influenza A n=2, Chikungunya n=1), [3] *Campylobacter* n=1, *Plesiomonas* n=1, *Shigella sonnai* n=1, rectal abscess n=1	

Nasopharyngeal swabs for respiratory multiplex PCR testing were collected in 95% (95/100) of participants. A pathogen was detected in 23.1% (22/95) of all respiratory PCR tests. The most commonly identified pathogens were Influenza A (8.4%; 8/95) and RSV (4.2%; 4/95) see figure 3.4). More than one virus was detected in two participants (2.1%; 2/95) one individual had

Parainfluenza virus 3 and Rhinovirus detected and the other had Influenza A and RSV detected (not included in totals for individual viruses).

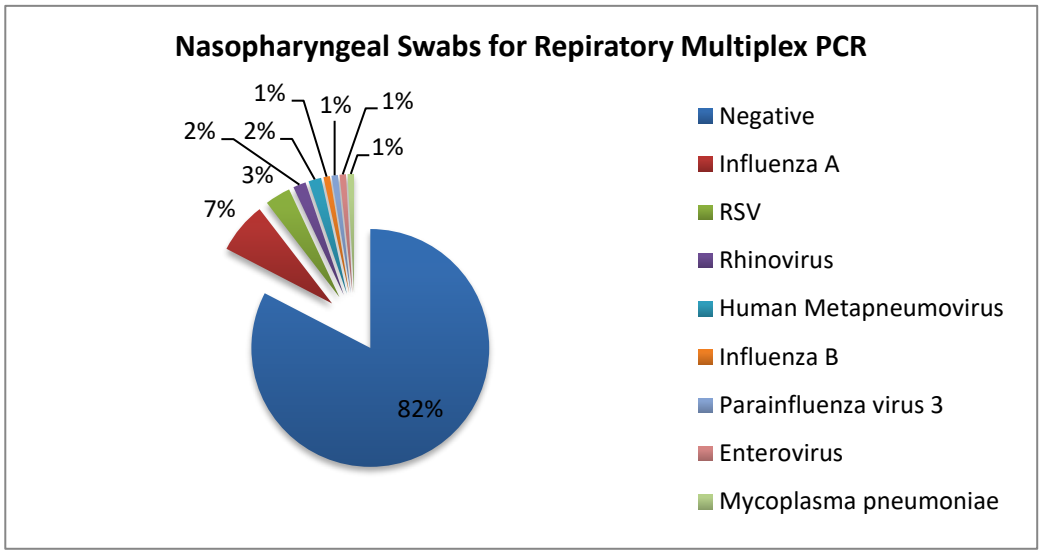


Figure 3.4 Results of respiratory multiplex PCR on nasopharyngeal swabs performed on participants with AUFI (n=95)

Of the pathogens detected by respiratory PCR, twelve results contributed to the final, clinically credible diagnosis (9 Influenza A, 1 Enterovirus, 1 Parainfluenza virus 3, 1 influenza B). Respiratory viral PCR contributed to 25% (12/48) of all final diagnoses.

Table 3.17 summarises the diagnoses made with virological and serological investigations including respiratory PCR and tropical geographical panel.

Table 3.17

Infections identified by virological and serological standard diagnostic testing	
Influenza A (Respiratory PCR)	9
Dengue virus (Serology and PCR)	5
Epstein barr virus (serology +/- PCR)	3
Streptococcus pyogenes (ASO + clinical history)	2
Mycoplasma pneumonia (serology and PCR)	2
Hepatitis E virus (serology and PCR)	1
Parainfluenza virus 3 (Respiratory PCR)	1
Norovirus (Stool)	1
Chikungunya virus (EDTA + serology)	1
Influenza B (Respiratory PCR)	1
Enterovirus (Respiratory PCR)	1

3.2.6.1.3 CSF, tissue, EDTA whole blood and other microbiological investigations microbiological investigations in participants with AUFI

Microbiological investigations which do not fit in the previously categories have been grouped together in table 3.18. Overall, there were 58 CSF, tissue, EDTA whole blood and 'other' investigations performed (0.58 per participant) and 17.2% (10/58) of these contributing to a clinically credible diagnosis.

Of sixteen CSF samples collected, bacterial cultures were negative in all (100%; 16/16). A CSF white cell count greater than five cells per mm^3 was seen in 31.3% (5/16) but PCR testing for HSV, VZV and Enterovirus was only positive in 18.8% (3/16), all of which were Enterovirus positive (Echovirus 6, Coxsackie virus B and one non-typeable Enterovirus due to low sample volume) in participants with features consistent with viral meningitis. Four participants had targeted CSF PCR testing for *Neisseria meningitidis* 25.0% (4/16) and one for *Streptococcus pneumoniae* 6.3% (1/16) all of which were negative. Overall, 18.8% (3/16) contributed to a clinically credible diagnosis.

Table 3.18 CSF, tissue, EDTA whole blood, other investigations, participants with AUFI, n=100

Total number investigations performed	58.0
Contributed to Diagnosis	10 (17.2)
Investigations per participant	5.8
Sample Type	
Cerebrospinal fluid examination	**16 (16)**
WCC < 5, negative PCR	9/16 (56.3)
WCC < 5 PCR not done	1/16 (6.3)
WCC < 5 Enterovirus PCR positive	1/16 (6.3)
WCC > 5 negative PCR	3/16 (18.8)
WCC > 5 Enterovirus PCR positive	2/16 (12.5)
Positive culture	0/16 (0.0)
Negative culture	16/16 (100.0)
EDTA blood sample PCR	**15 (15)**
Positive (1= Hepatitis E RNA, 1= *Neisseria meningitidis* DNA, 1= EBV DNA, 1=Enterovirus RNA)	4/15 (26.7)
Negative	11/15 (73.3)
Tissue Biopsy	**5 (5)**
Abnormal (2= lymph nodes; reactive, silicone deposits 2= skin; lymphoma, chronic dermatitis)	4/5 (80.0)
Normal	1/5 (20.0)

***Mycobacterium tuberculosis* diagnostic testing**	**6 (6)**
Positive	0/6 (0.0)
Negative	6/6 (100.0)
Skin swab VZV / HSV PCR	**7 (7)**
Positive (all HSV)	3/7 (42.9)
Negative	4/7 (47.1)
Monospot (Heterophile antibody)	**1 (1)**
Positive (serology and PCR consistent with acute EBV infection)	1/1 (100.0)
Negative	0/1 (0.0)
Bacterial culture on wound swabs performed	**3 (3)**
Positive (1= 'right thigh' methicillin sensitive *S. aureus* and *Pseudomonas aeruginosa* 1=neck *Cryseobacterium indologenes*)	2/3 (66.7)
Negative	1/3 (33.3)
All data presented as n/n (%)	

Tissue biopsies were performed in very few participants 5.0% (5/100) and contributed to a clinically credible diagnosis in two cases 40.0% (2/5) (Angioimmunoblastic T cell Lymphoma, and Silicone deposits from ruptured breast implants). A range of specimen types were sent (tongue 20.0% (1/5), lymph node 40.0% (2/5), skin 40.0% (2/5)). One participant (Angioimmunoblastic T cell Lymphoma) had multiple skin biopsies, lymph node biopsies and bone marrow biopsies but has been classified as 'skin' as this investigation led to the final diagnosis.

Seven participants had swabs of skin lesions tested for HSV and VZV 7.0% (7/100) and 42.9% (3/7) of these were positive for HSV. In all participants it was deemed to be HSV reactivation due to an inter-current illness as opposed to the principal cause of fever.

Investigations for *Mycobacterium tuberculosis* (including microscopy, culture and PCR) were performed in 6.0% (6/100) and were negative in all cases. Skin swabs for bacterial culture were taken in three participants 3.0% (3/100) of the sites 'right thigh', 'wound' and 'neck'. Two 66.6% (2/3) swab cultures were positive for bacterial with the swab of 'right thigh' growing methicillin sensitive *S. aureus* and *Pseudomonas aeruginosa* and the swab of the 'neck' growing *Cryseobacterium indologenes.* The bacterial skin swab results did not contribute to the final diagnosis in either patient.

A Monospot (Heterophil antibody test) was performed in 1.0% (1/100) and positive in this participant. The positive monospot result was taken from a participant with serologically confirmed acute EBV infection and was consistent with the final diagnosis.

Fifteen 15.0% (15/100) participants had EDTA blood samples tested by PCR for viral nucleic acid (see table 3.18), 27.0% (4/15) of these were positive. Of these, one participant's sample 6.7% (1/15) was positive for Hepatitis E RNA, this individual also had elevated Hepatitis E IgM and IgG levels consistent with acute Hepatitis E infection. One 6.7% (1/15) participant had EBV DNA detected in EDTA in addition to elevated IgM levels in their serum consistent with acute infection with EBV. One 6.7% (1/15) sample was positive for *Neisseria meningitidis* serogroup B DNA by PCR which was consistent with the clinical diagnosis of acute meningococcal septicaemia. A further participant 6.7% (1/15) had Enterovirus RNA detected in an EDTA blood sample and also in their respiratory PCR swab sample consistent with a diagnosis of acute enterovirus infection.

Table 3.19 summarises all diagnoses made from CSF, EDTA whole blood, tissue biopsies and PCR of EDTA blood samples.

Table 3.19

Diagnoses identified by CSF, tissue, EDTA whole blood and 'other' standard diagnostic investigations testing	
Enterovirus (n=3 CSF PCR, n=1 EDTA PCR)	4
Hepatitis E virus (PCR, also on serology)	1
Neisseria meningititis (PCR)	1
Epstein barr virus (monospot, also by PCR and serology)	1
Immunoblastic T cell lymphoma (skin biopsy)	1
Silicone deposits (lymph node biopsy)	1

3.2.6.1.4 Radiological Investigations of all patients with AUFI

Data was collection on all radiological investigations performed on participants with AUFI (table 3.20) Overall, 183 radiological investigations were performed (median of 1.8 per participant) but only 5.4% (10/183) contributed to a clinically credible diagnosis.

Table 3.20 Radiological investigations in participants with AUFI, n=100

Radiological investigations	
Total number of Radiological investigations performed	183.0
Contributed to Diagnosis	10.0 (5.4)
Radiological investigations per participant	1.8
Modality	
Chest X-ray performed	**90 (90%)**
Non-infective changes	8/90 (8.9)
Focal consolidation	7/90 (7.8)
Basal Atelectasis	5/90 (5.6)
Hilar prominence	2/90 (2.2)
Pleural thickening	1/90 (1.1)
Normal	67/90 (74.4)
Abdominal Ultrasound (USS)	**34 (%)**
Hepatobiliary abnormality	9/34 (26.4)
Hepatobiliary abnormality & Splenomegaly	6/34 (17.6)
Renal tract abnormality	6/34 (17.6)
Splenomegaly	2/34 (5.9)
Normal	11/34 (32.4)
Ultrasound other site	**7 (%)**
Abnormal[1]	4/7 (57.1)
Normal[2]	3/7 (42.9)
CT Imaging	**35 (%)**
Abnormal[3]	16/35 (45.7)
Normal[4]	19/35 (54.2)
MRI imaging	**6 (%)**
Abnormal[5]	2/6 (33.3)
Normal[6]	4/6 (66.7)
Transthoracic Echo (TTE)	**7 (%)**
Abnormal	5/7 (71.4)
Normal	2/7 (28.6)
Abdominal X Ray (AXR)	**4 (%)**
Normal	4/4 (100.0)

All data presented as n/n (%) and median [IQR], [1]USS neck n=4 (reactive lymph nodes n=1, thyroid nodule n=1, consistent with foreign material n=1, consistent with lymphoma n=1),[2]USS neck n=1, testicular n=1, knees and elbows n=1, [3]CT CAP/CT AP n=10, CT head n=5, CT neck n=1,[4]CT CAP/CT AP n=3, CT head n=16, [5]MRI spine n=1, MRI liver n=1, [6]MRI head n=4

Chest Radiographs (CXR) were performed for most participants (90%; 90/100) with the majority reported as normal by a radiologist (74%; 67/90). Non-infective changes such as cardiomegaly and bony abnormalities were identified most frequently, (9%; 8/90), followed by focal consolidation (7.7%; 7/90) and basal atelectasis (5.6%; 5/90) (see table 3.20). Of those with focal consolidation, three (3/7) had confirmed viral and bacterial pathogens (Influenza A H1N1, *Mycoplasma pneumonia* and Rhinovirus), although Rhinovirus would not typically be associated with focal consolidation, it may have contributed to a secondary bacterial co-infection. Of the remaining participants with focal consolidation on CXR (4/7), one participant was found to have silicosis from a ruptured breast implant (1/7) (with diagnostic lymph node histology) one lymphoma (skin biopsy confirming angioimmunoblastic T cell lymphoma) (1/7), and one participant was diagnosed with a urinary tract infection (1/7). The CXR changes in the final participant with focal consolidation were not deemed sufficient to diagnose a lower respiratory tract as this participant had no respiratory tract symptoms, normal saturations and no oxygen requirement (1/7). Four participants had an abdominal X-ray (4/100; 4%) all of which were reports as normal, and did not contribute to a definitive diagnosis.

Approximately one third (34/100; 34%) of participants had an abdominal ultrasound scan USS. Of these 32.4% (11/34) showed no abnormality. Of all the abdominal USS performed, only one 2.9% (1/34) contributed to the clinical management detecting a renal cell carcinoma (RCC) which was an incidental finding in a participant with a perianal abscess. It is likely the acute abscess was the cause of fever and in this participant as they presented with acute symptoms. The remaining abdominal USS 97.1% (33/34) may have contributed to the differential diagnosis in excluding significant pathology but none provided a conclusive diagnosis. Interestingly, in one participant 2.9% (1/34) USS abdomen revealed only splenomegaly whereas a follow up computed tomography imaging of the chest, abdomen and pelvis (CT CAP) demonstrated a multifocal Hepatocellular carcinoma and liver cirrhosis demonstrating the higher diagnostic sensitivity of CT over USS for intra-abdominal pathology.

Ultrasound of other sites was performed in 7% (7/100) (see table 3.20) the majority were USS neck scans 71.4% (5/7) results included 20% (1/5) incidental thyroid nodule, 20% (1/5) reactive lymph nodes, 20% (1/5) unhelpful, 20% (1/5) consistent with foreign material, 20% (1/5) consistent with lymphoma, the non-abdominal ultrasounds performed 28.7% (2/7) were diagnostic.

Thirteen 13% (13/100) participants had either a CT scan of their chest, abdomen and pelvis (CT CAP) or a CT scan of their abdomen and pelvis (CT AP) performed. The diagnostic yield was high with 76.9% (10/13) being reported as abnormal and 50% (5/10) of abnormal scans contributing

directly to the diagnosis; Pyelonephritis, (*E. coli* isolated from urine), peri-anal/rectal abscess, colitis (*S. enteritidis* isolated from stool), silicone implant disruption, lymphoma).

CT scans of the head (CT head) were performed in 21% (21/100) of participants and were abnormal in 23.8% (5/21), of the abnormal scans 60% (3/5) did not contribute to the final diagnosis and the results of one scan 33.3% (1/3) prevented a planned diagnostic lumbar puncture as it revealed low lying cerebellar tonsils which is a contraindication to lumbar puncture- this participant remained undiagnosed at the end of the study period. The remaining 40% (2/5) of abnormal CT head scans had findings consistent with the final diagnosis but would not have been diagnostic in isolation. The first, revealed prominent temporal horns, a finding consistent with increased intra-cranial pressure and the second revealed a tortuous optic nerve sheath, again a finding, which can be associated with raised intra-cranial pressure.

Both these participants had Enterovirus meningitis confirmed by PCR testing of CSF. The CT neck performed in one participant 1% (1/100) revealed an enlarged cerebellar tonsil but no definitive diagnosis.

Transthoracic echocardiograms (TTE) were performed in 7% (7/100) patients of which two were abnormal; one of these revealed a pericardial effusion consistent with a relapsed pericarditis, however the causative aetiology remained unknown. The second revealed moderate to severe aortic stenosis but excluded signs of infective endocarditis so neither scan yielded a definitive diagnosis supported by microbiological evidence.

MRI brain was performed in 4% (4/100) and was normal in all 100% (4/4). One each of MRI liver 1% (1/100) and MRI spine 1% (1/100) were performed. The MRI liver contributed to the patient management as confirmed the presence of a hepatocellular carcinoma but not the cause of AUFI. The MRI spine was reported as having possible signs of discitis, but this was not confirmed on consultant radiologist review and therefore this scan was not diagnostic. Table 3.21 summarises the diagnoses made or supported by radiological investigations.

Table 3.21

Causes of AUFI identified or supported by radiological investigations	
CXR (*influenza A, Mycoplasma pneumoniae*)	2
CT CAP/AP (Pyelonephritis, peri-anal/rectal abscess, colitis, silicone implant disruption, lymphoma)	5
CT head (both, signs of raised intracranial pressure)	2
TTE (pericardial effusion and relapses pericarditis)	1

3.2.6.2 Diagnostic yield of standard of care investigations

A large number of microbiological and radiological investigations were performed across the entire AUFI cohort as part of routine clinical care. In total, 1,401 individual investigations were performed with 5% (70/1401) of these investigations contributing to a clinically credible diagnosis in participants with AUFI (see table 3.22). Of the different investigations groups, the 'CSF, tissue, EDTA whole blood and other' investigations had the highest contribution to a clinically credible diagnosis 17.2% (10/58). This targeted group of investigations were performed in a very small number of participants and it is likely that a high pre-test probability of a disease process (such as meningococcal sepsis) and careful case selection for these targeted diagnostic tests is responsible for the high number of clinically useful results. By comparison, 'bacterial' 6.2% (15/242) and 'virology, serological and tropical' 3.8% (35/918) had similarly low diagnostic value in contributing to a clinically credible diagnosis in participants with AUFI.

Table 3.22 Diagnostic yield of standard of care investigations in all participants with AUFI, n=100

Diagnostic yield of standard of care investigations	
Investigation type	**Contributed to Diagnosis**
Bacterial	15/242 (6.2)
Virology, serological and tropical	35/918 (3.8)
CSF, tissue, EDTA whole blood and other investigations	10/58 (17.2)
Radiological	10/183 (5.4)
All investigations	**70/1401 (5.0)**
All data presented as n/n (%)	

3.2.6.3 Diagnosis made by standard of care investigations

Of all participants with AUFI, 48.0% (48/100) had a clinically credible diagnosis made by standard of care investigations. The majority of these were attributable to infection (93.8%; 45/48) as opposed to non-infection causes (6.3%; 3/48). Viral infection (60.0%; 27/45) (56.3% 27/48 of whole cohort) was most common, followed by bacterial infection (40.0%;18/45), (37.5%; 18/48 whole cohort). There were no clinically significant fungal or parasitic infections identified in this study.

Of the clinically credible viral infections, influenza A was the most commonly detected pathogen (lineage H3 (3), H1N1 (5), non-typeable (1)), 33.3% (9/27), followed by Dengue virus (18.5%; 5/27) and Enterovirus infection (18.5%; 5/27) (Coxsackie B (2), non-typeable (2), Echovirus (1)), then EBV (11.1%; 3/27). Hepatitis E virus infection, Parainfluenza virus 3, Norovirus 3.7% Chikungunya virus and Influenza B virus were detected from 1 participant each (see figures 3.3 and 3.4 and table 3.23).

Of the bacterial infections 11.1% (2/18) had no pathogen identified but had either clear radiological evidence of infection with a compatible clinical presentation (peri-rectal/anal abscess identified on CT and MRI in 1 patient) or had a compatible clinical history and admission blood tests with positive urine microscopy for WCC and bacteria (i.e. urinary tract infection in 1 patient). All other participants categorised as bacterial infection had a causative organism identified, 88.9% (16/18). The most common organism identified was *E. coli* 22.2% (4/18) (bacteraemia (1), and urine (3)), followed by *Campylobacter* species 11.1% (2/18), *Mycoplasma pneumoniae* 11.1% (2/18), and *S. pyogenes* 11.1% (2/18). *Plesiomonas shigellosis* 5.6% (1/18), *Salmonella enteritidis* 5.6% (1/18), *Clostridiodes difficile* 5.6% (1/18), *Neisseria meningitidis* serogroup B 5.6% (1/18), *Shigella sonnei* 5.6% (1/18), *Pseudomonas aeruginosa* 5.6% (1/18) were diagnosed in one participant each. The remaining 6.3% (3/48) of participants diagnosed by standard of care investigations had non-infection diagnoses. These were Angioimmunoblastic T cell lymphoma Silicosis secondary to breast implant rupture and seronegative autoimmune hepatitis in 1 patient each.

Table 3.23. Summary of standard of care diagnoses in all participants with AUFI n=48

All confirmed Infections	**45/48 (93.8)**
Viral infection	**27/48 (57.3)**
Influenza A (Respiratory PCR)	9/27 (33.3)
Dengue virus (Serology and PCR)	5/27 (18.5)
Enterovirus (PCR CSF and Respiratory PCR)	5/27 (18.5)
Epstein Barr virus (serology +/- PCR)	3/27 (11.1)
Hepatitis E virus (serology and PCR)	1/27 (3.7)
Parainfluenza virus 3 (Respiratory PCR)	1/27 (3.7)
Norovirus (Stool)	1/27 (3.7)
Chikungunya virus (EDTA + serology)	1/27 (3.7)
Influenza B (Respiratory PCR)	1/27 (3.7)
Bacterial infection	**18/48 37.5)**
E. coli (n= 1 bacteraemia, n=3 urine culture)	4/18 (22.2)
Streptococcus pyogenes (ASO + clinical history)	2/18 (11.1)
Campylobacter species (stool culture)	2/18 (11.1)
Mycoplasma pneumonia (serology and PCR)	2/18 (11.1)
Plesiomonas shigelloides (stool culture)	1/18 (5.6)
Salmonella enteritidis (stool culture)	1/18 (5.6)
Clostridiodes difficile (stool culture)	1/18 (5.6)
Neisseria meningitidis serogroup B (blood PCR)	1/18 (5.6)
Shigella sonnei (stool culture)	1/18 (5.6)
Pseudomonas aeruginosa (blood cultures)	1/18 (5.6)
Rectal abscess (CT and MRI scan)	1/18 (5.6)

UTI (microscopy)	1/18 (5.6)
Non-infective diagnosis	**3/48 (6.2)**
Angioimmunoblastic T cell lymphoma (tissue biopsy)	1/3 (33.3)
Silicosis secondary to breast implant rupture (tissue	1/3 (33.3)
Seronegative autoimmune hepatitis (tissue biopsy)	1/3 (33.3)
All data presented as n/n (%)	

3.2.7 Antibiotic, Outcome and Safety Data

3.2.7.1.1 Antimicrobial treatment and Confirmed Diagnoses

Antimicrobial data was available for all participants, of these, 81% (81/100) were treated with an antimicrobial agent. The majority of participants received an antibacterial agent alone 71.6% (58/81), and 26.8% (22/81) received an antibacterial agent in conjunction with one other agent (antiviral (86.4%; 19/22), antimalarial (9.1%; 2/22), antifungal (4.5%; 1/22)) and one participant received antibacterial, antiviral and antimalarial agents. All participants receiving either an antiviral, antifungal or antiparasitic agent also received an antibacterial agent (see Table 3.24).

Table 3.24 Summary of antimicrobials received in participants with AUFI, n=100

Antimicrobial	100 (%)	Duration (days)
Any antimicrobial agent given	81 (81.0)	5.3 [2.0-9.5]
Antibacterial alone	58 (58.0)	6.0 [2.5.0-10.0]
Antiviral agents	2 (2.0)	2.6 [1.1-5.8]
Antifungal agents	1 (1.0)	12.6 (NA)
Antimalarial agents	3 (3.0)	2.0 [1.6-2.5]
All data presented as n/n (%) and median [IQR], NA =not applicable		

3.2.7.1.2 Antibacterial Consumption

Twenty-four different antibacterial agents were prescribed (see figure 3.5). Ciprofloxacin was prescribed for the highest total number of days (137 days) across all participants followed by co-amoxiclav (91.7 days) and doxycycline (90.1 days). These three antibacterials represented 48.3% (318.8/660) of all antibacterial days prescribed. A single dose of Ertapenem and a single dose of Clindamycin were prescribed across the entire cohort and are not represented in figure 3.4.

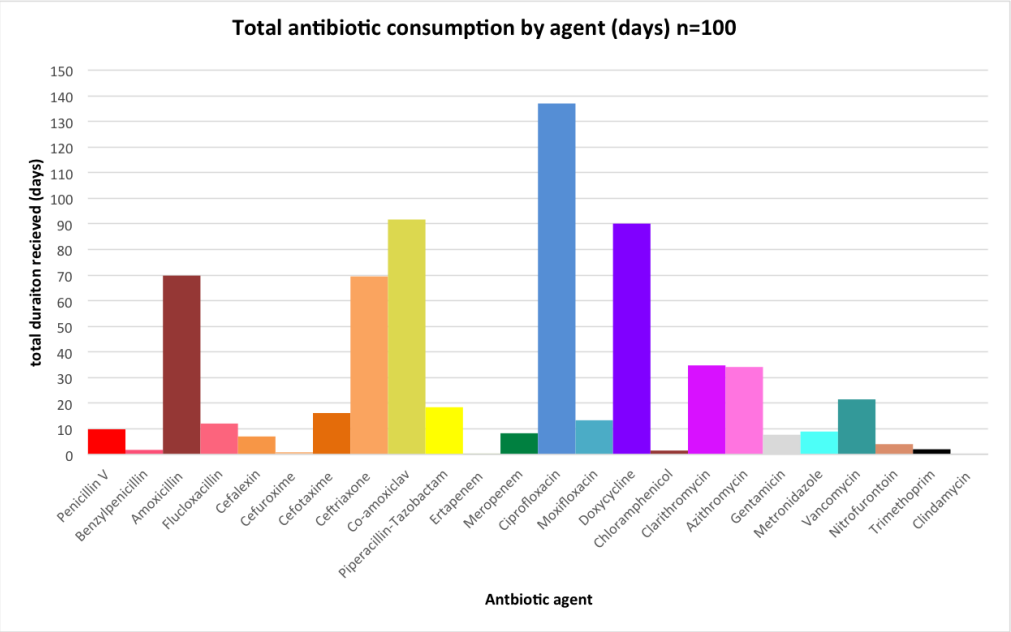


Figure 3.5 Total antibiotic consumption across entire AUFI cohort (n=100)

There was much variability in the number and type of antibacterial agents prescribed and the duration of treatment (see figure 3.5). Of those prescribed antibacterial agents, 37.0% (30/81) were prescribed a single agent, 24.7% (20/81) two agents and 12.3% (10/81) received three agents. Twenty-one participants (25.9%; 21/81) received four or more antibacterial agents with one individual (participant 70) receiving eight different agents although only three of these were prescribed for over 48 hours and one was a single dose of gentamicin.

Figure 3.6

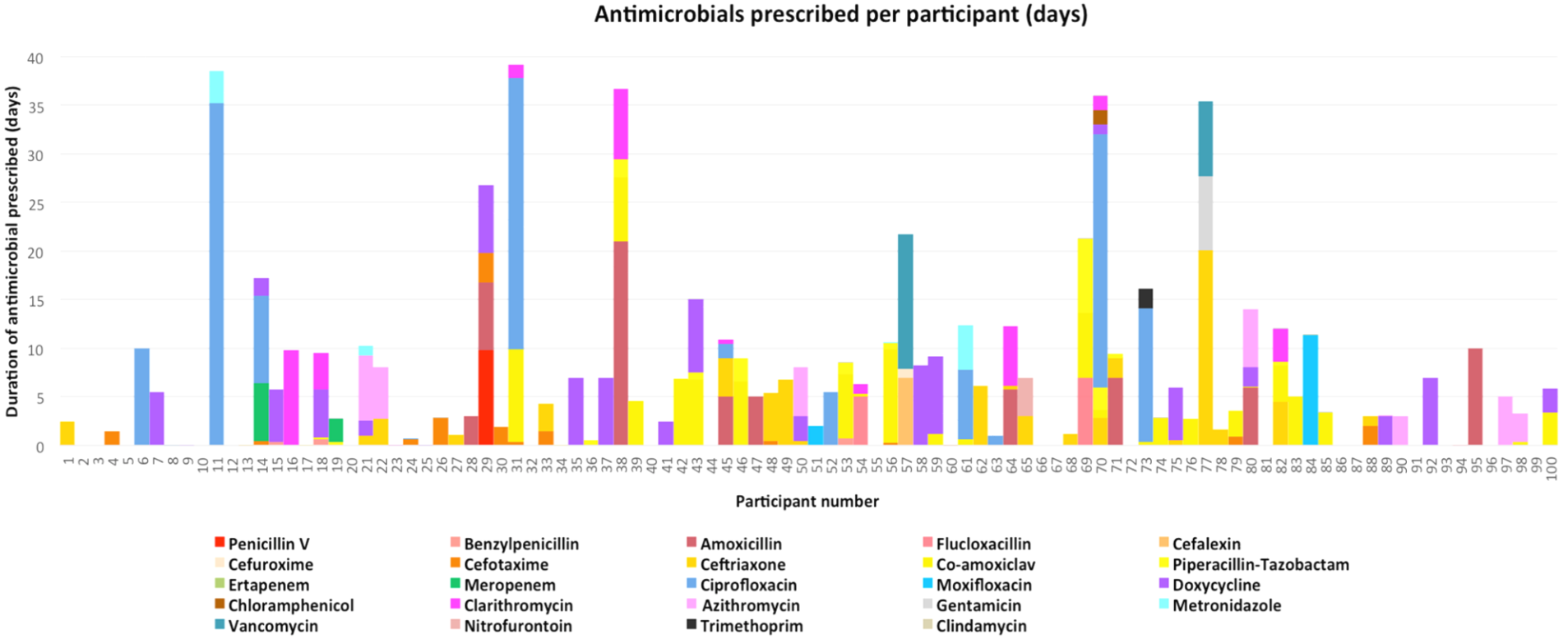


Figure 3.6 Duration of antimicrobials by type (days), per participant, illustrating the range of different agents used across the entire cohort of participants with acute undifferentiated febrile illness (n=100)

3.2.7.1.3 Antiviral and Antifungal consumption

Few participants received antivirals, antifungals or antimalarials and this data is presented in table 3.25. Antivirals were most commonly prescribed (23%; 23/100), only three participants received antimalarials and one participant received both antifungal treatments.

Table 3.25 Antiviral and antifungal treatment duration (days)

Antiviral	**20 (20)**	
Aciclovir	7/20 (35)	0.4 [0.02-0.7]
Oseltamivir	14/20 (70)	3.5 [1.4-6.2]
Valaciclovir	2/20 (10)	5.5 [4.5-6.5]
Antifungal	**1 (1)**	**Median (days) [IQR]**
Fluconazole	1/1 (100)	2.0 (NA)
Nystatin	1/1 (100)	10.6 (NA)
Antimalarial	**3 (3)**	**Median (days) [IQR]**
Artemether and Lumefantrine	3/3 (100)	2.0 [1.6-2.5]
All data presented as n/n (%) and median [IQR], NA, not applicable.		

3.2.7.2 Final diagnosis by standard of care investigation

Of the one hundred adult participants recruited, approximately half (52%; 52/100) remained undiagnosed at follow up. Of those who had a diagnosis made, the majority of diagnoses were infectious aetiologies, with a minority of non-infection diagnoses made (3%; 3/100) (see table 3.26)

Table 3.26 Summary of diagnoses made by standard of care investigations in participants with AUFI, n=100

Diagnosis made via standard of care	**n/n (%)**
Undiagnosed	52 (52)
Non-Infection diagnosis	3 (3)
Bacterial infection	18 (18)
Viral infection	27 (27)
All data presented as n/n (%)	

3.2.7.3 Outcome and Safety

Outcome and safety data was available for all one hundred participants (see table 3.27). The median length of stay for the entire cohort was 2 days (IQR 1.0-4.1, range 0.04-35.6). Three quarters of participants attended follow up 75.0% (75/100) and over half of those who attending had on-going symptoms at follow up 66.7% (50/75). No participant died during follow up and a small number were readmitted in with 30 days of recruitment (3.0%; 3/100).

Table 3.27 Outcome and safety data in all participants with AUFI, n=100

	Median [IQR]	**Range**
Length of Stay (days)	2.0 [1.0-4.1][1]	0.04-35.6
	n= (%)	
Admitted to ICU	1 (1.0)	
Died during follow up	0 (0.0)	
Readmission within 30 days	3 (3.0)	
Attended follow up	75 (75.0)	
Symptoms at Follow up	50/75 (66.7)	
Died during follow up	0.0 (0.0)	
[1] assessed in 97 participants		

Of those readmitted to hospital within 30 days where data was available, no participant was admitted for reasons related to participation in the study or infection related complications. One patient was admitted for further management of seronegative autoimmune hepatitis, which had been diagnosed on the same admission as they were recruited to the clinical study. A second patient was admitted with complications of a phaeochromocytoma and data is missing on the reason for the re-admission for the third participant.

3.3 Comparison of diagnosed and undiagnosed AUFI

3.3.1 Demographics

There were 48% (48/100) participants who achieved a diagnosis by standard of care investigations and 52% (52/100) who remained undiagnosed. There were no significant differences between the age, gender or history of travel in the prior three months in the diagnosed and undiagnosed groups. The median duration of fever at presentation was 4 days in the undiagnosed group and 3 days in the diagnosed group. Undiagnosed participants had a median of 1 [IQR 0.0-02.0] comorbidity and diagnosed participants a median of 2.0 [IQR 1.0-3.0](See table 3.28).

Table 3.28 Demographics in diagnosed versus undiagnosed participants with AUFI

Demographics	**Diagnosed**	**Undiagnosed**		
	n= 48	**n= 52**	**OR (95% CI)**	***p* value**
Female	21 (44.0)	23 (44.0)	0.98 (0.43-2.2)	1.00
White British	34 (71.0)	34 (65.0)	1.29 (0.56-3.1)	0.67
Travel in prior 3 months	24 (52.0)	26(52.0)	1.00 (0.45-2.24)	1.00
Age (years)	33.0 [24.0-50.0]	36.0 [25.0-48.0]	3.00 (-5.00-8.00)	0.63
No. of comorbidities	2 [1.0-3.0]	1 [0.0-2.0]	-1.00 (-1.00-0.00)	0.08
Duration of fever (days)	3 [1.0-6.0]	4 [3.0-6.0]	1.00 (0.00-2.00)	0.07
All data presented as n/n (%) and median [IQR],				

3.3.2 Presenting clinical features

The level of fever was similar between both undiagnosed 38.0°C (IQR: 37.0-38.7°C) and diagnosed participants 38.3 °C (IQR: 37.8-38.9 °C). All participants had a fever reported or measured at presentation, as fever was part of the inclusion criteria. There were no significant differences in the presenting clinical features of undiagnosed and diagnosed participants except for weight loss, which was more commonly reported, in the undiagnosed group (19.2%; 10/52, versus 2.1%; 1/48, OR 0.09, 95%CI 1.7-123.9; *p*=0.01). See table 3.29 for full details.

Table 3.29 Presenting clinical features of diagnosed versus undiagnosed participants with AUFI

Symptom	Diagnosed, n=48	Undiagnosed, n=52		*p* value
Temperature (°C)	38.3 [37.8-38.9]	38.0 [37.0-38.7]		0.173
Constitutional			**OR (95% CI)**	
Fever	48.0 (100.0)	52.0 (100.0)	-	1.00
Lethargy	42.0 (87.5)	47.0 (90.4)	0.74 (0.2-2.8)	0.75
Night sweats	38.0 (79.2)	42.0 (80.8)	0.90 (0.4-2.3)	1.00
Loss of appetite	32.0 (66.7)	32.0 (61.5)	1.25 (0.6-2.8)	0.68
Lymphadenopathy	17.0 (35.4)	18.0 (34.6)	1.00 (0.5-2.3)	1.00
Weight loss	1.0 (2.1)	10.0 (19.2)	0.09 (1.7-123.9)	**0.01**
Respiratory				
Cough	19.0 (39.6)	23.0 (44.2)	0.83 (0.4-1.9)	0.69
Sore throat	17.0 (35.4)	20.0 (38.5)	0.88 (0.4-1.9)	0.84
Coryza	11.0 (22.9)	14.0 (26.9)	0.81 (0.3-1.9)	0.82
Shortness of breath	11.0 (22.9)	9.0 (17.3)	1.42 (0.5-3.6)	0.62
Pleuritic chest pain	4.0 (8.3)	5.0 (9.6)	0.85 (0.2-3.1)	1.00
Wheeze	3.0 (6.3)	4.0 (7.7)	0.80 (0.2-3.1)	1.00
Gastrointestinal				
Diarrhoea	13.0 (27.1)	12.0 (23.1)	1.24 (0.5-3.0)	0.65
Vomiting	17.0 (35.4)	21.0 (40.4)	0.81 (0.4-1.8)	0.68
Abdominal pain	18.0 (37.5)	17.0 (32.7)	1.23 (0.6-2.8)	0.68
Jaundice	3.0 (6.3)	0.0 (0.0)	-	0.11
Hepatomegaly	2.0 (4.2)	1.0 (1.9)	2.2 (0.2-32.6)	0.61
Splenomegaly	3.0 (6.3)	0.0 (0.0)	∞ (0.96- ∞)	0.11
Cardiovascular				
Palpitations	3.0 (6.3)	3.0 (5.8)	1.10 (0.2-4.8)	1.00
Chest pain	1.0 (2.1)	3.0 (5.8)	0.35 (0.02-2.4)	0.62
Neurology				
Headache	38.0 (79.2)	43.0 (82.7)	0.80 (0.3-2.0)	0.80
Photophobia	14.0 (29.2)	17.0 (32.7)	0.85 (0.4-2.0)	0.83
Neck stiffness	7.0 (14.6)	14.0 (26.9)	0.46 (0.2-1.2)	0.15
Conjunctivitis	9.0 (18.8)	11.0 (21.2)	0.86 (0.3-2.4)	0.81
Focal neurology	2.0 (4.2)	2.0 (3.8)	1.08 (0.2-7.2)	1.00
Musculoskeletal				
Arthralgia	18.0 (37.5)	15.0 (28.8)	1.48 (0.7-3.5)	0.40
Myalgia	18.0 (37.5)	22.0 (42.3)	0.81 (0.4-1.9)	0.69
Genitourinary				
Dysuria	3.0 (6.3)	4.0 (7.7)	0.80 (0.2-3.1)	1.00
Frequency	1.0 (2.1)	0.0 (0.0)	-	1.00
Other	3.0 (6.3)	2.0 (3.8)	1.67 (0.33-9.67)	0.67

Rashes	Diagnosed, n=48	Undiagnosed, n=52		*P value*
Vesicular	0.0 (0.0)	5.0 (9.6)	0.00 (0.00-0.71)	0.06
Macular	14.0 (29.2)	10.0 (19.2)	1.73 (0.7-4.5)	0.35
Petechial	1.0 (2.1)	1.0 (1.9)	1.09 (0.06-21.0)	1.00
Other	2.0 (4.2)	1.0 (1.9)	2.22 (0.2-32.7)	0.61
All data presented as n/n (%) and median [IQR]				

3.3.3 Antibiotics, outcome and safety data

3.3.3.1 Antimicrobial therapy in diagnosed versus undiagnosed participants

Antimicrobials were prescribed in 84.6% (44/52) of the undiagnosed group and 77.1% (37/48) of the diagnosed group (OR 0.61; 95% CI 0.21-1.77, *p*= 0.45). The duration of antimicrobials 5.6 days in the undiagnosed groups [IQR 1.6-9.9] and 5.0 days in diagnosed group [IQR 0.1-12.1] (difference 0.63 (95% CI -2.46-1.96) *p*=0.99) see table 3.30 for full details.

Table 3.30 Antibiotic use in diagnosed versus undiagnosed participants with AUFI, n=100

	Diagnosed	**Undiagnosed**	**OR or difference (95% CI)**	***p* value**
	n= 48	**n= 52**		
Received antimicrobials	37 (77.1)	44 (84.6)	0.61 (0.21-1.77)	0.45
Duration of antibiotics (days)	5.0 (0.1-12.1)	5.6 [1.6-9.9]	0.63 (-2.46-1.96)	0.99
Duration of fever (days)	3 [1.0-6.0]	4 [3.0-6.0]	1.00 (0.00-2.00)	0.07
All data presented as n/n (%) and median [IQR],				

3.3.3.2 Outcome and Safety

Median length of hospital stay was significantly longer in undiagnosed participants (2.9 days [IQR 1.6-4.9]) versus diagnosed participants (1.7 days [IQR 0.8-1.6]) (*p*=0.036). There was a trend for undiagnosed participants 59.6% (31/52) to have more symptoms at follow up than diagnosed participants 39.6% (19/48) however, this was not statistically significant (OR 0.44, 95% CI 0.2-4.9; *p*=0.07). Interestingly, all adverse events occurred in the diagnosed group (n=1 GICU admission, n=3 readmission within 30 days) (see table 3.31) and the odds ratio of experiencing an adverse event was higher in the group of diagnosed patients (8.3%; 4/48 versus 0%; 0/52, *p*=0.05).

Table 3.31 Follow up and outcome data for diagnosed (n=48) and undiagnosed (n=52) participants with AUFI.

	Diagnosed	**Undiagnosed**	**Difference (95% CI)**	***p* value**
	n=48	n=52		
Length of stay (days)	1.7 [0.8-1.6]	2.9 [1.6-4.9]	1.2 (0.1-1.7)	**0.036**
			OR (95% CI)	
Admitted to ICU	1 (2)	0 (0)	∞ (0.12-∞)	0.48
Died during follow up	0 (0)	0 (0)	-	-
Readmitted within 30 days	3 (6)	0 (0)	∞ (0.96-∞)	0.11
Attended follow up	32 (67)	41 (79)	0.53 (0.22-1.34)	0.18
Symptoms at follow up	19 (40)	31 (60)	0.44 (0.2-4.9)	0.07
Any adverse event	4 (8.3)	0 (0)	∞ (1.09-∞)	**0.05**
All data presented as n/n (%) and median [IQR]				

3.4 Summary

This proof of concept study is the first identified to examine adults presenting with AUFI to a UK hospital. Although numbers are small, there are some important key findings. Firstly, approximately half (52%; 52/100) of participants remained undiagnosed following thorough standard of care investigations and follow up. Secondly, the majority of participants (93.8%; 45/48) in which a diagnosis was made had an infectious cause identified by standard diagnostics. Thirdly, a broad range of pathogens was identified. Fourthly, antimicrobial use was common in both diagnosed and undiagnosed participants. Fever resolved in all participants with the exception of one who was diagnosed with angioimmunoblastic T cell lymphoma. The most significant finding however, was an increased length of stay in undiagnosed participants, who stayed an average of 1.2 days longer in hospital, suggesting that diagnostic uncertainty may contribute to longer hospital admissions. All assumptions made from this data set should be caveated with the single site and small sample size, and this will be further discussed in Chapter 6 (Discussion).

Chapter 4 DNA extraction of EDTA whole blood: results

Metagenomic NGS depends on efficient, unbiased extraction of nucleic acids from a biological sample. In studies of infection, depletion of host nucleic acids is of equal importance to increase the proportion of pathogen derived nucleic acids compared with the much more abundant host material, thereby optimising the detection of organisms that may be present in low numbers. There is currently no standardised approach to these essential pre-analytical steps and a wide range of commercial and novel approaches have been reported in the literature.

This chapter describes a series of experiments performed to compare two manual DNA extraction and enrichment techniques using EDTA whole blood from healthy volunteers spiked with known concentrations of *Escherichia coli* ATCC 2955 (*E. coli*). Methods were chosen for their ability to deplete human DNA whilst preserving or 'enriching' for prokaryotic DNA.

The aim was to identify which technique was superior for depleting human genomic material (hereafter referred to as gDNA) and enriching for prokaryotic genomic material, in this case *E. coli*. The enrichment processes used were non-selective for *E. coli* species, but as *E. coli* is a common cause of human blood stream infection it is a useful surrogate for these experiments [203]. *Candida albicans* ATCC 90028 (*C. albicans*), also a common cause of human infection, was selected to ensure the methodology was also robust against eukaryotic pathogens such as fungi. The first, section 4.1, will present the results of *E. coli* ATCC 2955 and *C. albicans* ATCC 90028 growth curve experiments to calculate mid-log growth phase and generation of stock culture dilution series used for the DNA extraction experiments. The second section (4.2) will present the findings of the extraction and enrichment experiments.

4.1 Generation of a *E. coli* ATCC 2955 and *C. albicans* ATCC 90028 dilution series stock culture

4.1.1 Generating *E. coli* and *C. albicans* growth curves

To accurately quantify extraction efficiency, it is necessary to be able to reproducibly spike samples with known concentrations of bacteria sampled at mid-log growth phase where organisms are dividing rapidly. To facilitate this, growth curves to determine Log growth phase of *E. coli* and *C. albicans* were generated. *E. coli* and *C. albicans* growth cultures (GC) were set up as described in section 2.2.1.2 to 2.2.1.4. Regular measurements of optical density (OD) were performed at time intervals post-inoculation. Optical density values obtained during the generation of the *E. coli* (see figure 4.1) and *C. albicans* growth curves to determine the mid-log growth phase are shown below.

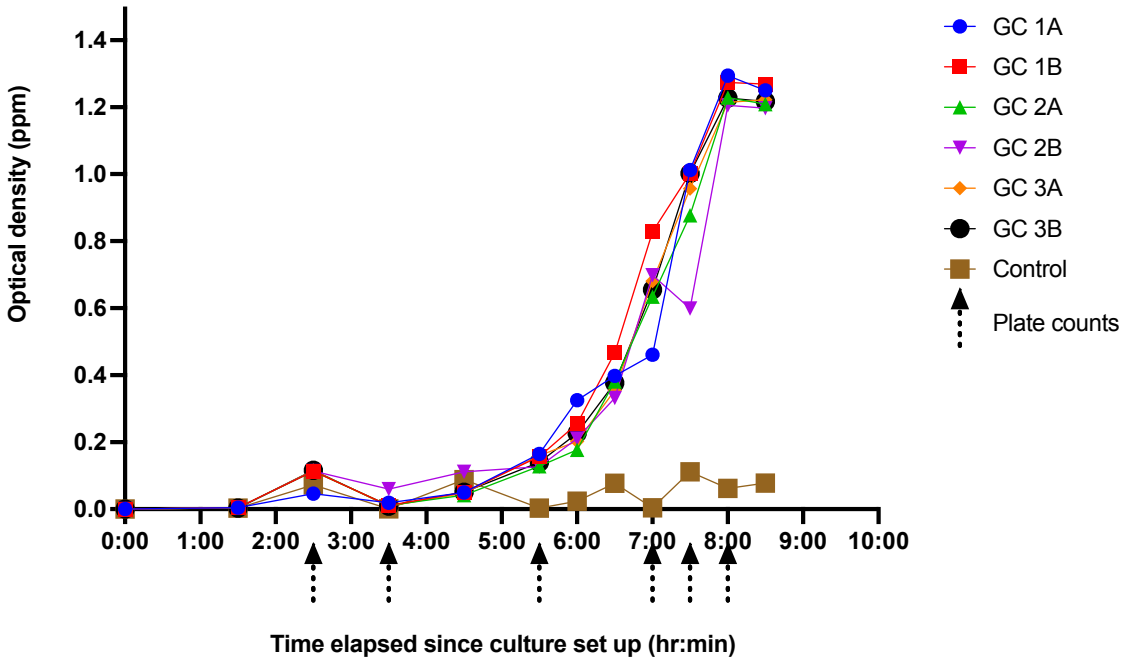


Figure 4.1 *E. coli* growth curve used to calculate mid-logarithmic growth phase. GC 1 to GC 3 each represents a duplicate growth culture (GC) inoculated from the starter culture at time zero. The control used was uninoculated BHI broth.

Plate counts for *E. coli* were performed at six time points (figure 1.2, dotted black arrow). *C. albicans* plate counts were performed when OD reached 0.5 ppm (see figure 1.2, orange arrow). The growth curves generated were then used to determine mid-log growth phase.

4.1.2 Calculating mid-logarithmic growth phase

Mid-log phase was calculated as below after plotting OD to a logarithmic scale. For *E. coli* mid-log growth occurred at 6.5 hours after the culture was inoculated, and represented an optical density of 0.460 ppm.

Equation 4.1 Calculation of mid-logarithmic growth phase *E. coli*

Mid-logarithmic growth point time = ((07:30– 05:30)/2) + 05:30

= 06:30 (hrs:mins from culture set up)

Mid-logarithmic growth point OD = ((0.794 – 0.126)/2) + 0.126

= 0.460 ppm

For *C. albicans,* mid-logarithmic growth occurred at eight hours after the culture was set up at an optical density of 0.564 ppm (see equation 4.2).

Equation 4.2 Calculation of mid-logarithmic growth phase *C. albicans*

Mid logarithmic growth point time = ((09:00– 07:00)/2) + 07:00

= 08:00 (hrs:mins from culture set up)

Mid logarithmic growth point OD = ((0.818– 0.311)/2) + 0.311

= 0.564 ppm

4.1.3 Quantification of *E. coli* and *C. albicans* using plate counts

The calculated timing of mid-log phase (section 1.1.2) was used to inform the timing of culture sampling to generate a stock culture dilution series with plate count quantifications for *E. coli* (table 4.1) and *C. albicans* (Table 4.2).

Table 4.1 reports *E. coli* plate counts and table 4.2 the results of *C. albicans* plate counts taken at mid log growth phase (sample 'neat') and subject to a serial dilution in PBS (samples B to J). As expected, plate counts and calculated CFU/mL decreased with each dilution step (samples A to J) for both organisms. Overall, *C. albicans* plate counts were lower, this is likely due to the slower rate of replication of *C. albicans* (optimal growth 1 hour) versus *E. coli* (optimal growth 20mins).

Table 4.1 Plate counts of *E. coli* dilution series

		GC 1 (0.608 ppm)		GC 2 (0.528 ppm)		GC 3 (0.600 ppm)	
Sample [3]	**Dilution**	**Plate count[1]**	**CFU/mL**	**Plate count**	**CFU /mL**	**Plate count**	**CFU/mL**
A	Neat	-	-	-	-	-	-
B	1:10	TMTC[2]	$>10^4$	TMTC	$>10^4$	TMTC	$>10^4$
C	$1:10^2$	TMTC	$>10^4$	TMTC	$>10^4$	TMTC	$>10^4$
D	$1:10^3$	TMTC	$>10^4$	TMTC	$>10^4$	TMTC	$>10^4$
E	1: 10^4	>1000	$>10^4$	>1000	$>10^4$	>1000	$>10^4$
F	$1:10^5$	101	1.0×10^3	66	6.6×10^2	265	2.7×10^3
G	$1:10^6$	19	1.9×10^2	9	90	28	2.8×10^2
H	$1:10^7$	1	10	0	0	2	20
J	$1:10^8$	0	0	0	0	0	0

[1]All plate counts performed using 100 µL culture broth, [2]TMTC= Too many to count, [3]control sample plate counts were negative at 'neat' with no evidence of contamination (not shown).

Table 4.2 Plate counts of *C. albicans* dilution series

		GC 1 (0.533ppm)		GC 2 (0.585ppm)		GC 3 (0.561ppm)	
Sample[b]	**Dilution**	**Plate count[a]**	**CFU/ml**	**Plate count**	**CFU /mL**	**Plate count**	**CFU/ml**
A	Neat	-	-	-	-	-	-
B	1:10	TMTC	$>10^4$	TMTC	$>10^4$	TMTC	$>10^4$
C	$1:10^2$	>1000	$>10^4$	>1000	$>10^4$	>1000	$>10^4$
D	$1:10^3$	586	5.9×10^3	818	8.2×10^3	651	6.5×10^3
E	1: 10^4	43	4.3×10^2	88	8.8×10^2	86	8.6×10^2
F	$1:10^5$	6	60	2	20	9	90
G	$1:10^6$	0	0	3	30	1	10
H	$1:10^7$	1	10	0	0	0	0
J	$1:10^8$	0	0	0	0	0	0

[a]All plate counts performed on 100 µL culture broth

[b]control sample plate counts were negative at 'neat' with no evidence of bacterial contamination (not shown), [c]TMTC= too many to count

Less than 10 colonies per mL was obtained by dilution $1:10^6$ (row G and H, table 4.1) in the *E. coli* dilution series and by $1:10^5$ (row F and G, table 4.2) in the *C. albicans* dilution series demonstrating an appropriate dilution series. Stock cultures were stored at −20°C.

4.2 Comparison of two manual DNA extraction and enrichment techniques applied to EDTA whole blood samples spiked with known quantities of *E. coli: Results*

The following section reports the findings of a comparison between two manual bacterial DNA extraction and enrichment techniques; the Molysis Complete 5 Extraction Kit (Molzym, Bremen, Germany) and the Blood and Cell culture DNA Maxi Kit (Qiagen, Hilden, Germany) followed by the NebNext Microbiome DNA extraction kit (NEB, Ipswich, MA). EDTA whole blood inoculated with known quantities of *E. coli* was subjected to DNA extraction, enrichment of prokaryotic DNA by depletion of gDNA. Semi-quantitative real-time PCR was performed using an *E. coli* specific assay (uidA, Glucuronidase) and a human specific gDNA target (assay detects single a copy of non-transcribed region which is highly conserved in the human genome) (see section 2.2.2.2). The two extraction methods were compared to determine their efficiency at preserving bacterial (*E. coli*) DNA and depleting host genomic material (gDNA). Here, PCR cycle thresholds (*Ct*) were used as a proxy for target concentration with depletion of unwanted gDNA indicated by an increased *Ct* value.

Following the *E. coli* extraction comparison experiments the aim was to repeat this work with *C. albicans* to evaluate the robustness of the extraction methodology against a non-bacterial pathogen. Due to a collaboration opportunity with a site with expertise in metagenomic sequencing of clinical samples the extraction experiments with *C. albicans* were not performed and so this section refers to *E. coli* extraction only.

4.2.1 Comparison of extraction and enrichment methodologies of bacterial DNA

The enrichment of prokaryotic genomic material in a biological sample relies on depletion of the more abundant host genomic DNA, and capture of prokaryotic DNA to enhance the proportion of prokaryotic nucleic acids present in a given volume of sample (see section 2.2 for full details on these methods). In this section, the two-step approach of the Qiagen kit is referred as 'pre-enrichment' (before magnetic bead depletion using NebNext) and 'post-enrichment' (after NebNext).

4.2.2 Enrichment of prokaryotic DNA

Table 4.3 shows the comparison between the Molysis and Qiagen extraction methods for enriching *E. coli* DNA. As expected, the mean *Ct* value for both methods increased as the sample concentration of *E. coli* decreases (10^4 to 10^1 CFU/mL). The Molysis kit performed slightly better

across all concentrations with *E. coli* detected in both replicates down to 1×10^1 CFU/ mL. The lower *Ct* values for Qiagen-based extractions without enrichment suggest a better efficiency of *E. coli* DNA retention at higher concentrations, but as indicated there was no consistent detection at the lowest concentration (10^1 CFU/mL). No significant difference in mean *Ct* was seen between the Molysis, Qiagen pre-enrichment and Qiagen post enrichment steps (see ANOVA, table 4.4).

Table 4.3 Comparison of semi-quantitative PCR for *E. coli* ATCC 2955 following Molysis and Qiagen enrichment extraction methods

	Molysis (*Ct*)			Qiagen pre-enrichment (*Ct*)			Qiagen post-enrichment (*Ct*)		
***E. coli* CFU/mL[a]**	**Rep. 1**	**Rep. 2**	**Mean (SD)**	**Rep. 1**	**Rep. 2**	**Mean (SD)**	**Rep.1**	**Rep. 2**	**Mean (SD)**
10^4	29.54	23.96	26.75 (3.95)	24.21	26.55	25.38 (1.66)	26.55	24.03	25.29 (1.78)
10^3	33.26	32.97	33.12 (0.21)	29.45	30.95	30.20 (1.06)	30.95	27.61	29.28 (2.36)
10^2	46.21	36.12	41.17 (7.14)	29.70	31.33	30.52 (1.15)	37.21	31.33	34.27 (4.16)
10^1	49.51	39.25	44.38 (7.26)	49.23	-	-	-	48.56	-
[a] Concentration refers to *E. coli* concentration inoculated into sample (1×10^X CFU/mL)									

Table 4.4 *E. coli* analysis of variance at each concentration for each method

***E. coli* CFU/ mL**	**Comparison**	**Mean diff.**	**95% CI intervals**	**Summary**	***p* Value**
10^4	Molysis Vs. pre Qiagen	1.37	-12.89 to 15.63	ns	>0.999
10^4	Molysis Vs. post Qiagen	1.46	-12.80 to 15.72	ns	>0.999
10^4	Pre Qiagen Vs. post Qiagen	0.09	-14.17 to 14.35	ns	>0.999
10^3	Molysis Vs. pre Qiagen	2.915	-11.35 to 17.18	ns	0.999
10^3	Molysis Vs. post Qiagen	3.835	-10.43 to 18.10	ns	0.994
10^3	Pre Qiagen Vs. post Qiagen	0.92	-13.34 to 15.18	ns	>0.999
10^2	Molysis Vs. pre Qiagen	10.65	-3.610 to 24.91	ns	0.221
10^2	Molysis Vs. post Qiagen	6.895	-7.365 to 21.16	ns	0.739
10^2	Pre Qiagen Vs. post Qiagen	-3.755	-18.02 to 10.51	ns	0.995
10^1	Molysis Vs. pre Qiagen	-4.85	-22.32 to 12.62	ns	0.992
10^1	Molysis Vs. post Qiagen	-4.18	-21.65 to 13.29	ns	0.998
10^1	Pre Qiagen Vs. post Qiagen	0.67	-19.50 to 20.84	ns	>0.999
[a] Concentration refers to *E. coli* concentration inoculated into sample (1×10^X CFU/mL)					

Furthermore, an analysis of variance was performed comparing the total mean *Ct* values for Molysis, and the pre and post-enrichment Qiagen results and there was no significant differences detected (see table 4.5 and figure 4.4) demonstrating that in this set of experiments, the two methods are equally as efficient at enriching *E. coli* DNA extracted from EDTA whole blood inoculated with four different known quantities of viable *E. coli*.

Table 4.5 Analysis of variance *E. coli* enrichment

***E. coli* Analysis of Variance [204], mean difference, *p* value, (95% confidence interval)**		
	Molysis	**Qiagen pre-enrichment**
Vs. Qiagen pre-enrichment	2.52, *p*=0.93, (-16.47 to 21.51)	-
Vs. Qiagen post-enrichment	2.01, *p*=0.95, (-16.99 to 21.00)	-0.52, *p*=1.00, (-19.51 to 18.47)

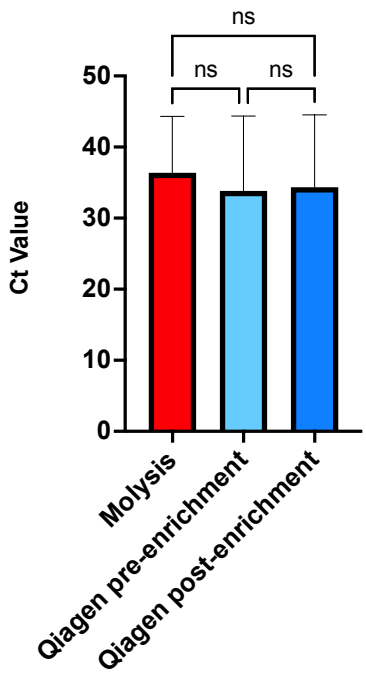


Figure 4.2 ANOVA *E. coli* enrichment

4.2.3 gDNA depletion

Table 4.6 shows the results of gDNA depletion for both methods. Whole blood in EDTA from the same volunteer was used for these experiments and therefore gDNA quantity (and therefore *Ct* value) would be expected to be consistent across the replicates at each concentration.

There was a significant increase in mean *Ct* value for gDNA (consistent with a fall in corresponding gDNA copies/mL) between pre-enrichment (mean *Ct* 16.92, range 15.30-18.37) and post enrichment steps (mean *Ct* 25.22, range 23.02-27.60) (difference *Ct* 8.30, 95% confidence interval

6.35-9.91 p=0.02) demonstrating a significant depletion of gDNA between these two steps (analysis not shown).

Table 4.6 Comparison of semi-quantitative PCR for gDNA following Molysis and Qiagen extraction methods

gDNA[a]	Molysis extraction (*Ct*)			Qiagen pre-enrichment (*Ct*)			Qiagen post-enrichment (*Ct*)		
	Rep. 1	Rep. 2	Mean (SD)	Rep. 1	Rep. 2	Mean (SD)	Rep. 1	Rep. 2	Mean (SD)
10^4	31.63	33.60	32.62 (1.39)	16.08	18.01	17.05 (1.37)	23.02	27.60	25.31 (3.24)
10^3	35.86	33.49	34.68 (1.68)	18.37	17.11	17.74 (0.89)	25.86	25.43	25.65 (0.30)
10^2	31.62	26.16	28.89 (3.86)	15.90	17.67	16.79 (1.25)	23.20	24.98	24.09 (1.26)
10^1	27.16	27.02	27.09 (0.10)	16.17	15.30	15.74 (0.62)	27.12	23.11	25.12 (2.84)

[a] Concentration refers to *E. coli* concentration inoculated into sample (1×10^X CFU/mL)

Table 4.7 summarises the analysis of variance comparing each methodology at each concentration, demonstrating that, with the exception of two results (Molysis Vs. post Qiagen at concentration 1×10^2 CFU/mL and Molysis Vs. post Qiagen at concentration 1×10^1 CFU/mL) *Ct* values at each comparison were statistically significantly different. At all concentrations of *E. coli*, all Molysis versus pre-Qiagen and all pre Qiagen versus post Qiagen results were statistically significant with the highest *Ct* values consistently found in the Molysis results.

Table 4.7 Analysis of Variance at each concentration gDNA for each method

gDNA	Comparison	Mean diff	95% CI	Significant?	*p* Value
10^4	Molysis Vs. pre Qiagen	15.57	8.804 to 22.34	Yes	**<0.0001**
10^4	Molysis Vs. post Qiagen	7.305	0.5391 to 14.07	Yes	**0.030**
10^4	Pre Qiagen Vs. post Qiagen	-8.265	-15.03 to -1.499	Yes	**0.013**
10^3	Molysis Vs. pre Qiagen	16.94	10.17 to 23.70	Yes	**<0.0001**
10^3	Molysis Vs. post Qiagen	9.03	2.264 to 15.80	Yes	**0.006**
10^3	Pre Qiagen Vs. post Qiagen	-7.905	-14.67 to -1.139	Yes	**0.017**
10^2	Molysis Vs. pre Qiagen	12.11	5.339 to 18.87	Yes	**0.001**
10^2	Molysis Vs. post Qiagen	4.8	-1.966 to 11.57	No	0.291
10^2	Pre Qiagen Vs. post Qiagen	-7.305	-14.07 to -0.5391	Yes	**0.030**
10^1	Molysis Vs. pre Qiagen	11.36	4.589 to 18.12	Yes	**0.001**
10^1	Molysis Vs. post Qiagen	1.975	-4.791 to 8.741	No	0.991
10^1	Pre Qiagen Vs. post Qiagen	-9.38	-16.15 to -2.614	Yes	**0.005**

[a] Concentration refers to *E. coli* concentration inoculated into sample 1×10^X CFU/mL

An ANOVA was performed comparing all the Molysis results with the Qiagen pre- and post-enrichment methods, which found statistically significant differences, with Molysis performing better i.e. removing more human gDNA across the *E. coli* concentrations. It is also of note that the

Qiagen post-enrichment process significantly improved human gDNA removal in comparison to the Qiagen method alone. (see table 4.8 and figure 4.4).

Table 4.8 Analysis of variance gDNA depletion

gDNA Analysis of Variance [204], mean difference, p value, (95% confidence interval)		
	Molysis	**Qiagen pre-enrichment**
Vs. Qiagen pre-enrichment	12.99, *p*= 0.0001, (8.53 to 17.45)	-
Vs. Qiagen post-enrichment	4.78, *p*= 0.04, (0.32 to 9.42)	-8.21, *p*=0.0016, (-12.67 to -3.75)

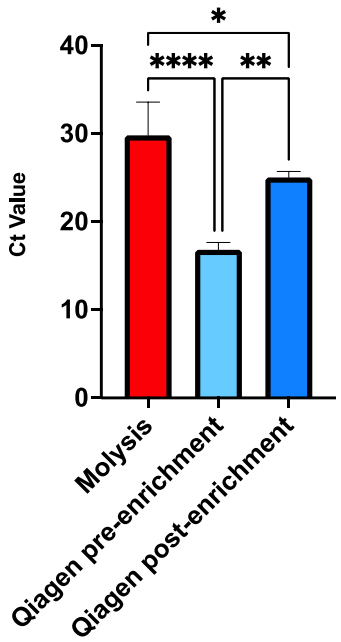


Figure 4.3 ANOVA gDNA depletion

4.3 Summary

This piece of work compared two commercially available extraction and enrichment techniques using EDTA whole blood spiked with known quantities of live *E.coli.* This comparison had not previously been described in the published literature and only a small number of studies reported extraction methods suitable for mNGS on EDTA whole blood [208-211]. These experiments found no statistically significant differences in detection of *E. coli* using semi-quantitative PCR, demonstrating that the Molysis Complete 5 method and the Qiagen Blood and Cell Maxi Kit plus NebNext Microbiome DNA extraction kit were equally as efficient at enriching pathogen DNA by depleting host gDNA.

Both the Molysis and Qiagen with enrichment methods resulted in significantly reduced quantities of human gDNA (demonstrated by higher *Ct* values). Of note, the differences between human gDNA *Ct* values following the Molysis method and the Qiagen post enrichment methods was statistically significant with the Molysis performing better (demonstrated by higher *Ct* values).

The new data here is largely consistent with the published literature. The Molysis basic 5 kit (host DNA depletion only) and the Molysis Complete 5 kit (DNA depletion, microbial enrichment and isolation) and Nebnext Microbiome DNA enrichment kit have been used on a range of biological samples for the purposes optimising molecular analysis for pathogen detection. These include, urine [150], joint sonicate fluid [208], CSF and NPA [209] and whole blood [210-211]. A range of organisms have been studied including; *E. coli* [210, 212], *K. pneumonia* [212] *S. aureus* [209, 213], *S. pneumoniae* [209], *S. agalactiae* [209]*, H. influenza* [209], *N. meningitidis* [209], *B. pertussis* [209], HSV [209], Adenovirus [209], Influenza A [209], *P. aeruginosa* [211], *C. albicans* [211], *S. epidermidis* [208], *E. faecalis* [208]. Schimidt *et al* found the Molysis basic kit to be superior to the NebNext Microbiome DNA enrichment kit when combined with centrifugation, bacterial lysis buffer and proteinase K (Roche, Basel Switzerland) for human DNA depletion of infected 4-10 mL urine samples (CFU/mL= 10^7-10^8) [150]. However, this work was performed on a small number of specimens (total n=15, clinical samples= 10/15, spiked samples 5/15), all organisms examined were Gram negative (*E. coli*, *E. cloacae* and *K. pneumonia*) and urine has fewer host cells than an equal volume of whole blood which may limit the relevance to work based on more cellular sample types.

Thoendel *et al* found the Molysis basic 5 kit superior to the NebNext microbiome DNA enrichment kit paired with the Mobio BiOstic Bacteremia DNA Isolation Kit (MO BIO Laboratories, Carlsbad, CA at gDNA) for depletion of joint sonicate fluid reporting a 76-fold enrichment of bacterial DNA in the sample by depletion of host DNA with the Molysis method versus a 6 fold enrichment with the NebNext method [208]. In contrast to the work presented here, this examined only Gram

positive organisms, and joint sonicate fluid is less cellular than whole blood and so is likely to contain less host gDNA than whole blood. More relevant perhaps, McCann *et al*, found the Molysis complete 5 method was superior to a bead-based extraction method in a Beagle *S. aureus* bacteraemia model. Beagles were intentionally infected with S. aureus to simulate bacterial sepsis and whole blood samples obtained. Samples processed with the Molysis method detected 70.7% (58/82) of *S. aureus* DNA from whole blood samples versus 59.8% (49/82) processed by an organic bead-based extraction method [210]. Similarly, Loonen *et al* found the Molysis Complete 5 kit was reliable and sensitive when used on 5 mL whole blood samples spiked with *S. aureus*, *P. aeruginosa* and *C. albicans* with culture concentrations ranging from 1-1000 CFU/ mL [211].

The Qiagen blood and cell maxi kit has a long-history of use across multiple biological specimens, to the best of our knowledge there has been no prior study in which It was combined with the NebNext Microbiome DNA enrichment for the purposes of isolating and enriching microbial DNA from clinical samples. Furthermore, no studies were identified directly comparing the Molysis Complete 5 method with the Qiagen blood and cell maxi kit plus NebNext Microbiome DNA enrichment.

From a practical perspective, the Molysis method was cheaper than the Blood and Cell culture DNA Maxi Kit (Qiagen, Hilden, Germany) plus NebNext Microbiome DNA enrichment kit (£12.96/sample versus £124.28/sample) due to the high costs of using magnetic beads for host gDNA depletion of a large volume sample. The Molysis method was also less time consuming and less technically difficult as extraction and enrichment could be performed in one continuous workflow. Conversely, a benefit of the Qiagen method was that it could be performed in a two-step approach allowing for samples to be frozen prior to gDNA depletion, which allowed more flexibility. Neither workflow is accredited for diagnostic purposes or would be directly amenable to automation or high-throughput analysis. Both issues would need to be addressed for any workflow intended for clinical diagnostic purposes.

There are some limitations of the experimental approach and the methods themselves, which must be considered. Firstly, this work is on small numbers of samples and focuses on a single organism (*E. coli*) and does not evaluate the efficiency of these methods against Gram-positive bacteria, intracellular bacteria, fungi or viruses. Secondly, the Molysis method uses differential lysis of cells to release and degrade host DNA. As a result, cell free nucleic acids may be degraded meaning non-viable organisms may not be represented in analysis. Similarly, organisms lacking a robust cell wall (e.g. *M. pneumoniae*) or residing in host cells (*C. pneumonia*) may be degraded.

In summary, these experiments demonstrate that both the Molzym Molysis Complete 5 Extraction Kit and the Qiagen Blood and Cell culture DNA Maxi Kit were able to deplete host gDNA

whilst preserving *E. coli* DNA across a range of known concentrations of *E. coli* (10^1 to 10^4). The Molysis method had fewer manual steps, was completed in one workflow, was cheaper and performed better at host DNA depletion. For the purposes of mNGS of EDTA whole blood samples, the Molysis method would be a more appropriate choice as this method performed better at gDNA depletion but equally well at *E. coli* DNA extraction resulting of enrichment of prokaryotic DNA in the samples.

Chapter 5 Metagenomic DNA sequencing: results

There is currently no standardised method for mNGS analysis of human clinical samples for diagnostic pathogen detection and no laboratory accredited mNGS pathway for clinical samples in the UK or USA despite a growing body of literature on this topic. Any new diagnostic test used in the clinical setting must be rigorously testing, produce reliable and reproducible results and be accessible to the patients requiring diagnostic testing. This chapter reports the findings of DNA mNGS of selected EDTA whole blood and serum samples from participants with AUFI and healthy controls. The aim was to determine the clinical diagnostic utility of mNGS compared to standard diagnostic testing and to evaluate whether mNGS provided any additional diagnostic information to participants with AUFI who were undiagnosed by standard diagnostic testing. This work was performed in collaboration with Dr Gemma Kay (Senior Researcher) at the University of East Anglia led by Professor Justin O'Grady under a material transfer agreement.

5.1 Sample selection

Metagenomic sequencing of 30 matched EDTA whole blood and 29 serum samples (missing sample for SEP070) was done at the collaborating institute. Samples were selected by the lead investigators at the primary site to include a range of diagnosed and undiagnosed patients (with a range of different pathogens) as well as healthy volunteer samples (see table 5.1). All samples were anonymised prior to transfer and the receiving laboratory was blinded to the results of standard diagnostic testing. Four negative controls of molecular grade water were also included.

Table 5.1 Samples sent for metagenomic NGS

Pt. no.	Category	Results of standard of care	Diagnosis
SEP004	Undiagnosed	None	Undiagnosed
SEP007	Undiagnosed	None	Undiagnosed
SEP009	Undiagnosed	None	Undiagnosed
SEP014	Undiagnosed	Parainfluenza virus 3	Inconsistent with clinical picture
SEP026	Undiagnosed	None	Undiagnosed
SEP030	Undiagnosed	Lymphocytic CSF	Undiagnosed- no pathogen detected
SEP077	Undiagnosed	None	IE suspected- no pathogen detected
SEP005	Viral	HEV, IgM, IgG, RNA +ve	Acute Hepatitis E
SEP020	Viral	Dengue IgM, IgG and RNA +ve	Acute Dengue virus infection
SEP031	Viral	Influenza A H1N1 pdm09& HBV DNA	Influenza A & active HBV infection
SEP041	Viral	Dengue IgG and RNA +ve	Acute Dengue virus Infection
SEP059	Viral	Enterovirus RNA (blood and throat)	Acute Enterovirus infection
SEP066	Viral	Chikungunya RNA and IgM +ve	Acute Chikungunya virus infection
SEP099	Viral	EBV IgM, IgG and DNA positive	Acute EBV infection
SEP008	Bacterial	*P. aeruginosa* in blood cultures	*P. aeruginosa* bacteraemia- likely UTI
SEP016	Bacterial	*M. pneumoniae* serology & DNA +ve	Acute *M. pneumonia* infection
SEP033	Bacterial	*N. meningitidis* B PCR +ve blood	Meningococcal serogroup B
SEP032	Bacterial	*E. coli* bacteraemia MSU 10-20 WCC	*E. coli* bacteraemia
SEP052	Bacterial	*C. jejuni* from stool	*C. jejuni* gastroenteritis
SEP057	Bacterial	*C. difficile* 005 from stool	*C. difficile* colitis
SEP061	Bacterial	*S. enteritidis* from stool	*S. enteritis* gastroenteritis
SEP011	Bacterial	Perirectal abscess on CT	Pyogenic abscess
SEP029	Bacterial	ASOT consistent with *S.pyogenes*	*S. pyogenes* pharyngitis
SEP054	Not infection	None	Ruptured breast implant
SEP070	Not infection	None	Angioimmunoblastic T cell lymphoma
HEAL08	Healthy	Not applicable	Healthy
HEAL09	Healthy	Not applicable	Healthy
HEAL10	Healthy	Not applicable	Healthy
HEAL135	Healthy	Not applicable	Healthy
HEAL137	Healthy	Not applicable	Healthy
CONT01	Control	Not applicable	Water
CONT02	Control	Not applicable	Water
CONT03	Control	Not applicable	Water
CONT04	Control	Not applicable	Water

5.2 Metagenomic NGS analysis of EDTA whole blood and serum samples

Figure 5.1 illustrates the proportion of human, non-human and unclassified sequence reads identified from each participant sample from EDTA whole blood. Known contaminants of the sample preparation and sequencing process were identified and acknowledged in the analysis (see section 2.3 for a list of expected contaminants). Unclassified reads relate to those that could not be mapped to either a human or non-human genome within the reference database. There is much variety in the sequencing output between the samples but human genomic material predominates. The median proportion of human reads in EDTA whole blood amongst the participants (SEP004 to SEP099) was 90.7% (IQR 82.9-92.6%). Similar proportions of non-human (4.8%, IQR 3.7-8.8%) and unclassified reads (4.6%, IQR 3.6-8.3%) were identified across these samples although there was a wide range of results (human read range: 6.0-93.7%, non-human read range 3.1-56.0%, unclassified read range 3.1-48.8%). The control samples (Control01 to Control04), which consisted of molecular grade water, also contained some human reads despite containing no biological material albeit at a much lower proportion of the total sequence reads than the biological samples (median 5.6%, IQR 6.0-20.2%), this suggests contamination with human biological material at some stage during sample preparation and processing.

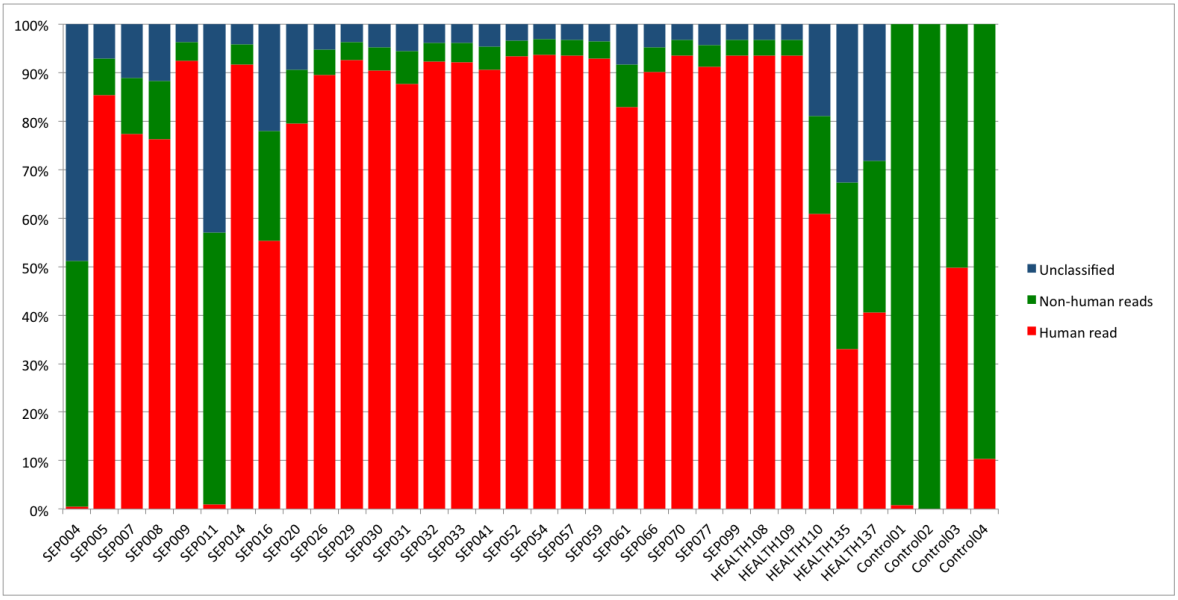


Figure 5.1 Summary of mNGS reads per participant EDTA whole blood sample, showing proportions of human (red), non-human (green) and unclassified (blue) reads proportionate to the total sample sequence reads.

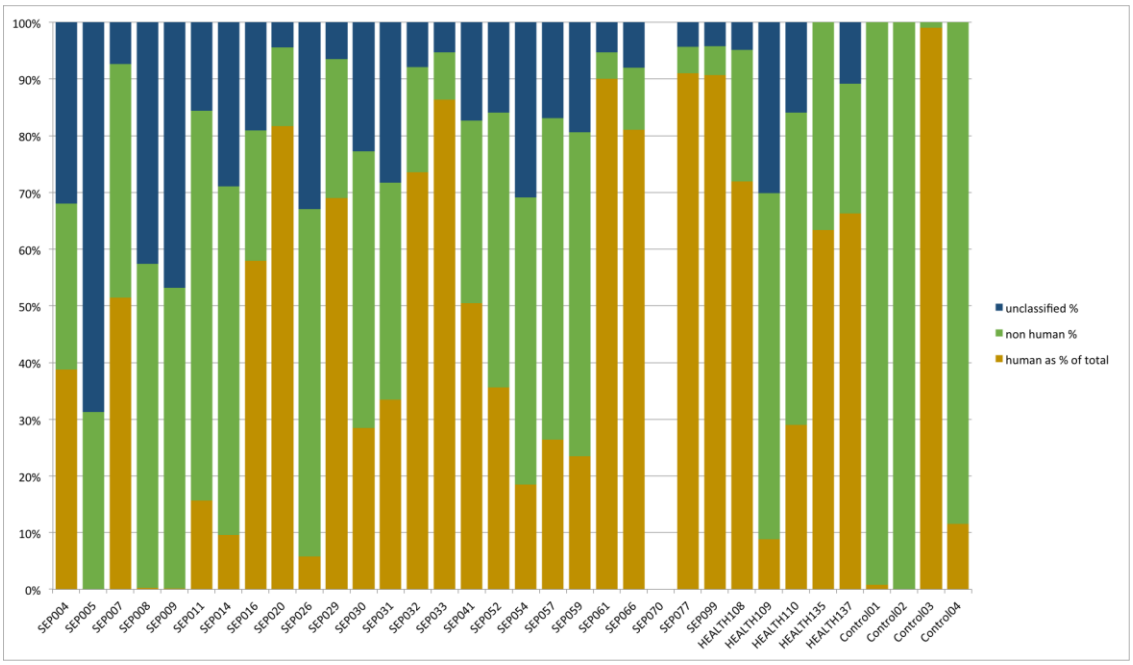


Figure 5.2 Summary of mNGS reads per participant serum sample, showing proportions of human (yellow), non-human (green) and unclassified (blue) reads proportionate to the total sequence reads.

Figure 5.2 shows the same data for serum samples. No serum sample was available for participant SEP070 and therefore this data is not shown. A lower proportion of human reads were detected in serum samples than from EDTA whole blood in AUFI participants (serum median 35.9%, IQR 18.3-74.7%; EDTA whole blood 90.7%, IQR 82.9-92.6%) and this difference was significant (difference -54.75, 95% CI -60.3 to -16.5, p<0.0001) (figure 5.3) which may be expected as serum is separated from the cellular component of blood and therefore lower levels of host genomic material would be expected.

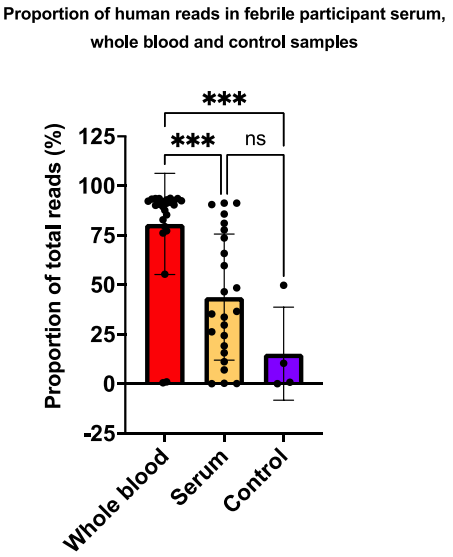


Figure 5.3 Proportion of human reads identified in EDTA whole blood (red) and serum (orange) in AUFI participants and molecular grade water control samples (purple)

A statistically significantly higher proportion of non-human reads were detected in serum from AUFI participants (median 40.3%, IQR 17.2-54.8%) than from EDTA whole blood (whole blood median 4.8%, IQR 3.7-8.8%) (Difference 35.5, 95% CI 13.7 to 45.3, p<0.0001) (see figure 5.4). The molecular grade water controls contained significantly higher proportions of non-human reads than both serum and EDTA whole blood which as expected, although it should be remembered that these samples had much lower read numbers in comparison.

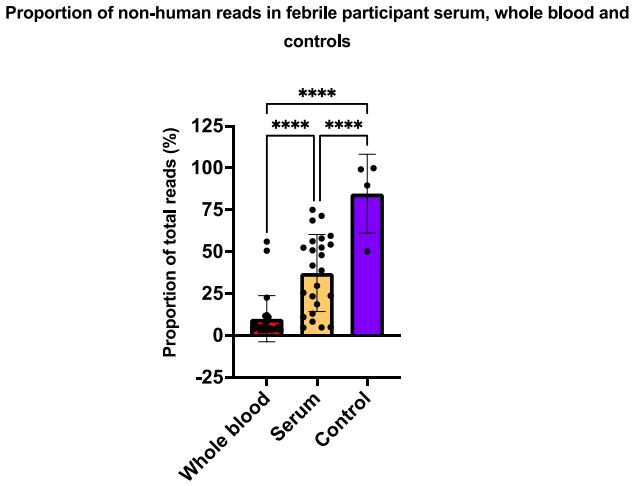


Figure 5.4 Proportion of non-human reads detected in EDTA whole blood (red) and serum (orange) in AUFI participants and molecular grade water control samples (purple)

Unclassified reads were significantly more common in serum samples from AUFI participants (median 16.4%, IQR 6.7-31.3%) than in EDTA whole blood samples (median 4.6%, IQR 3.6-10.6%) (difference 35.5, 95% CI 13.7 to 45.3, p<0.0001) (see figure 5.5).

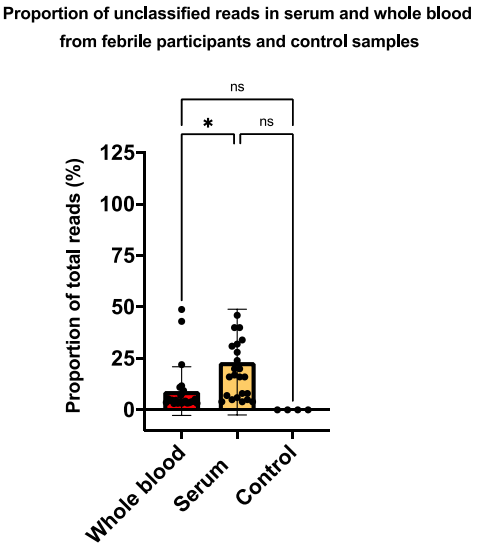


Figure 5.5 Proportion of unclassified reads from EDTA whole blood (red) and serum (orange) obtained from AUFI participants and molecular grade water control samples (purple)

5.2.1 AUFI participants versus healthy controls

No significant differences were identified in the proportions of human, non-human and unclassified reads in serum and EDTA whole blood from healthy volunteers and AUFI participants (see figure 5.6) however it must be noted that the number of healthy volunteers analysed was very small (n=5).

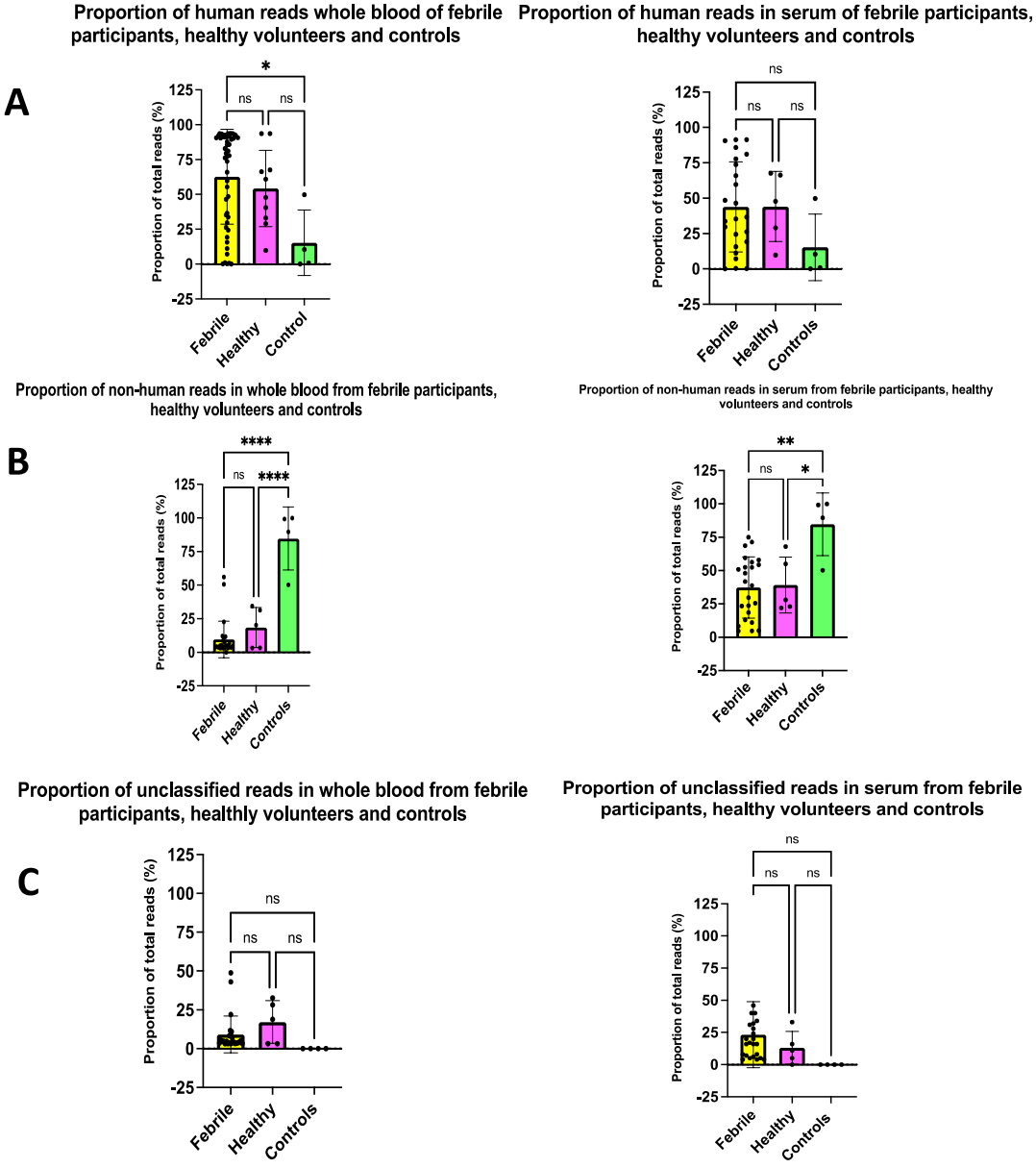


Figure 5.6 Comparison of proportion of human reads (A), non-human reads (B) and unclassified reads (C) in serum and EDTA whole blood from febrile participants and healthy volunteers

5.2.2 Non-human mapped reads

5.2.2.1 Contamination from methodology

Figure 5.7 (EDTA whole blood) and figure 5.8 (serum), illustrates the proportion of non-human reads (shown in green in figures 5.1 and 5.2) mapped to different organisms within each participant sample. All organisms coloured blue are known contaminants associated with the method (see section 2.3 for details).

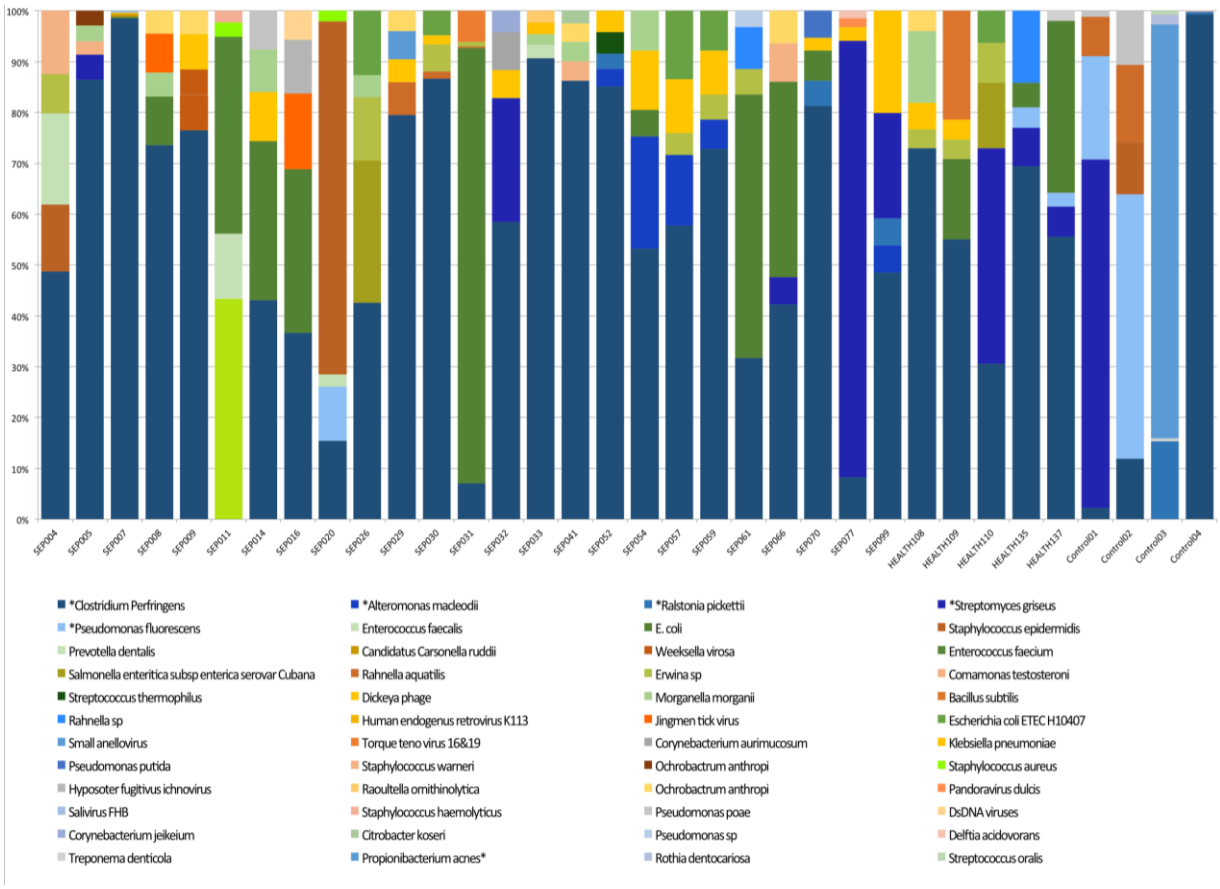


Figure 5.7 Non-human reads in EDTA whole blood by participant. Organisms identified as contaminants from the method are marked with a * in the legend and are depicted in blue.

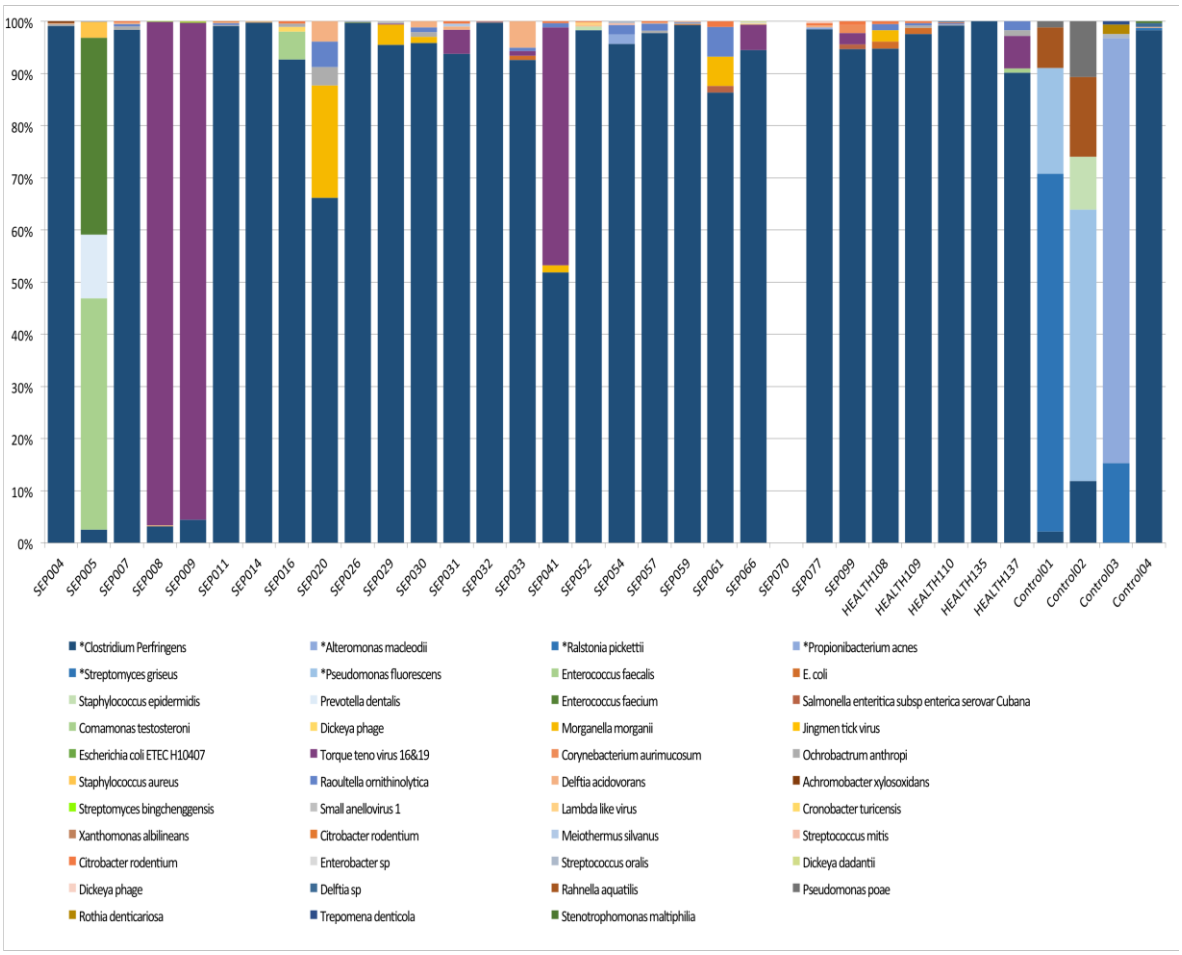


Figure 5.8 Non-human mapped sequence reads in serum by participant. Organisms identified as contaminants from the methodology are marked with a * in the legend are depicted in blue.

Contaminating organisms reflect a large proportion of the non-human mapped reads in EDTA whole blood and serum across all AUFI participants with a significantly higher proportion of contaminants in serum than EDTA whole blood (EDTA whole blood median 76.5%, IQR 43.1-86.3%; serum median 95.7%, IQR 91.1-98.9%) (difference 21.45, 95% CI 11.3-25.9, p<0.0001) (see figure 5.9) although wide variability is seen (EDTA whole blood, range 0.0-98.5%, serum range 0.0-98.5%).

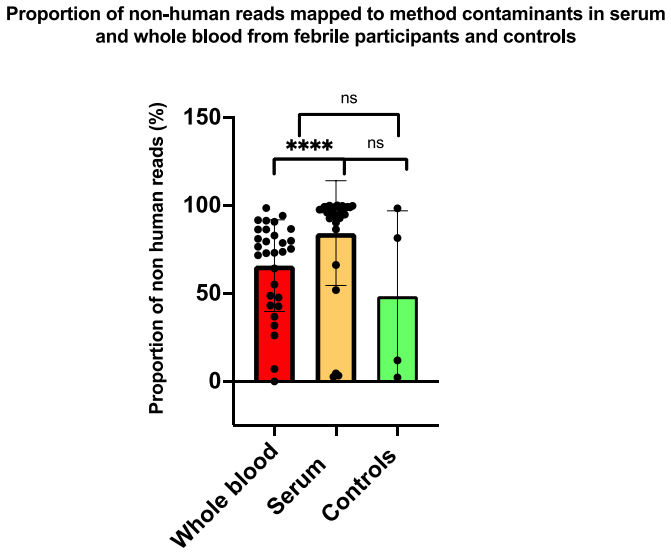


Figure 5.9 Proportion of non-human reads mapped to known contaminants of sequencing methodology in EDTA whole blood (red) and serum (orange) taken from AUFI participants and molecular grade water controls (green)

There were no significant differences between the proportion of reads mapped to contaminating organisms in febrile participants versus healthy controls (difference 1.6, 95% CI -9.9 to 21.0, *p*= 0.53), consistent with a common source of contamination across all specimen types which would fit with methodological contamination (see figure 5.10).

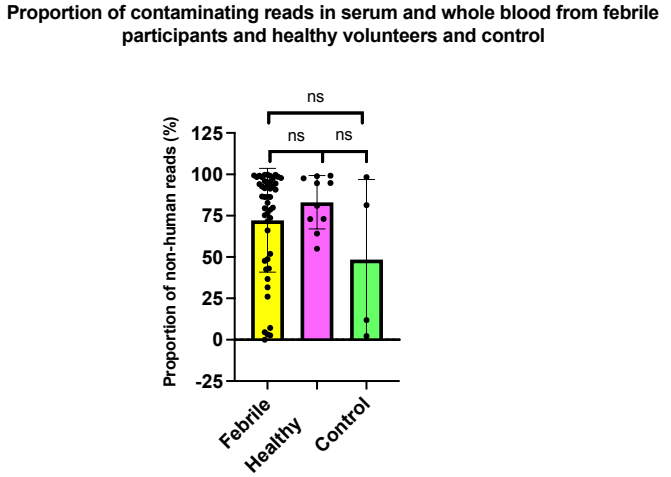


Figure 5.10 Proportion of reads mapped to contaminating organisms in combined serum and EDTA whole blood from AUFI participants (yellow) and healthy volunteers (pink) and molecular grade water controls (green)

5.2.2.2 Non-human mapped reads with contamination removed

Once known laboratory contaminants were removed, the remaining mapped non-human reads identified by mNGS can be seen in figure 5.11 (EDTA whole blood) and figure 5.13 (serum). However, the non-human reads reflect a very small proportion of the overall sequence output for

each sample with only one EDTA whole blood sample (SEP011) and one serum sample (SEP005) having a relatively high number of reads remaining following contamination and human read removal (see figure 5.12 EDTA whole blood and figure 5.14 serum).

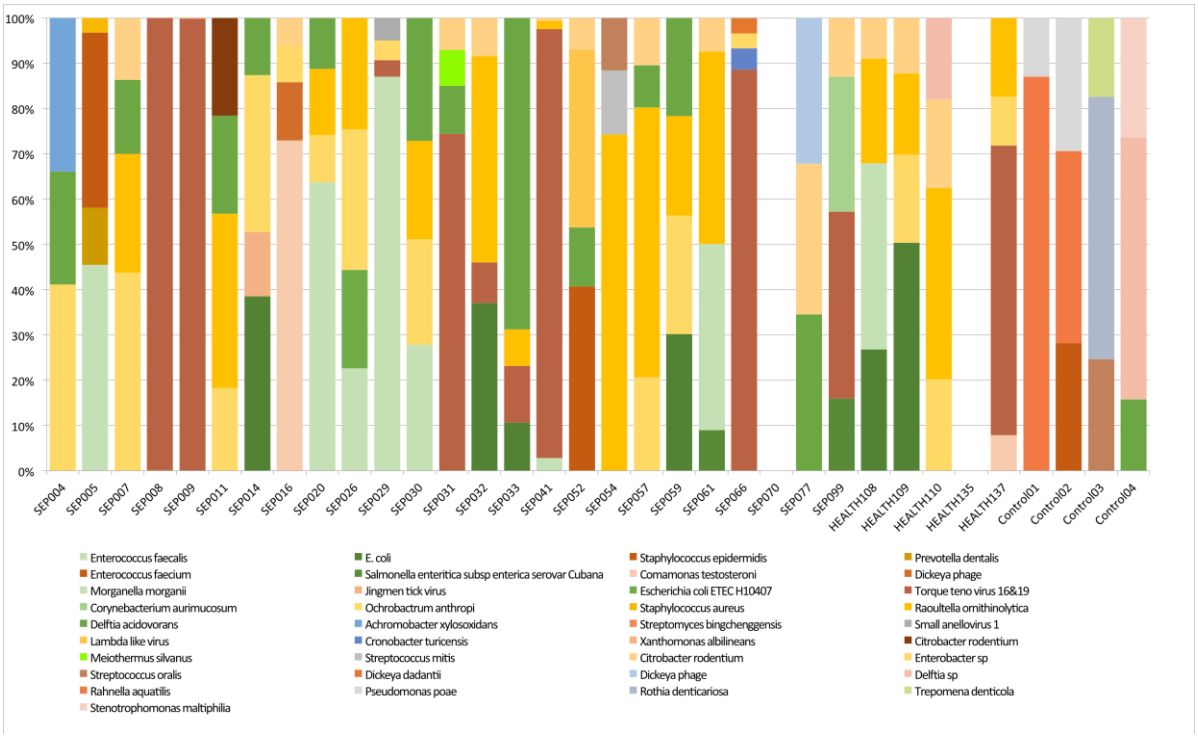


Figure 5.11 Proportion of non-human mapped reads by participant in EDTA whole blood once contaminants are removed

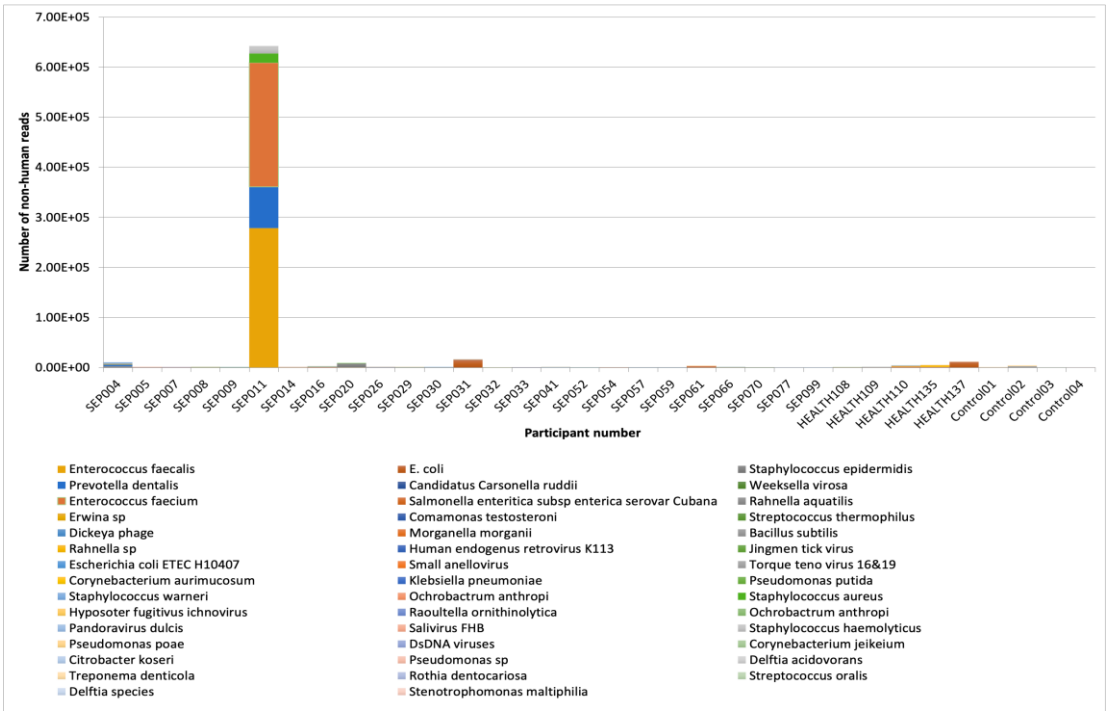


Figure 5.12 Total non-human mapped reads by participant EDTA whole blood sample, contamination removed

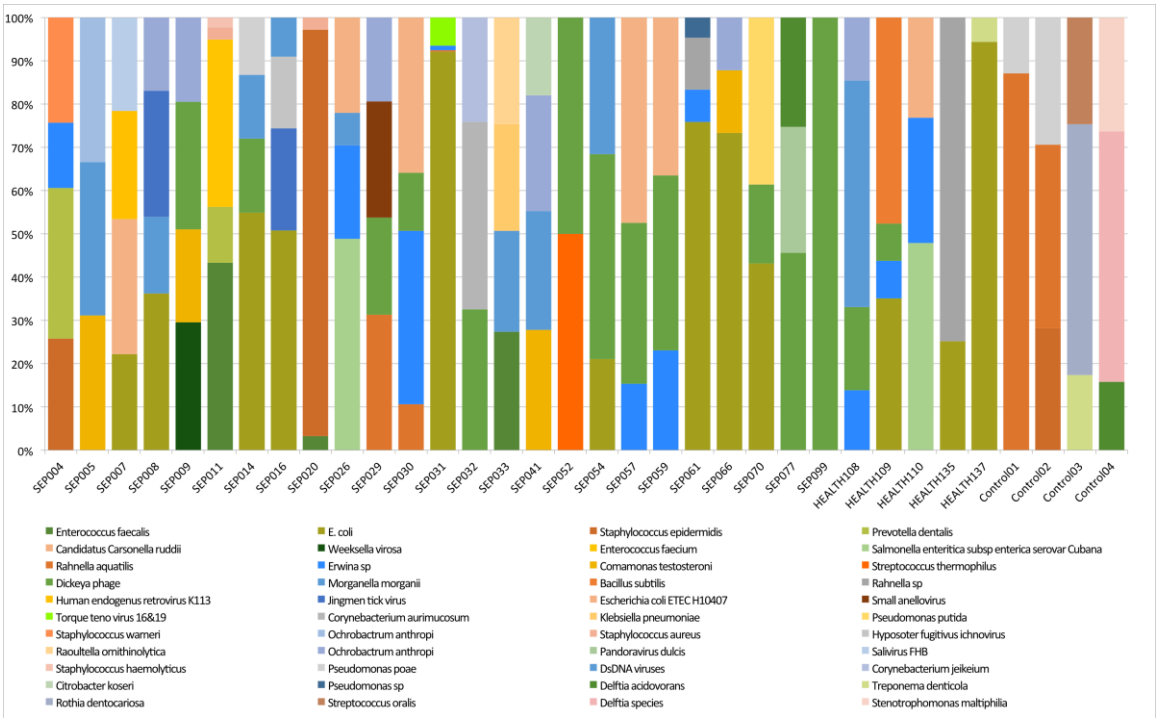


Figure 5.13 Proportion of non-human mapped reads by participant in serum once contaminants are removed

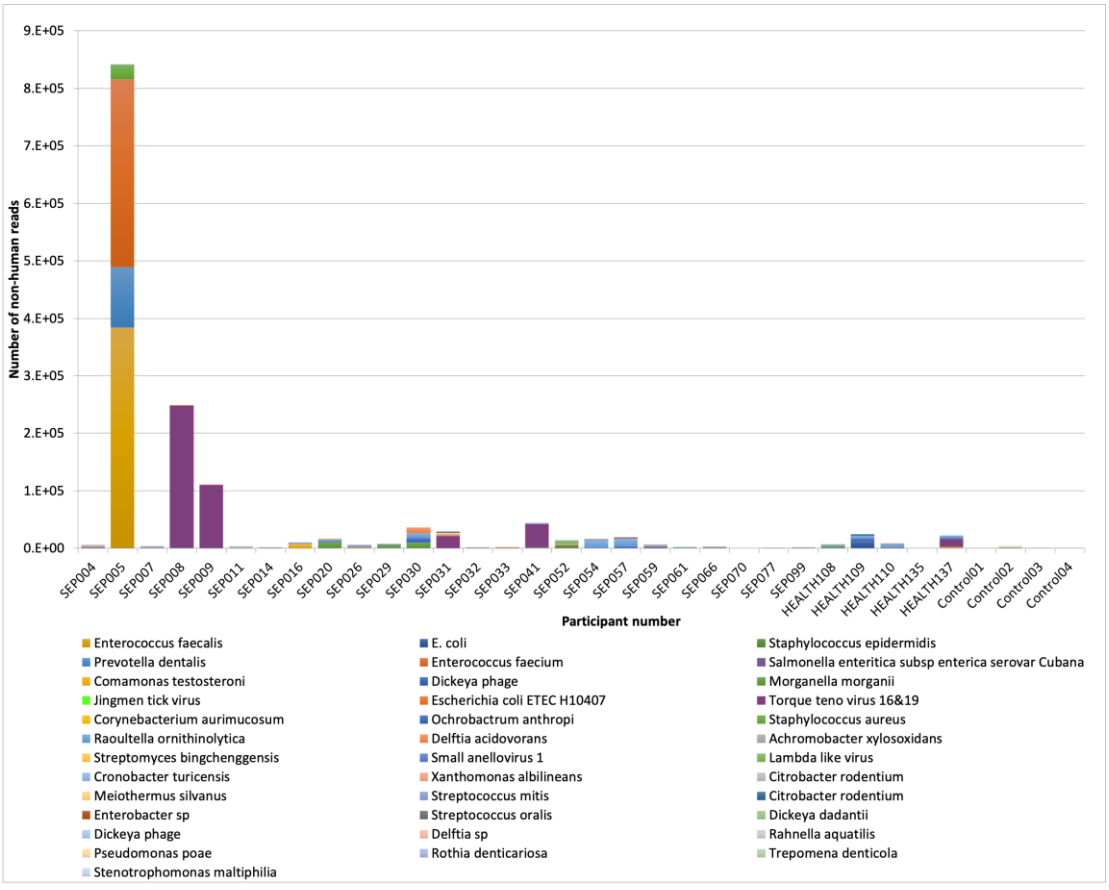
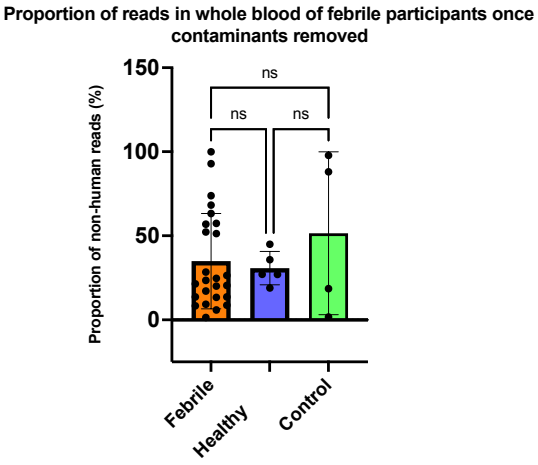


Figure 5.14 Total non-human mapped reads by participant serum sample, contamination removed

There was no significant difference between the proportions of reads mapped to non-contaminating organisms in EDTA whole blood AUFI participants compared with healthy volunteers (figure 5.15, A). Although there was a higher proportion of reads mapped to non-contaminating organisms in serum (figure 5.15, B), this difference was not statistically significant (difference 2.1, 95% CI -31.5 to 2.5, *p*=0.36).

A



B

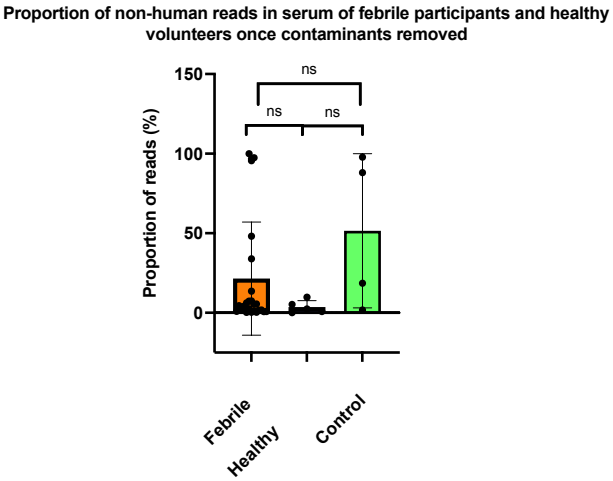


Figure 5.15 Proportion of mapped non-human reads once contamination removed in EDTA whole blood (A) and serum (B) for febrile participants (orange) and healthy volunteers (purple) and molecular grade water controls (green)

5.3 Clinical correlation of mNGS data

The following section reports the mNGS findings excluding known contaminants and focusing on organisms known to cause human disease in the Health and Safety Executive 'The Approved List of biological agents' [207] (see section 2.3.5 for full details) for each participant analysed.

5.3.1 Bacterial infection

The mNGS results of participants diagnosed with bacterial infection by standard of care diagnostics are presented in table 5.2 and figure 5.16.

Table 5.2 MNGS results of participants diagnosed with bacterial infection by standard of care diagnostics

Pt. no.	standard diagnostics	mNGS serum and EDTA whole blood (total reads)
SEP008	*P. aeruginosa*	*M. morganii* (112), *E. coli (*343), TTV (248280)
SEP011	Perirectal abscess- no pathogen	*E. faecalis* (278362*), E. faecium* (248696), *P. dentalis* (82274), *S. aureus* (18119), TTV (251)
SEP016	*M. pneumoniae*	*E. coli* (1319)
SEP029	*S. pyogenes*	*M. morganii* (6018)
SEP032	*E.coli*	*C. jeikeium* (20), TTV (35), *C. aurimucosum* (138), *E. coli* (144)
SEP033	*N. meningitidis*	*M. morganii* (52), *K. pneumoniae* (18), *E. coli* (215), TTV (230)
SEP052	*C. jejuni*	*S. thermophilius* (52)
SEP057	*C. difficile* 005	*E. coli* (27)
SEP061	*S. enteritidis*	*M. morganii* (641), *S. enteritica* (140)

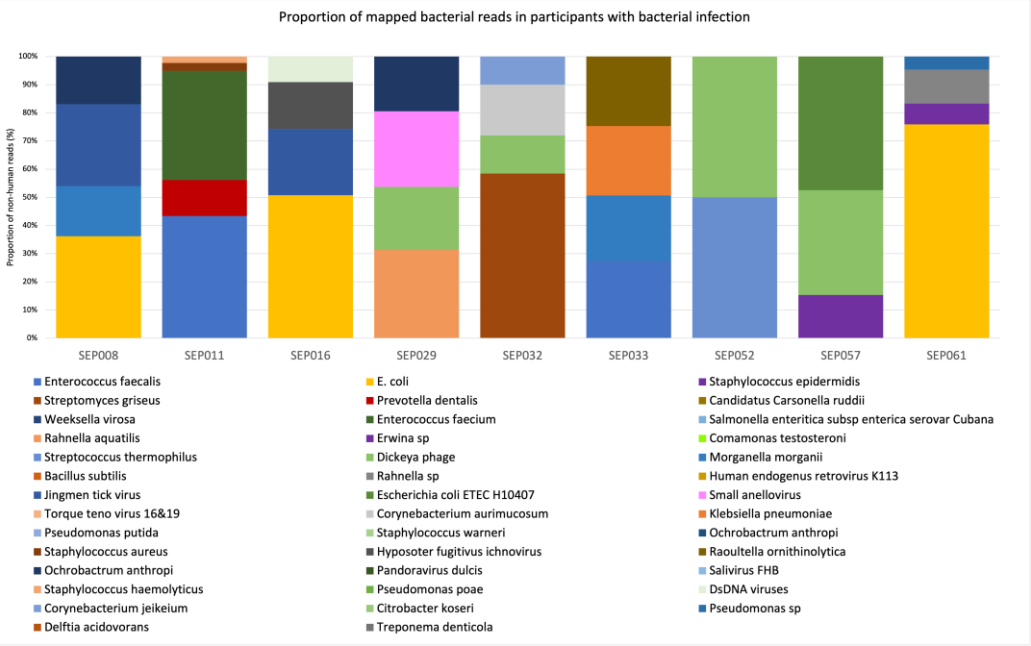


Figure 5.16 Proportion of mapped reads in participants with confirmed bacterial infection

Whilst for the majority samples, the organism(s) identified through standard diagnostic testing were not seen in mNGS there are two possible exceptions. Firstly, participant SEP032 was diagnosed with *E. coli* bacteraemia by standard diagnostic testing. *E. coli* was also the most prevalent potential pathogen identified by mNGS analysis in this individual (144 reads). However, if compared to other participants within the group, three participants (SEP005, SEP016 and SEP033) all had higher read numbers and higher proportions of non-human reads mapped to *E. coli* (depicted in yellow in figure 5.16) despite alternative clinically credible infections being diagnosed by standard diagnostics (*P. aeruginosa, M. pneumoniae, N. meningitidis*, respectively). Direct sequence comparison of the *E. coli* isolated from blood cultures in participant SEP032 to that identified by mNGS was not performed. It is therefore not possible to know whether the reads detected by mNGS reflect the infecting organism, however with the wide presence, and higher burden of *E. coli* mapped reads in participants without proven infection with *E. coli*, it seems most likely these findings are due to environmental or procedural contamination as opposed to a clinically significant finding.

The second finding of interest is participant SEP011, this individual had a perirectal and perianal abscess identified by CT imaging but no infecting organism identified by standard of care diagnostics. Metagenomic NGS identified reads mapped to four bacterial species (*Enterococcus faecalis*, *Enterococcus faecium, Prevotella dentalis, S. aureus*) and TTV. *E. faecium and E. faecalis* are part of the human gastrointestinal microbiome but can also cause invasive disease. Most commonly associated with urinary tract infections, infective endocarditis and infections associated with the GI tract [209]. *P. dentalis* is a strict anaerobe found in the oral and

gastrointestinal tract and associated with dental abscess formation [214]. Case reports in the literature also document deep seated infection with *P. dentalis* at other sites including intra-abdominal abscess formation in a peritoneal dialysis patient [215] and a pyogenic liver abscess in an immunocompetent male associated with a dental abscess [216]. *S.aureus* is a common skin commensal but also well established as a significant human pathogen causing multi-site disease including abscess formation [217]. All four organisms could plausibly be implicated in causing a perirectal abscess in this participant. Unlike *E. coli*, *E. faecium*, *P. dentalis* and *S. aureus* were not widely identified from other participants, *E. faecalis* reads were only detected in one other participant (participant SEP020 Dengue viral infection), *P. dentalis* in two participants (participant SEP005, Hepatitis E (HEV) infection, undiagnosed participant SEP004) and *S. aureus* two participants (participant SEP005 HEV infection and participant SEP020 Dengue virus infection). None of these organisms were detected in healthy controls or non-infection diagnosed participants. As standard of care microbiological tests were negative in participant SEP011 it is impossible to confirm whether the organisms identified by metagenomic NGS are of clinical significance, there is clinical plausibility due to the association of organisms detected with the focus of infection however these findings may also be due to chance only. Of note, participant SEP005 had confirmed acute HEV infection had a similar range and proportion of bacterial species identified by mNGS as participant SEP011 (see samples SEP005 and SEP011 in figures 5.7, 5.8, 5.12 and 5.14) which could be explained by several reasons. These include; participant SEP005 having undetected bacterial translocation from their GI tract to the blood as a result of an active viral hepatitis which was not detected by standard diagnostic testing, cross-contamination between the participant's samples, that the serum and EDTA whole blood samples for these participants were mixed up, or that the findings in participant SEP005 and SEP011 are not clinically significant and a result of chance. Unfortunately, it is not possible to draw a conclusion on the currently available data as no standard of care diagnosis was made in SEP011 and human genome was not examined and therefore in cannot be concluded if the serum from SEP005 and EDTA whole blood from SEP011 were actually from the same participant.

Of note, three of the participants (SEP052, SEP057, SEP061) had a diagnosis made by pathogens isolated from stool; *C. jejuni*, *C. difficile* and *S. enteritidis* respectively. These three organisms isolated often cause localised gastrointestinal tract infection and are only very rarely recovered from blood cultures. Therefore, there may have been very little/no pathogenic DNA present in the blood samples sent for analysis.

5.3.2 Viral infection

Table 5.3 shows the mNGS findings for participants diagnosed with viral infection by standard investigations. All the viruses identified by standard investigations had an RNA genome with the exception of EBV. This is significant as only DNA mNGS was performed. Furthermore, RNA is more labile and vulnerable to degradation than DNA so RNA levels within the sample when analysed may be lower than the time of sample acquisition. No clinically significant viruses were detected by mNGS in this cohort of participants. TTV was the only virus detected by mNGS and this was detected in 57.1% (4/7) of participants with known viral infection compared with 44.4% (4/9) with bacterial infection, 50% (1/2) healthy controls, 14.3%(1/7) of participants with undiagnosed AUFI and non-infection cases. Overall, DNA mNGS did not demonstrate clinically diagnostic utility when compared to standard of care diagnostics in patients with a range of viral infection.

Table 5.3 MNGS of participants with viral infection confirmed by standard of care diagnostics

Pt.	Pathogens by Standard	Metagenomic NGS results (reads)
SEP005	HEV	*M. morganii* (108), *S. aureus* (26118), *P. dentalis* (105596), *E. faecium* (326306), *E. faecalis* (383467)
SEP020	Dengue virus	*S. aureus* (248), *E. faecalis* (290), *M. morganii* (10835)
SEP031	Influenza A H1N1 pdm09 & HBV	*E. coli* (15136), TTV (22705)
SEP041	Dengue virus	*M. morganii* (1282), TTV (40951)
SEP059	Enterovirus	*E. coli* (2009)
SEP066	Chikungunya	*E. coli* (492), TTV (1710)
SEP099	EBV	*S. enteritica* (54), *C. aurimucosum* (101), TTV (140)

5.3.3 Non-infection and healthy volunteers

Metagenomic NGS data from participants with non-infection diagnoses and healthy volunteers are presented in table 5.5. *E. coli* is present in all participant samples but as previously discussed the clinical significance of this is not clear. Equally *M. morganii*, *S. enteritica*, and *T. denticola* detected in healthy controls who were well two weeks following recruitment are very unlikely to be of any clinical significance as symptoms would be expected to develop from a bacterial blood stream infection by this time. Samples from participant SEP054 (non-infection) found *S. mitis* and *S. oralis*, which are both oral streptococci and often contaminate clinical specimens, only felt to be significant if recurrently isolated or in the context of infective endocarditis.

Table 5.5 MNGS results of participants with non-infection diagnoses and healthy volunteers

Pt. no.	Pathogens by Standard	Metagenomic NGS results (reads)
SEP054	Ruptured breast implant	*S. mitis* (2302), *S. oralis* (1877), *E. coli* (4), *M. morganii* (6)
SEP070	Angioimmunoblastic T cell	*E. coli* (19),
HEALTH08	Healthy control	*E. coli* (1659), *M. morganii* (2609)
HEALTH09	Healthy control	*E. coli* (12646)
HEALTH10	Healthy control	*E. coli* (1068), *S. enteritica* (2208)
HEALTH135	Healthy control	*E. coli* (1256)
HEALTH137	Healthy control	*T. denticola* (700), *E.coli* (11723), TTV (14438)

5.3.4 Undiagnosed participants

Table 5.6 shows mNGS results for participants with an unknown cause of AUFI. Again, *E. coli* was widely present in 71% (5/7) of participant samples and in isolation in two participants (participants SEP007 and SEP077) but in very low numbers. *E. coli* would have been expected to be readily cultured from standard of care diagnostic tests such as blood cultures if causing clinically significant infection and so its role as the causative pathogen is unlikely. *M. morganii* was isolated from 42.9% (3/7), however, *M. morganii* infection would be inconsistent with a lymphocytic meningitis (SEP030) and should be easily cultured from standard of care investigations. TTV was present in high levels and in isolation in participant SEP009, however, the role of TTV's in human infection is yet to be fully understood and TTV has been frequently identified in healthy individuals [195] so it's clinical significance here is unclear.

Table 5.6 MNGS results of participants with no diagnosis made by standard of care diagnostics

Pt. no.	Pathogens by Standard	Metagenomic NGS results (reads)
SEP004	None	*P. dentalis* (3640),
SEP007	None	*E. coli* (39)
SEP009	None	TTV (109963)
SEP014	Parainfluenza virus 3	*M. morganii* (51), *E. coli* (51)
SEP026	None	*S. enteritica* (292), *M. morganii* (1453), *E. coli* (1481)
SEP030	Lymphocytic CSF	*E. coli* (51), *M. morganii* (10108),
SEP077	None	*E. coli* (39)

5.4 Limitations of bioinformatics analysis

This study used a single bioinformatics tool (Kraken) for the DNA bioinformatics analysis. RNA was not analysed. Using a single bioinformatics tool limits the analysis to species represented in that database and so misclassification errors may not be identified. Organisms may be missed or misclassified as a result. When directly compared using the same dataset, different bioinformatics pipelines have been shown to have different sensitivities [218] supporting the use of more than one bioinformatics tool. Pooling the output from several different bioinformatics analyses may have provided more reliable and comprehensive results.

In addition, further analysis of the current results could have been performed to compare with the mNGS analysis. For example, whole genome sequence of organisms detected by standard of care diagnostics could be directly compared to the sequence data generated by mNGS to help determine whether mNGS was picking up the clinically significant organism or background contamination. This would be particularly useful in case SEP032 where *E.coli* was detected by both standard of care testing and mNGS. In addition mNGS reads could be mapped to a reference genome to determine whether they map to regions across the entire target genome which would be more consistent with true presence of that organism. Reads mapping to isolated regions of the genome would be more consistent with contamination or DNA present in the reagents or blood sample collection tubes.

5.5 Summary

Metagenomic NGS was performed on EDTA whole blood and serum samples from a selection of thirty individuals comprising of twenty-five participants with AUFI and five healthy volunteers. Overall, serum samples were found to have significantly fewer human reads, more non-human reads and more unclassified reads with a significantly increased proportion of reads mapped to contaminating organisms when compared to EDTA whole blood samples. The high proportion of contamination highlights an on-going and unsolved issue with this approach. There were no significant differences in the proportion of human, non-human and unclassified reads between AUFI participants and healthy controls although the numbers studied were very small.

In one participant, (SEP011) the non-human reads in EDTA whole blood represented a high proportion of the total sample reads. Moreover, the organisms detected were clinically plausible agents in the participant's disease process (peri-anal abscess) and were not widely detected in the other participants. However, as no microbiologically confirmed diagnosis was made using standard of care diagnostics, it is impossible to determine whether these findings were due to chance or a true reflection of infecting organisms.

Overall, with the possible exception of one participant, DNA mNGS performed in this study failed to detect bacterial and viral pathogens detected by standard of care investigations. Furthermore, it did not clearly identify a clinically credible cause of AUFI in participants with undiagnosed AUFI. Both *E. coli* and TTV were widely identified in participants with AUFI and healthy volunteers but their clinical significance could not be determined. Further research is needed into the clinical significance of TTV in AUFI.

There are several limitations to this work, firstly, a very small and diverse selection of samples were analysed with differing numbers of each sample type (bacterial n=9, viral n=7, unknown n=7, non-infected n=2, healthy controls n=5, negative controls n=4) making it challenging to draw direct comparisons between the groups. This approach was taken to try and maximise the yield from a limited resource without compromising further analysis of the research samples. A larger, more balanced selection of samples is required to understand the true diagnostic utility of mNGS. Secondly, the rationale for using the HSE 'The Approved List of biological agents' as an objective reference guide was to prevent falsely attributing a disease to ubiquitous environmental/ contaminating organisms - a complication described of using mNGS as a diagnostic tool [219]. However, by using this approach, it is possible that novel pathogens may have been excluded from the analysis. Thirdly, all samples were frozen prior to analysis with no prior treatment to preserve genomic material. Both freeze-thawing and endogenous DNAses or RNAases present in the sample may have caused the breakdown of pathogen associated genomic material prior to

analysis. Fourthly, of the participants with confirmed bacterial infection, three participants had organisms isolated from stool specimens, which are rarely identified in the blood stream, which may limit diagnostic utility of mNGS performed on blood samples. Finally, this piece of work only performed DNA analysis. Dengue virus, Chikungunya virus, Enterovirus, Influenza virus and Hepatitis E virus, being RNA viruses, would therefore have been missed by this approach.

Despite the findings of this study, the evidence base for mNGS as a diagnostic tool for infection is growing. Publication bias must be acknowledged as a potential confounder as studies with positive outcomes are more likely to be accepted and published in journals. However, in spite of this, there is convincing data that mNGS should have a role in clinical diagnostic laboratory although its exact positioning is yet to be fully understood.

Case studies demonstrate that mNGS can diagnose infection where other methods have failed and include a severe case of neuroleptospirosis (identified in CSF) [190], two cases of encephalitis attributable to Astrovirus [191] (CSF and brain tissue) [220] (CSF, stool, serum and throat swab), a paediatric case of encephalitis due to *Balamuthia mandrillis* (CSF) [194], a case of *Naegleria fowleri* encephalitis (CSF, brain tissue) [221], a case of Leishmaniasis [192] (peripheral blood), and assisted in the post mortem diagnosis of Rabies encephalitis (brain tissue) [193]. Untargeted mNGS had identified novel infectious agents such as Bas Congo virus [170] and a novel Paramyxovirus in a returning traveller from Sudan and Uganda who presented with fever, headache and myalgia [222].

Adding to the case studies, broader clinical studies have been performed. Of note, a study of meningoencephalitis revealed mNGS alone identified a plausible causative pathogen in (22%) 13/58 of cases. Of the diagnoses confirmed by standard testing, mNGS identified 42.2% (19/45) of pathogens. Those not identified by mNGS were either diagnosed on serology alone (42.3%; 11/26), tissue samples rather than CSF (26.9%; 7/26) or had low titres of the infective pathogen 30.8% (8/26) [223]. This study is particularly relevant to the clinical field as the mNGS results were available in a clinically relevant time frame and 53.8% (7/13) of mNGS results directly influenced patient management. Two studies found mNGS of CSF samples moderately sensitive and very specific for the diagnosis of TB meningitis [224,225] with Zhou *et al* finding the best diagnostic sensitivity when standard testing was combined with mNGS [224. This is particularly relevant as TB meningitis is notoriously difficult to diagnose and associated with poor clinical outcomes [226]. Schimidt *et al* found nanopore mNGS to be a useful tool to identify organisms and acquired resistance genes in urine samples in around eight hours [150] and in another study, were able to accurately identify organisms causing lower respiratory tract infection and their antimicrobial sensitivity profile from respiratory samples [199].

The studies described here illustrate how mNGS may be used to improve infectious disease diagnostic capabilities on a range of sample types, however, more relevant to this UK study are studies of serum and EDTA whole blood samples. Yozwiak *et al* used mNGS to analyse serum of children with undiagnosed fever and found a previously undetected virus in 37% (45/123), however, there were two key limitations of this work, firstly it is unlikely that all the viruses identified by mNGS were specifically looked for by standard diagnostic testing due to a lack of assays available to detect them (Dengue virus, hepatitis A virus (HAV), HHV-6, parvovirus, TTV, laviviridae, Circoviridae, Anelloviridae, Asfarviridae, and Parvoviridae). This may mean the true sensitivity of mNGS may be lower than stated when compared to standard diagnostics. Secondly, with the exception of Dengue virus, HAV, HHV-6 and Parvoviridae (specifically Parvovirus B19) the role of most of these viruses in paediatric febrile illness is not understood and therefore their clinical significance in these patients cannot be confidently assigned.

EDTA whole blood requires little or no manipulation once obtained and would be a practical choice for untargeted mNGS for patient diagnostics. Despite this, mNGS of EDTA whole blood has not been well studied. Greninger *et al* analysed whole blood from patients with Ebola virus, plasma from patients with Chikungunya and serum from patients with HCV infection. They found mNGS was able to correctly identify the viral pathogens within 6 hours using real-time nanopore sequencing on all samples [227]. Vijayvargiya *et al* analysed whole blood samples from eight participants with known vector borne infection (*Babesia microti*, *Borrelia hermsii*, *Plasmodium falciparum*, *Mansonella perstans*, *Anaplasma phagocytophilum* or *Ehrlichia chaffeensis).* Four different bioinformatics tools were used and mNGS identified five of eight specimens to species level [228]. A poster presentation at European Congress of Clinical Microbiology and Infectious Disease (ECCMID) 2018 by Hedberg *et al* found mNGS of whole blood correctly identified organisms in four bacteraemic participants and six whole blood samples spiked with *E. coli* and *S. aureus* when two bioinformatics pipelines were used in parallel, however this work is yet to be published in a peer reviewed journal [229].

The published literature supports mNGS as a diagnostic tool for infectious disease with most studies showing mNGS can add clinically useful diagnostic information. This view is also supported by recent meta-analysis of mNGS diagnostic accuracy studies [198]. There are several possible reasons why this UK based study deviates from the published literature.

This study did not confirm or exclude the diagnostic potential of mNGS in adults presenting to hospital with AUFI, however, the problems of contamination and high levels of human DNA encountered with the analysis are well described in the published literature. Salter *et al* demonstrate the significant burden of contaminating organisms/DNA in reagents used for both

16S PCR and mNGS such that in samples with a low abundance of a particular pathogen, contaminating sequence reads dominate the sequence output. They document a list of identified contaminants and recommend this approach be adopted in laboratories undertaking mNGS [230]. Glassing *et al* support this approach, their work identified sequence reads mapping to 88 different contaminating organisms in extraction and sequencing reagents when molecular grade water was run in place of blood. Within this list, Enterobacteriaciae, Enterococci and Staphylococci [219], all well established human pathogens and all present in the mNGS data generated in this UK based study of AUFI. Additional to contamination associated with reagents, Greninger *et al* demonstrated reads mapping to Chikungunya could be detected on a subsequent run of a sample from a patient with Ebola due to contamination of the flow cell occurring during library preparation of the previous Chikungunya positive sample illustrating another potential source of contamination [227].

High proportions of human DNA in clinical samples has been recognised as a key limitation of mNGS [150, 206, 231, 232]. Much like contamination from prokaryotes, human DNA can mask pathogens present in the sample but dominating the sequence output. In this study over 90% of mapped reads from EDTA whole blood were mapped to the human genome, the proportion was lower in serum but still accounted for approximately a third of all mapped reads. Yang *et al*, found similar proportions of human reads in nasopharyngeal swabs of children which accounted for 91.6% of all mapped reads [233] and others report proportions even higher [191]. Much work has tried to address this problem and some success demonstrated with saponin based detergents [150, 209, 232]. High proportions of human DNA may make EDTA whole blood less amenable to analysis by mNGS for pathogen detection and this may partly explain why some of the more successful studies have focussed on less cellular sample types such as CSF and more recent approaches have included cell free DNA analysis [234].

In addition to the methodological limitations, there is currently no standard mNGS method for pathogen detection in clinical specimens, no clear guidance on what thresholds or parameters must be met for an detected organism to be deemed significance and no standardised interpretation of mNGS results.

Further work should include both DNA and RNA sequencing performed on either fresh samples or samples treated to protect nucleic acids from degradation. Preventing exposure to freeze-thaw events to protect against further loss of genomic material and analysing a wider range of specimens such as stool samples should be considered. Analysis should be performed on a larger selection of samples ideally with equal numbers of samples from participants with fever of ‘bacterial’, ‘viral’, ‘non-infection’ and ‘unknown’ aetiologies. An equal number of matched

healthy controls should also be included to improve the validity of any differences. The data from this set of experiments could be used to help inform a power calculation to inform the number of samples needed to confidently examine any differences/ similarities between the groups.

Chapter 6 Discussion

6.1 Key findings and interpretation in the current literature

This proof of concept study found that over half of adults presenting to hospital with AUFI remain undiagnosed by standard diagnostic testing. A large number and range of standard diagnostic tests were performed but only 5% of these contributed to the final diagnosis. The clinical features of the AUFI group were diverse and, with the exception of weight loss, which was significantly more common in undiagnosed participants, no clinical features could distinguish between diagnosed and undiagnosed participants. Infection was the predominant cause of AUFI, with twenty different pathogens identified (viral n=9, bacterial n=11) and three non-infection diagnoses made. Antimicrobials were prescribed in the vast majority, and broad-spectrum antimicrobial use was common. Of note, the length of hospital stay of was significantly higher in patients who remained undiagnosed than diagnosed participants (difference 1.2 days (IQR 0.04-1.66, *p*=0.036). The majority of participants had on-going symptoms at follow up and there was a trend towards prolonged symptoms in the undiagnosed when compared with the diagnosed group but this was not statistically significant. Conversely, all four adverse events occurred in diagnosed participants but were unrelated to study participation. To our knowledge, this is the first study of AUFI in UK adults and accepting of its limitations, provides insight into the clinical features, burden of diagnostic testing, aetiology, antimicrobial use and outcome of adults presenting to the UK with AUFI. The significantly longer length of stay and trend towards ongoing symptoms in the undiagnosed group is clinically plausible as diagnostic uncertainty may increase the time needed for a physician to investigate and confidently discharge a patient with AUFI and those with undiagnosed/untreated conditions may experience prolonged symptoms, however, due to the single site and small size of this proof of concept study, these findings must be interpreted with some caution and further explored by future studies before firm conclusions can be drawn.

The proportion of undiagnosed AUFI in this study is higher than that in the most other published studies, median 52% versus 36.0% (IQR 33.9-47.7%, range 3.2-62.0%) [1, 8, 11, 13, 17, 20-22, 24, 25, 29, 32, 34-37, 39-41, 44, 62-66, 68], however, there are no UK studies for comparison and there are some important differences in study design and sample population in the existing literature which may contribute to this difference.

Studies including diagnoses made on clinical grounds without microbiological confirmation may result in a lower proportion of undiagnosed cases being reported. D'Acremont *et al* report 3.2% (32/1005) of participants were undiagnosed but that 11.8% (119/1005) had no microbiological

cause identified in a cohort of febrile paediatric outpatients [65], equally, in a study of 270 participants in Northern India, Shelke *et al* report 10.7% (9/87) of diagnoses were attributable to 'viral', 'urinary tract infection (UTI)', 'lower respiratory tract infection (LRTI)', and 'gastroenteritis' but not all had an associated pathogen confirmed by standard diagnostic testing [37]. In a prospective study of AUFI in Tamil Nadu, Abrahamsen *et al* recorded a low proportion of undiagnosed participants 13.0% (13/100) but the 'diagnosed' cohort included five cases of TB meningitis and one culture negative infective endocarditis, despite no pathogen being identified [8]. In contrast, including clinical diagnoses did not always result in a low proportion of undiagnosed participants. Gur *et al* included clinical diagnoses in their three month study of febrile adults presenting to ED in Israel, and report 48%; (28/58) of participants remained undiagnosed [32].

Another key difference is study location, the majority of AUFI studies took place in tropical and sub-tropical regions in a mixture of rural and urban settings unlike this UK based study performed in a single city hospital. Agricultural work is common in these studies [12, 13, 21, 35, 66] whilst no participant in this UK based study listed agriculture as a main occupation. Agricultural work facilitates frequent, direct exposure to animals, wildlife and arthropod vectors such as mosquitos, ticks and mites, associated with specific infectious aetiologies. This difference in pathogen exposure is reflected in the high proportion of AUFI attributable to vector borne and zoonotic infections in the AUFI literature [8, 11, 13, 17, 20-22, 24, 34, 35, 37-40, 44, 62, 66, 235] in contrast to this UK based study where vector borne infections were only detected in those who had travelled outside the UK (see table 3.15).

There are no UK based AUFI studies in adults, as a result, the range of pathogens expected in UK AUFI is not well understood. Therefore, a predetermined set of investigations for the most probable pathogens could not be employed which may have increased diagnostic yield. Diagnostic testing in this study largely relied on a detailed history, examination and non-specific investigations, with a lower diagnostic sensitivity than specific tests. This may have contributed to a higher proportion of undiagnosed participants. AUFI is better studied in tropical and subtropical regions therefore more specific tests (such as blood films for malaria, and PCR for Dengue virus) can be employed based on known prevalence of these infections at the study site.

In contrast to this argument, several studies from regions known to be endemic for malaria and zoonotic infections had similar proportions of undiagnosed participants to this UK based study. A surveillance study of AUFI from 33 rural and urban hospitals in India performed by the Manipal Academy of Higher Education reported rural and urban regions found 52% (14232/27586) of all febrile participants aged between 5-65 years remained undiagnosed despite extensive

investigation [66]. Similarly, Chheng *et al* found 53.1% (758/1333) hospitalised children with febrile illness in Cambodia remained undiagnosed [44] and a study of AUFI in Queensland, Australia reported 56.8% (193/340) remained undiagnosed [51].

Access and cost of diagnostic testing, particularly where a test needs to be repeated (such as convalescent serology) may impact on the proportion of diagnoses made. Some AUFI studies were performed in either ED [32, 36] or outpatient/ community settings with limited or no follow up [24, 40, 44, 65], these studies had a higher median proportion of undiagnosed participants 41.3% (IQR 28.7-51.8%), when compared to the wider literature. Incomplete follow up was a recognised limitation by Susilawati *et al* who report a 3.3% attendance rate. Several studies cite lack of access to diagnostic testing as a limitation, this includes a lack of access to PCR for a broad range of respiratory pathogens [62], diagnostics limited by patient costs [22] and no access to stool microscopy, urine culture, TB culture or HIV testing [44]. The UK health system is fortunate to have access to a wide range of diagnostic testing and the majority of participants attended follow up in this study so these factors are unlikely to have contributed significantly to the undiagnosed cohort.

Participant age in AUFI studies may impact on the range of responsible pathogens identified, children may not have developed immunity or are yet to be vaccinated against some infections as you may reasonably expect in a UK cohort of adults to be. The Global Burden of Disease Study (described more fully in section 1.1.1) supports this concept, with twelve infectious diseases in the top twenty-five diseases for the 0-9 age group and only three infectious diseases in the top twenty-five for people over 25 years of age [10]. Children (less than 16 years) were the sole participants in several AUFI studies [44, 65, 68] and co-recruited with adults in a number of others [17, 20, 24, 29, 37, 39, 40, 63, 66]. Illustrating the impact of age in the aetiology of AUFI, D'Acremont *et al* detected 1232 pathogens in 1005 febrile Tanzanian children, of these 7.9% children had HHV-6 infection [65]. Primary infection with HHV-6 frequently occurs in childhood and can be associated with AUFI [236]. It then establishes latent infection but rarely causes disease in non-immunocompromised adults. Immunocompromised adults were excluded from this UK based study and so HHV-6 is unlikely to relevant to this population. Similarly, rotavirus and adenovirus cause asymptomatic or mild disease in immunocompetent adults [237] but accounted for 30% of gastroenteritis in this same paediatric study. Respiratory viruses showed a similar pattern, with over half of viruses responsible for the admitting illness being non-influenza and non-enterovirus. In contrast, only one clinically significant non-influenza, non-enterovirus respiratory virus was identified in this UK study of adults (Parainfluenza virus 3). Chheng *et al* report in 1225 admission episodes of Cambodian children with fever, 64.7% (44/68) of respiratory

viruses were non-influenza and 20.6% (14/68) were respiratory syncytial virus (RSV), an uncommon pathogen in hospitalised adults [44].

Antimicrobial usage in this study was high (82%; 82/100). Data for comparison in AUFI is scarce, but where documented, a smaller proportion of participants received antimicrobials (median 52.3%, IQR 50.2-61.1%) [13, 24, 38, 42, 43]. Study design may account for some of these differences, in this UK based study, any antimicrobial use was recorded including single doses. This is in contrast to Nhiem le Viet *et al* who recorded 49.7% (188/378) of participants received antimicrobials but only recorded treatment durations of 5-10 days [13]. If only participants receiving antimicrobials for 5 days or more were considered in this UK study, the proportion of participants on antimicrobials is more comparable to the published literature (54%, 54/100) (see figure 3.4). The median duration antimicrobials prescribed in this study was five days, but there is very limited data for comparison. Phuong *et al* examined AUFI in Vietnam community health centres, finding 77.2% (1618/2096) of participants were prescribed empirical antimicrobials, of which, only 54.4% (880/1618) were prescribed an appropriate duration and only 47.3% (765/1618) were prescribed an appropriate duration and dose. However, there was no data on included on what the prescribed durations were other than it was compared to a local standard [38].

The most used antimicrobials in this study were ciprofloxacin, co-amoxiclav, doxycycline, amoxicillin and ceftriaxone, which accounted for 69.4% (458.1/660) of all antimicrobials prescribed. Only three AUFI studies were identified which report on antimicrobial choice. Nhiem Le-Viet *et al* found 80.9% (152/188) received one antibiotic, and 19.2% (36/188) received two or three antibiotics. Most commonly used were amoxicillin, doxycycline, 2^{nd}/3^{rd} generation cephalosporins, fluoroquinolones, macrolides and antimalarials [13]. At three rural health centres in Laos, Phommasone *et al* found Penicillin A, cephalosporins, quinolones, macrolides, penicillin M&V and 'others' had been prescribed at the point of assessment [42]. Penicillin M&V is likely to equate to phenoxymethylpenicillin, however Penicillin A is not a commonly used term and is not clear from the text which agent this refers to which makes direct comparison to other studies difficult. Phuong *et al* report that the two most frequently prescribed antimicrobials were amoxicillin and cephalexin in Vietnam and they also noted 5.4% (82/1524) were prescribed combinations of antibiotics but details were not provided [38].

There is a clear lack of data in the published literature on antimicrobial usage, such that drawing comparisons between this UK study and the available data is unreliable. What can be concluded is that penicillins, cephalosporins and quinolones are frequently prescribed both in this study and

the wider literature and 'broad spectrum' antimicrobials represented three of the five most frequently prescribed agents (ciprofloxacin, co-amoxiclav and ceftriaxone) in this UK based study.

Length of hospital admission was significantly longer in undiagnosed participants compared to those with diagnosed AUFI, however, there is little available data for comparison. Interestingly, in a study of AUFI in North Queensland Susiliwati *et al* found significantly longer stays in the diagnosed cohort compared to the undiagnosed cohort (1 day versus 0 days p=0.001) [51] in a study of 340 participants with AUFI. Radhi *et al* found a significantly increased length of stay in febrile compared with afebrile children however, this was a paediatric study and not limited to those with AUFI [68].

This UK study found that 68% (50/73) of participants who attended follow up had symptoms 4-6 weeks following their illness onset and ongoing symptoms were more common in those without a clinically credible diagnosis (59.6% undiagnosed, 39.6% diagnosed, p= 0.07). However, no published studies of AUFI were identified that reported the proportion of participants who have symptoms at follow up. One large paediatric study concluded that diagnostic uncertainty may be responsible for the variation in hospital admission in febrile children and this was particularly notable in fever without focus (FWF) [238]. A second study found less experienced doctors used more resources when diagnosing and treating paediatric patients with fever [239]. It is clinically plausible that diagnostic uncertainty may lead to more investigations, more empirical treatment and potentially longer hospital admissions and that an undiagnosed, untreated illness may result in prolonged symptoms. However, there are no studies that have explored these issues in adults with AUFI and more good quality data is needed to better understand this area.

Adverse events were less common in this study than in the published literature. One participant was admitted to general intensive care and no participants died compared to 8.7% (IQR 5.3-11.2) ICU [34, 44, 51, 64, 235] and 6.2% (IQR 1.3-10.8) died [8, 9, 13, 20, 34, 40, 43, 44, 51, 52, 64, 235] respectively in the published literature. The diverse range of AUFI study designs make drawing any firm conclusions on this difference difficult, however some key factors may contribute to the difference in adverse events. Firstly, all published studies reporting outcome data took place in malaria endemic countries which is a prominent cause of global mortality [10]. Secondly, eleven of twelve studies reporting outcome data were performed in LMIC where access to healthcare, diagnostic testing and appropriate treatment may be less readily available than the UK. For example, Punjabi *et al* found 57% of participants had running water and half had a shared outdoor toilet [9]. Financial constraints were cited in two studies and healthcare is rarely free at the point of access [9, 235]. Other studies recruited participants over multiple sites for example, Susiliwati *et al* received a number of participants from neighbouring islands which would have introduced

significant delay in participants gaining healthcare support and so may have impacted on their outcome.

In addition to the clinical study findings, this study was the first identified to compare the Molysis Complete 5 method and the Qiagen Blood and Cell Maxi Kit plus NebNext Microbiome DNA extraction, commercial kits. Both methods were found to be equally effective at extracting *E. coli* DNA from EDTA whole blood samples. The Molysis method performed better at gDNA depletion with statistically significantly higher gDNA *Ct* values (compatible with lower gDNA PCR targets in the sample). The Molysis methodology was cheaper, more straightforward and could be performed in one continuous workflow allowing for multi-sample processing. Based on this data the Molysis method would be a more appropriate choice for preparation of EDTA whole blood samples for mNGS for pathogen detection.

A collaboration opportunity arose during the study meaning that the trialled extraction methods were not used on the clinical study samples prior to mNGS, but a patented technique developed at the collaborating site. All mNGS data generated was returned to the primary research site for analysis.

With the exception of one participant in whom the organisms identified by mNGS were clinically plausible pathogens (but not microbiologically confirmed). Metagenomic NGS did not confirm any diagnosis made by standard of care diagnostics in participants with AUFI and added no clinically credible diagnosis to those who remained undiagnosed. The proportion of total reads mapped to the human genome remained high following extraction and depletion steps (>90% for EDTA whole blood and approximately one third for serum) and reads associated with contaminating organisms represented a high proportion of non-human reads. Serum samples had lower human DNA levels but a higher proportion of contamination. There were no significant differences in the proportion of human and non-human reads between those who presented with AUFI and healthy volunteers.

Overall, the clinical diagnostic utility of mNGS for the diagnosis of adults presenting to hospital with AUFI cannot be confirmed or refuted by this small UK based study. There are some key limitations, which are likely to have contributed to this finding; these are described in the next section. However, some findings were consistent with the published literature. These findings include; that AUFI often remains undiagnosed and a wide range of pathogens can be causative, that many diagnostic tests are employed to diagnose AUFI but the diagnostic yield of standard testing is low. In the mNGS of clinical samples performed in this study, environmental/reagent contamination and high burdens of human DNA are highlighted as key barriers to overcome if mNGS is to move from the research laboratory to the routine clinical diagnostic laboratory.

If this body of work was to be repeated now, key improvements could be made to better assess mNGS as a clinical diagnostic tool. Firstly, analysing fresh samples in real-time, close to the patients would minimise RNA/DNA degradation and the impact of freeze-thaw and results could be provided in closer to real-time. Secondly, utilising newer host DNA depletion techniques such Saponin [199] and adaptive sequencing [240] could help increase the proportions of non-human genomes sequenced making mNGS a faster and more cost-effective option. Thirdly, this study has demonstrated RNA viruses are responsible for causing a high proportion of AUFI, 57.8% (26/45). Any future mNGS for AUFI should involve analysis of both RNA and DNA. Lastly, where a diagnosis is made by standard of care diagnostics, WGS of these organisms and comparison to the mNGS analysis could clarify whether mNGS is picking up the infecting organism and distinguish them from reads mapped to that same species which are present for another reason such as contamination. Until the diagnostic utility of mNGS for AUFI in adults is better understood, the focus should be on performing large, multicentre, well designed studies with pre-determined standard diagnostic tests and clinical follow up in well phenotyped cohorts of participants.

6.2 Limitations

The main limitation of this clinical study and experimental work is the small sample size of one hundred AUFI participants, fifty healthy volunteers with thirty participant samples analysed by mNGS. Additionally, the clinical study was performed at a single site on the acute medical unit excluding patients presenting directly to medical specialties (such as cardiology and neurology). Pregnant women and immunocompromised people were excluded, as were children and those who could not give consent and so extrapolation of study findings to these populations may be unreliable. People who are willing to take part in research may be more amenable than the general population to be extensively investigated and attend follow up, which may limit the generalisability of some study findings. However, better participant engagement is likely to have underestimated rather than overestimated the proportion of patients with undiagnosed AUFI and so the true figure may actually be higher in a non-research setting. Furthermore, length of hospital stay, antimicrobial use and participant outcome should not be affected by this bias.

Another important consideration is the pre-COVID-19 era in which this study took place. At this time it was not routine for all adults admitted to hospital to have a respiratory viral screen which is now commonplace. Unlike current times, at the time this study took place, global travel was frequent, resulting in a reasonably high number of travel related infections (7%; 7/100). Additionally, the presence of SARS-CoV-2 itself or the restrictions imposed during the pandemic may alter the epidemiology of other infectious diseases as we saw with low levels of Influenza in the winter of 2020 and are seeing again so far this year (2021) [242]. Additionally, SARS-CoV-2 may interact with circulating infectious diseases in a way we do not yet understand changing the epidemiology of infection.

Accepting of the limitations, the strengths of this proof of concept study are that it is the first of its kind to investigate adults with AUFI hospitalised in the UK, documenting detailed clinical data and diagnostic data contributing to the understanding of outcome, aetiology and antimicrobial use in AUFI in the UK.

As well as small sample sizes, there are some specific limitations of the experimental methodology to acknowledge. Within the clinical study, EDTA whole blood and serum samples were taken in standard tubes used for routine hospital analysis. The EDTA whole blood samples were kept on ice until aliquoted and frozen, however no other steps were taken to prevent the degradation of genomic material which may have impacted on the poor yield of mNGS. Equally, standardised blood collection tubes are not optimised for mNGS analysis so may carry a risk of contamination.

With regards to the extraction and enrichment work, when using live bacterial cultures there is inherent variability. It is difficult to precisely control for the quantity of bacteria in each sample; therefore, it is unlikely the exact number of organisms were present in each dilution step inoculated due to on-going bacterial replication. Equally, identical numbers of *E. coli* organisms are unlikely to be present in the aliquot used for analysis due to variability in the portion of sample selected. At lower concentrations of *E. coli* it becomes less likely that a portion of the sample containing *E. coli* will be sampled for analysis introducing more variability to the results at lower concentrations of *E. coli*. This variability was demonstrated across the Molysis results demonstrated by the higher standard deviation, most marked in the lower concentrations of *E. coli.* It is conceivable that at the higher concentrations the bacteria will be more evenly distributed throughout the blood sample allowing more reproducible results.

The Molysis method uses enzymatic degradation of DNA, the efficiency of enzymes can be affected by the specific environment and temperature at which they are working. If there is variability for example in temperature across the samples, this may impact upon the optimal working of the enzymes producing inconsistent results. Variability may also be accounted for by inconsistency in the manual extraction steps such as inaccurate pipetting, reagents which are not stored properly, or out of date or genuine variability in the extraction method itself. Two replicates were performed in these experiments, increasing the number of replicates and repeating these experiments may improve the consistency of the results. Another key limitation of this set of experiments was the use of frozen donor EDTA whole blood. Differential lysis relies on first lysing the human cells and degrading the host DNA before lysing the bacterial cells and extracting their DNA. Prior freezing of healthy volunteer blood samples can cause host and pathogen cells to lyse releasing endogenous enzymes and intracellular DNA. The host enzymes may break down pathogenic DNA of interest impacting on the amount of DNA which can be subsequently extracted for analysis. Multiple freeze-thaw events can also affect protein structure and may impact the affinity of enzymes in the extraction kit binding to target proteins.

Further limitations include the focus on a single organism (*E. coli*) and lack of inclusion of Gram-positive bacteria, intracellular bacteria, fungi or viruses, if this work was to be repeated, analysis of *S. aureus* as a significant Gram positive organism, *C. albicans*, a common fungal infection and Enterovirus, a common RNA viral infection should be included to broaden the implication of this work.

When interpreting the mNGS data there are a number of limitations, which should be considered. Firstly only a small selection of samples were analysed from 25 participants with AUFI and 5 healthy volunteers, rather than the whole cohort of participants and healthy volunteers which

may mean the diagnostic yield from mNGS is falsely low. Secondly, there was a delay from sample collection to mNGS analysis, although samples were frozen within one hour, they were later transferred to a collaborating laboratory. Although all measures were taken to protect against samples thawing, it is well understood that freeze-thawing events can impact on the integrity of sample DNA and RNA, which may in turn reduce the abundance of pathogen nucleic acid in the sample. Thirdly, no nucleic acid stabilisers were added to the sample prior to freezing. Endogenous DNAases and RNAases present in the sample may have broken down pathogen genomic material if the samples were not frozen promptly. Fourthly mNGS of EDTA whole blood has not been as widely studied as serum and plasma and so there are fewer 'tried and tested' methods, which can be adapted from the literature, which may have impacted on the methodology used. Fifthly, this study did not include RNA analysis despite 57.8% (26/45) of all infection diagnoses being with RNA viruses and 24% (6/25) of samples send for mNGS being from patients with confirmed RNA virus infections. This is a significant limitation, which impacts on the interpretation of these results in pathogens with a RNA genome. Lastly, the types of infections participants had may have reduced the potential yield of mNGS as three participants had fever associated with an infectious gastroenteritis (*C. difficile, S. enteritidis, C. jejuni*) organisms largely isolated to the gastrointestinal tract.

Due to the limitations noted above further work is needed to understand the aetiology, burden of antimicrobial use and outcome of AUFI across the UK to determine whether there is a true difference in the gDNA depletion between the methods examined here and whether mNGS really does have a clinically diagnostic value in AUFI. However, any further work must take into account new developments in the field since this piece of work was performed.

6.3 Implications of study

This study provides the first comprehensive review of hospitalised adults presenting with AUFI in the UK and provides important information regarding clinical characteristics, antibiotic use, aetiology and outcome in this under-studied condition. It demonstrates that patients with AUFI frequently remain undiagnosed despite extensive investigations and a wide range of diagnoses including viral, bacterial and non-infective aetiologies in those who do achieve a diagnosis. Despite the predominance of viral aetiology in those with a diagnosis, empirical treatment with broad-spectrum antibiotics remains very common. In addition, those who remain undiagnosed may have worse outcomes including longer length of hospital stay and less rapid return to premorbid levels of health. The findings highlight the urgent unmet need for improved diagnostics in patients with AUFI.

Prolonged hospital admissions impact on hospital capacity and resources, particularly as patients with AUFI are often housed in a side room (at greater cost) with additional infection control precautions in place. Multiple diagnostic tests add to the burden on healthcare resources and require additional expertise to perform and interpret the results. Patients who have persistent unexplained symptoms are likely to require more follow up appointments than those whose symptoms resolve and may be referred to additional specialties if the diagnosis remains unclear.

Undiagnosed AUFI may contribute to inappropriate antimicrobial use, in turn contributing to AMR as it can be difficult to confidently stop or de-escalate antimicrobials in a patient with a persisting fever without clear diagnosis. In addition, less easy to measure is the potential for unrecognised transmission of an infectious disease. At the time of writing, the SARS-CoV-2 pandemic continues to close borders, push healthcare systems to breaking point and threaten any sustained return to a more normal way of life. In common with most pandemic infections, the first few human cases of what would be named COVID-19 would undoubtedly have presented to healthcare with a febrile illness. If a nasopharyngeal swab for respiratory virus PCR was taken, this test would have been negative as the PCR primers required to detect SARS-CoV-2 were yet to be developed. Those first patients may have been cohorted in a bay with others or sent home to continue with their normal daily life, facilitating person-to-person transmission. Without access to rapid, untargeted diagnostic testing for infectious disease patients may not get the care they need, unnecessary pressure is put on healthcare systems and AMR and we may lose a vital opportunity to recognise the first case of a pathogen with the potential to cause the next pandemic.

This clinical study although small, presents some new insights in the field of AUFI in the UK. The extraction and enrichment experiments completed in this study also contribute to the body of literature on sample preparation prior to mNGS analysis. This study was the first identified to

compare the two chosen commercial kits and the first to use EDTA whole blood spiked with *E. coli* for this purpose. The findings were largely compatible with the literature. Both methods have been used in isolation or compared to other extraction and enrichment methodologies on a range of organisms including *E. coli* [210, 212] *K. pneumonia* [212], *S. aureus* [209, 213], *S. pneumoniae* [209], *S. agalactiae* [209]*, H. influenza* [209], *N. meningitidis* [209], *B. pertussis* [209], HSV [209], Adenovirus [209], Influenza A [209], *P. aeruginosa* [211], *C. albicans* [211], *S. epidermidis* [208], *E. faecalis* [208]. Since this work was completed there have been significant scientific developments that may minimise its impact on the wider field particularly, in the depletion of human DNA where Saponin based depletion has been demonstrated to remove over 99% of host DNA [199]. Advances in sequencing, for example Nanopore's adaptive sampling sequencing method can also optimise sequencing of the microbial composition of a sample by selectively rejecting host reads from the nanopore before they are fully sequenced [240].

The key implications of this study are that improved, untargeted diagnostic tests are needed to better diagnose patients with AUFI in the UK and beyond. Further data is needed to characterise AUFI in the UK and research needed to explore and develop mNGS as a potentially useful diagnostic test in for AUFI and undiagnosed infection. Focus on the pre-analytical steps of human DNA depletion and reducing method contaminants may bring mNGS closer to the high standards required for a reliable clinical diagnostic test, however robust clinical studies will be needed to confirm the diagnostic utility of any new mNGS diagnostic pathway prior to widespread implementation.

6.4 Recommendations for future work

This proof of concept study provides insight into the aetiology, antimicrobial usage and outcome of adults hospitalised AUFI in in the UK. It explores some of the pre-analytical and analytical barriers of mNGS as a potential diagnostic tool in AUFI and highlights some of the potential limitations of this approach. Further work is needed to better understand AUFI presenting in the UK and examine whether mNGS can be incorporated into a robust diagnostic pathway for infection.

One large non-UK study is the multisite FIEBRE study. This study aims to characterise and protocolise the presentation of fever in sub-Saharan Africa and Southeast Asia where AUFI associated mortality rates are high. Recruitment was completed on December 14th 2021 with some early data presented this year [243] and full results to follow in 2022 [244]. This study is the largest of its kind and promises to provide invaluable data to guide the diagnosis and management of fever in the sites studied. There may be lessons from this impressive study which will be applicable to UK based AUFI including appropriate sample types, diagnostic yield of diagnostics tests and reliable data on imported infections to the UK from the countries studied. However, early data from the FIEBRE study suggests much of the mortality is attributable to complications of uncontrolled HIV, findings which are not comparable to the UK setting [243]. For comparison, in 2020 there was a total of 614 HIV related deaths in the UK [245] compared with 12,000 in Malawi, one of the countries enrolled in the FIEBRE study [246]. In addition, all countries enrolled in FIEBRE are malaria endemic which again is not applicable to the UK setting.

Targeting research in areas where mortality is high is a clear and appropriate priority, particularly where there are highly effective interventions to improve outcomes such as improving access to HIV care. However, understanding AUFI in the UK is still of great importance as although mortality was low in this study, outcome from AUFI across the UK is not understood. Equally, the threat of AMR and need to detect infections with pandemic potential affects the entire population and developing diagnostic strategies effective in UK AUFI may be applicable non UK settings. Large, multi-centre studies are needed to confirm whether the findings of this UK study are reproducible in more diverse cohorts of participants with AUFI across the UK. More data is needed on the aetiology of fever throughout the UK as spatial heterogeneity of infectious disease has been demonstrated within the same country in other studies and we have no data to understand whether this is the case in the UK. For example, it is very plausible that the aetiology of fever on a remote Hebridean island is very different from that in a densely populated city such as London. The inclusion of pregnant women, immunocompromised people and children should be considered but would warrant a suitable power calculation to inform an appropriate sample

size to account for the variability within such a sample. Matched asymptomatic testing of healthy controls should be considered to understand the prevalence of asymptomatic infection in healthy volunteers as this may influence how much significance is attributed to a certain pathogen when detected in a febrile patient.

This study has illustrated that appropriate sample collection, preparation and storage is essential to optimise mNGS diagnostic output. If this work was to be repeated, consideration should be taken to ensure the whole diagnostic pathway could be completed as close to the patient as possible, in a clinically relevant timeframe and that steps are taken to minimise contamination, deplete human DNA and stabilise non-human RNA and DNA in the sample prior to storage or analysis. Successful studies have used a number of different bioinformatics pipelines and pooled the output, maximising diagnostic yield and this should be considered for any future work as it allows for a broader assessment of the mNGS output to include common and novel pathogens. Wilson *et al* employed a weekly clinical microbial sequence board (CMSB) to critically evaluate mNGS output in the clinical context [185]. This approach needs to be incorporated into future real-time mNGS clinical studies for clinical governance and quality control, particularly to ensure contamination is not misinterpreted as significant and low abundance pathogens are not discounted without proper consideration of the clinical context. As in this study, ethical consideration to the use or removal of human genomic material and the potential to discover life changing infections such as HIV must be clearly set out in the study design and maintained in the storage, analysis and publication of study findings. Pairing pathogen detection with analysis of the host response to infection [247] has been studied but the host response to infectious diseases is yet to be fully understood and so further work is needed to explore this. Additionally, appropriate ethical approval must be in place to analyse host genomic material.

In 2017 the Center for Next-Gen Precision Diagnostics, California became the first site to offer a validated mNGS diagnostic pathway for CSF and plasma, this approach has a turnaround time of 1-2 weeks so may not directly influence patient management. Cell free DNA mNGS analysis has shown promise in septic ICU patients with a turnaround time of 30 hours [234] and this technology has been harnessed by Karius who have developed a commercial mNGS pathway analysing cell free DNA from blood samples or frozen plasma, however the full results of their sepsis trial SEP-SEQ are still awaited [248]. In the UK, a group led by Professor Judith Breuer at Great Ormond Street Hospital are offering mNGS as an experimental tool for undiagnosed infection on CSF and fresh tissue based on a recent literature review [249] and successful implementation in selected cases [241, 242] but clearly state the laboratory methods are yet to be UK Accreditation Service (UKAS) accredited.

No study to date has devised an mNGS diagnostic pathway which has consistently superior sensitivity and specificity to standard diagnostic testing across all samples, and there are some key limitations of mNGS which should be acknowledged. MNGS requires a broad range of expertise including clinicians, molecular scientists, bioinformaticians and research ethics advisors. Expensive equipment is needed and is specific to the type of work being performed. Secure and abundant data storage facilities are needed and access to a range of well curated, comprehensive, up to data reference sequence databases. The results of any diagnostic test must be available in a clinically useful timeframe and with mNGS results; a panel of experts may be needed to ensure appropriate interpretation of results, which may take time. Samples must be collected, stored and protected from degradation (or propagation) appropriately and analysed in a timely fashion to reflect the current clinical picture. Equally, the timing of sample collection within the infectious illness is crucial as pathogen loads can be low or present in different body fluids as different phases of the illness [179]. Due to the complications described here, mNGS at present cannot be provided at all diagnostic laboratories particularly those with in low resource settings. However, well-organised, centralised government run diagnostic laboratories such as the Manipal Centre for Virus research in India provide a useful model as to how state of the art diagnostics can be offered across a wide geographic area with results available in a clinically useful timeframe [66]. At present, it is unlikely that mNGS will completely replace standard diagnostics but there is clear evidence to support mNGS augmenting standard of care diagnostic testing for infectious disease, particularly in those where a diagnosis is not forthcoming. Since 2018, the NIH alone has invested over five million dollars in clinical diagnostic studies of the use of mNGS as a diagnostic tool for febrile illness and emerging infection [251], the results of which are still being seen [196, 223, 231]. However, AUFI in the UK remains underexplored and larger, multi-site studies are needed to address this, to date, to our knowledge there is no such study open to recruitment in the UK.

6.5 Conclusion

Many key questions remained unanswered regarding AUFI; comprehensive, well-designed studies are key to better understand the burden of AUFI such that appropriate diagnostic and management pathways can be developed and the clinical diagnostic value of mNGS should be evaluated in this same context. Since this study was performed, the STROBE-metagenomics guidance has been published to guide the design and reporting of clinical studies employing mNGS as a diagnostic tool and this guide should be used to inform future clinical studies.

Understanding the true extent of AUFI and improving diagnostics in the UK could allow antimicrobials to be used more appropriately, reducing the pressure on AMR and hospital resources. Manock highlighted the importance of diagnostics in AUFI and its impact on public health and antimicrobial use quoting: "Until simple, affordable tests become available to accurately determine the etiology of AUFI early in its course, a combination of epidemiologic surveillance, focused public health efforts and broad- spectrum empiric therapy will have to suffice."[17].

With science rapidly advancing in the fields of mNGS and genomic surveillance, catalysed in the last two years by the SARS-CoV-2 pandemic, now is the time to generate robust, well designed studies of AUFI, concentrating on those who remain undiagnosed and exploiting ground-breaking science such as mNGS to enhance the diagnostic capacity of clinical laboratories ensuring appropriate patient management, targeted use of antimicrobials and the diagnostic infrastructure needed to detect the next novel pathogen of pandemic potential.

Appendix A Consent form

A.1 Participant consent form: patients with febrile illness

Patient trial ID number:

The use of unbiased Next Generation Sequencing for pathogen detection in adults hospitalised with acute undifferentiated febrile illness and suspected infection: an observational pilot study (SePSI)

Chief Investigator: Dr Tristan Clark

1. Please put your INITIALS in each box and sign at the bottom to indicate your consent to the following: I confirm that I have read and understand the Participant Information Sheet, Version_____, dated _____ for the above study. I have spoken to_____ and had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation in the study is voluntary and that I am free to withdraw at any time, without giving any reason and without my care or legal rights being affected. ☐

3. I agree that my GP will be informed about my participation in the study.

4. I agree to have a blood sample, urine sample and a nose and throat swab taken to be tested by genetic testing for microorganisms. If I have a lumbar puncture or a biopsy as part of hospital treatment, I agree to having some of the sample sent for research too. ☐

5. I understand that the results of these tests will not change my care or treatment during this illness. ☐

6. I give permission for relevant sections of any of my medical records and research data collected during the study to be looked at later by responsible individuals from regulatory authorities, the University of Southampton and University Hospital Southampton NHS Foundation Trust for the purposes of data analysis, audit and monitoring. ☐

7. I understand that I will be asked for a follow up visit and for repeat samples of blood, urine and nose and throat sample to be taken between 2-6 weeks later ☐

8. I agree to take part in this study. ☐

RESEARCH PARTICIPANT SIGNATURE:.....

NAME [138]: ..... DATE:.....

RESEARCH STAFF: SIGNATURE:.....

NAME:.....

DATE:.....ROLE (PLEASE CIRCLE): DOCTOR / NURSE

A.2 Participant consent form: healthy controls

Patient trial ID number:

The use of unbiased Next Generation Sequencing for pathogen detection in adults hospitalised with acute undifferentiated febrile illness and suspected infection: an observational pilot study (SePSI)

Chief Investigator: Dr Tristan Clark

Please put your INITIALS in each box and sign at the bottom to indicate your consent to the following:

1. I confirm that I have read and understand the Participant Information Sheet, Version 3.0, dated 22/12/2015 for the above study. I have spoken to_____ and had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation in the study is voluntary and that I am free to withdraw at any time, without giving any reason and without my care or legal rights being affected. ☐

3. I agree that my GP will be informed about my participation in the study. ☐

4. I agree to have a blood sample, urine sample and a nose and throat swab taken to be tested by genetic testing for microorganisms. ☐

5. I give permission for relevant sections of any of my medical records and research data collected during the study to be looked at later by responsible individuals from regulatory authorities, the University of Southampton and University Hospital Southampton NHS Foundation Trust for the purposes of data analysis, audit and monitoring. ☐

6. I understand that next generation sequencing has the ability to detect blood-borne viruses including HIV, Hepatitis B and Hepatitis C and I will be informed in the unlikely event that one of these infections is detected and referred to the appropriate specialty with my consent. ☐

7. I agree to take part in this study ☐

RESEARCH PARTICIPANT SIGNATURE:.....

NAME [138]: ..... DATE:.....

RESEARCH STAFF: SIGNATURE:.....

NAME:..... DATE:.....

ROLE (PLEASE CIRCLE): DOCTOR / NURSE

Appendix B Patient information leaflet

The use of unbiased Next Generation Sequencing for pathogen detection in adults hospitalised with acute undifferentiated febrile illness and suspected infection: an observational pilot study (SePSI).

Chief Investigator: Dr Tristan Clark BM MRCP DTM&H MD, Associate Professor and honorary consultant in Infectious Diseases.

Research Study

You are being invited to take part in the research study named above. Before you decide if you want to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part. Thank you for reading this information sheet.

What is the purpose of this study?

Acute fevers are a common reason for admission to hospital and it can be very difficult to diagnose the cause, which is usually an infection. Standard laboratory tests are often unable to find the cause of the infection, which may mean that an inappropriate treatment is used such as antibiotics being given when they are not needed (for example where a viral infection is actually the cause of the fever). Recent developments have shown that a technique called next generation sequencing performed on samples can sometimes diagnose infection when standard tests have not been able to. This study will look at the usefulness of this new method of testing by comparing the results to those from routine laboratory tests, taken at the same time. This study will also collect and test samples from a healthy volunteer group, which will then be compared to the results from group who have a fever.

Why have I been asked?

The symptoms and clinical signs that you display are consistent with an 'acute febrile illness' and so we would like to take samples from you and use next generation sequencing to look for a wide range of viruses, bacteria and other organisms which might have caused your illness. This is in addition to the standard tests you will have as part of your routine clinical care. We are therefore

asking for volunteers over the age of 18 who present to hospital with an acute fever illness. In addition we need well, healthy volunteers (over the age of 18) to act as controls and so we are also asking fit and well adults to take part and provide samples.

Do I have to take part?

No. It is completely up to you to decide whether or not to take part in this study. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part, you are free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive.

What will happen to me if I take part?

After you have finishing reading this, you will have the opportunity to discuss the study in more detail with a member of the research team. If you are happy to take part, then you will be asked to read and sign a consent form. All women, who have the potential to be pregnant (ie; are pre-menopausal), will be offered a urine pregnancy test as pregnancy could potentially affect the results of the study tests. Results of pregnancy tests done as part of this study will be given to you as soon as they are available by a member of the research team.

<u>Febrile Hospital Inpatients</u>

For those in hospital with acute febrile illness, three samples will be taken in addition to the standard tests that will be done by your medical team. The samples will be; a blood sample (the volume of this will be around 20mls – roughly 1 tablespoon full) this will include a PAX RNA gene tube which will allow us to better preserve any genetic material from organisms present in your blood, a urine sample and a swab from your nose and throat (which is like a 'cotton bud' briefly put up your nose and in the back of your throat). If you have had diarrhoea we will also ask you to provide a stool sample. If the medical team looking after you think that you need to have a lumbar puncture (a test where fluid is taken from the spine) or a tissue biopsy, we would also like to test these samples using the research test (if there is enough sample left over). All samples will be frozen and analysed at a later date. In addition information on your background, medical history, travel, symptoms and routine test results will be recorded by the research team.

You will then be asked to return around 4 weeks later to a clinic to see how you are and to take further samples, blood, urine and a swab from your nose and throat (there is no need to repeat the PAX RNA gene tube this time), this will most likely take place when you attend for follow up of the condition that led you to become unwell, if we ask you to return purely for the purpose of the

study, reasonable travel expenses can be provided. Again, information on your symptoms, the results of routine tests and the details of your hospital stay will be recorded at this time.

Healthy volunteers

For the healthy volunteers, blood (approximately 20mls), for 25 participants (50%) this will include a PAX RNA gene tube which will allow us to better preserve any genetic material from any organisms present in your blood, urine and nose and throat samples will be taken as above and information on your background and medical history will be collected. A telephone follow up will take place approximately 2 weeks after you enrol in the study to ask whether you have developed any signs of infection as this will affect our interpretation of your results. There is no need for collection of a second set of samples. All samples will be frozen and analysed at a later date looking for evidence of viruses, bacteria and other organisms.

What about confidentiality?

We take participant confidentiality very seriously. Only a very limited amount of personal identifiable information is requested from you, and when we come to look at and publish any results then information is presented anonymously – i.e. your details and personal information are never made available. Your GP will be informed that you are participating in this study.

What are the risks?

The risks of having a simple swab taken from your nose and throat in this study are minimal. Having the swab taken could be mildly uncomfortable for some people, but it is over very quickly. The blood sample may also be uncomfortable and occasionally cause mild bruising but the volume of blood taken will not have a detrimental effect on you. Next generation sequencing has the ability to detect blood borne viruses including HIV, Hepatitis B and Hepatitis C. If you are a hospital patient you will be routinely tested for these infections as part of your standard care and the results given to you by a member of the clinical team.

In the unlikely event that next generation sequencing detects unknown HIV, Hepatitis B or Hepatitis C infection in a healthy volunteer, the lead investigator on this study, Dr Tristan Clark is a fully trained and certified infectious diseases consultant who is experienced in management all infections including HIV, Hepatitis B and Hepatitis C and will inform you of the results, advise on next steps for management and refer you on to an appropriate medical professional with your consent. A confirmatory blood test may be required as part of routine care. The medical teams who look after patients with Hepatitis B or C are liver specialists (hepatologists), and patients with

HIV are looked after by infectious diseases and genito-urinary medicine specialists. All specialties are based in the local area.

What happens when the research study stops?

This study will end when we have tested 100 patients in hospital with fever and 50 healthy volunteers. We will use the data generated from this small 'pilot' study to plan a larger one.

What will happen to the results of the research study?

We intend to publish the results of our research in medical journals and to present the results at scientific meetings. The information from these journals will be available on the Internet. All results are anonymous in these publications and presentations. Any leftover samples collected during the study will be stored anonymously and may be used by our research group for future research studies, the same confidentiality rules will apply for any future research. We would like any useful results to lead to other studies looking at this new technology with the eventual aim that it will allow medical professionals to improve the way we manage and treat patients.

Who is organising and funding this research?

The NHS and the University of Southampton are funding this research.

Who has approved this study?

An ethics committee has reviewed the design of this study and approved it to go ahead. The local NHS Research and Development department has approved this study too.

Who can I talk to further?

The research team are very happy to answer your questions and discuss things further with you. You are welcome to talk to your doctors and nurses, family and friends, should you wish, about participating. Should you have any specific concerns you are welcome to discuss these with a research doctor, or the chief investigator, or Patient Support Services, about how you might take your concerns further.

Dr Tristan Clark, Chief investigator

Tel: 02381218410. T.W.Clark@soton.ac.uk

Dr Patrick Lillie / Dr Rebecca Houghton, Co-investigators

Tel: 02380777222. Patrick.Lillie@uhs.nhs.uk Rebecca.Houghton@uhs.nhs.uk

Patient Support Services, University Hospital Southampton NHS Foundation Trust

Tel: 023 8120 6325. patientsupportservices@uhs.nhs.uk.

Appendix C Case report form febrile patients

Participant Trial Number:	

SePSI Trial Case Report Form: Febrile Patient **Version 2.1**

FORM A

Enrolment

You cannot proceed without a 'Yes' to both questions	**Yes**	**No**
Has given fully informed written consent		
Agreed to blood borne virus testing		

Date of birth	/ /

Details of hospitalisation and recruitment		
Date of Hospitalisation (dd/mm/yyyy)	/	/
Time of hospitalisation (24hour clock):	:	
Admitted via ED, AMU or other (specify)		
Hospital ward location at time of recruitment		
Isolated in Side room (please circle one)	**Yes**	**No**
Date Recruited: (dd/mm/yyyy)	/	/
Time Recruited (24hour clock):	:	
Duration of Illness prior to hospitalisation(days)		

Demographics		
Sex: (please circle)	**Female**	**Male**
Age: (years)		
Occupation:		

Ethnicity	**(tick one box only)**
White British	
White Other (please specify)	
Indian	
Pakistani	
Bangladeshi	
Afro-Caribbean	
Black African	
Chinese	
Other- please specify	

Co-morbidities	**Tick/list all that apply**
Diabetes mellitus	
Chronic cardiovascular disease	
Chronic respiratory disease	
Chronic kidney disease	
Cerebrovascular disease	
Malignancy (please specify)	
Current Smoker	
Chronic neurological disease	
Other	

Travel Details		
Recent Travel: Includes travel abroad &UK within three months prior to admission (circle one)	**Yes**	**No**
Interval from travel to onset of illness (days, weeks, months)		

Details of Recent Travel (within three months of admission)			
Region Of Travel	**Specify Countr(ies)**	**Reason for Travel** **B:** Business **T:** Tourism **F:** Visiting friends / relatives **V:** volunteer/ aid work **R:** Research/education **O:** other	**Duration of travel / days**
Europe			
United States of America, Canada			
South and central America			
North Africa / Middle east			
Sub-Saharan Africa			
Southeast Asia			
South Central Asia			
Oceania (excluding Australasia)			
Australasia			
Caribbean			

Exposure (last 6 weeks or longer if felt to be relevant)		
Animal / insect	**Date if known**	**Details**
Insect bite		

Animal bite		
Other Animal/ bird contact		
Water contact		
Sea water		
Fresh water (river, stream, lake)		
Food / Drink / Other substances		
Unpasteurised dairy products		
Untreated water		
Raw meat/fish/bush meat		
Illicit substances		
Sexual contacts / unwell contacts		
Unwell contacts (please give details of symptoms & proximity of contact)		
New sexual contacts		
Healthcare contact		
Working in healthcare facility		
Sought healthcare attention		
Environmental/ Other exposures felt relevant (eg caves, mines,etc)		
1.		
2.		

Symptoms/Signs (Tick all that apply, give duration and any relevant details)			
Symptom	**Tick all that apply**	**Duration** (days)	**Specify details if required**
Constitutional			
Fever			
Lethargy			
Loss of appetite			
Night sweats / chills			
Lymphadenopathy			
Weight loss (specify how much)			
Respiratory			
Cough (specify details if productive)			
Shortness of breath			
Pleuritic chest pain			
Wheeze			
Coryza (rhinorrhoea, congestion, sneezing etc)			
Sore throat			

Gastro-intestinal			
Diarrhoea (three or more loose or liquid stools per day)			watery bloody please circle
Vomiting			
Abdominal pain			
Jaundice			
Hepatomegaly			
Splenomegaly			
Cardiovascular			
Chest pain			
Palpitations			
Neurology / opthalmology			
Headache			
Photophobia			
Neck stiffness			
Focal Neurology			
Seizures			
Conjunctivitis			
Musculoskeletal			
Arthralgia (specify which joints)			
Myalgia (specify sites)			
Genitourinary			
Dysuria /frequency			
Genital discharge /ulceration (specify)			
Rashes (please specify sites)			
Vesicular			
Macular			
Petechial			
Other			
Other signs / symptoms			
1.			
2.			
3.			

Working / Differential Diagnosis (list all documented, most senior doctor if possible)
1.
2.
3.
4.

Admission Observations	
Pulse	Beats per minute
Blood pressure	mmHg
Temperature	Degrees Celsius
Respiratory Rate	Breaths/minute
Saturations	%

Inspired oxygen (please circle unit measured)	FiO2 % OR L/Min

Antimicrobials (include antimalarials & those prescribed in the community for this illness)			
Name	**Dose**	**Route**	**Start date(dd/mm/yyyy) and time(00:00)**
1.			
2.			
3.			
4.			
5.			

Drug History (regular medications taken for chronic pre-existing co-morbidities)
1.
2.
3.
4.
5.
6.
7.
8.
9.
10.

Samples taken as part of standard clinical care (document N/A if not indicated)		
Sample type (number if more than one)	**Specify type, site or tests requested**	**Date**
Blood cultures		/ /
Respiratory sample –virology (sputum, BAL etc)		/ /
Respiratory sample –bacterial (sputum, BAL etc)		/ /
Mycobacterial cultures (eg. respiratory, blood, biopsy)		/ /
Nasopharyngeal swab- viral		/ /
Pharyngeal swab- bacterial		/ /
Whole blood (specify tests requested)		/ /
Urine MC&S		/ /
Cerebrospinal fluid (specify tests requested)		/ /
Biopsy material (specify tests requested)		/ /
1.		/ /
2.		/ /

Serology test	**Date**	**Serology test**	**Date**	**Serology test**	**Date**
HIV		EBV IgG		Parovirus IgM	
HBSAg		EBV IgM		Parvovirus IgG	
HBCore Ab		CMV IgG		Mycoplasma	
HCV Ab		CMV IgM		Toxoplasma	

Syphilis		ASOT		HAV	
HEV IgG		HEV IgM		Lyme antibody	
Measles IgG		Mumps		leptospirosis	
Other.....		Other.....		Other.....	
Other.....		Other.....		Other.....	

Study Samples Taken			
Sample type	**Number/volume taken (mls)**	**Date**	**Time (24 hour clock)**
Nasopharyngeal swab		/ /	:
Serum (10mls)		/ /	:
Whole blood (10mls)		/ /	:
PAXgene RNA (2.5mls)		/ /	:
Urine		/ /	:
Cerebrospinal fluid		/ /	:
Biopsy material		/ /	:
Stool sample		/ /	:

Study samples taken by				
Name (please print)	**Date**	**Time**	**Signature**	**Role** (please circle)
	/ /	:		Doctor
				Nurse
CRF Completed by				
Name (please print)	**Date**	**Time**	**Signature**	**Role** (please circle)
	/ /	:		Doctor
				Nurse

Participant Trial Number:	

SePSI Trial Case Report Form: Febrile Patient **Version 2.1**

FORM B

Follow up Review

Participant Trial ID:	
Date of Birth:	/ /
Date of Review:	/ /

Repeat Study Samples Taken		
Sample type	**Date**	**Time (24 hour clock)**
Nasopharyngeal swab	/ /	:
Serum (10mls)	/ /	:
Whole blood (10mls)	/ /	:
Urine	/ /	:

Repeat samples taken as part of routine care(All infection related investigations taken between enrolment and & last follow up should be included), document N/A if not indicated		
Sample type (number if more than one)	**Specify type, site or tests requested**	**Date**
Blood cultures		/ /
Respiratory sample –virology (sputum, BAL etc)		/ /
Respiratory sample –bacterial (sputum, BAL etc)		/ /
Mycobacterial cultures (eg.respiratory, blood, biopsy)		/ /
Nasopharyngeal swab- viral		/ /
Pharyngeal swab- bacterial		/ /
Serum (includes HIV, HBV, HCV)		/ /
Whole blood (specify tests requested)		/ /
Urine MC&S		/ /
Stool sample		/ /

Antimicrobials (during hospitalisation and on discharge)						
Antimicrobial	**Dose**	**Route**	**Time first dose**	**Date first dose**	**Time last dose**	**Date last dose**
1.						
2.						
3.						
4.						
5.						
6.						

Ongoing symptoms give duration and any relevant details)	**Duration (days, weeks)**
Constitutional	
Respiratory	
Gastro-intestinal	

Cardiovascular	
Neurology /ophthalmology	
Musculoskeletal	
Genitourinary	
Rashes (please specify sites)	
Other signs / symptoms	

Repeat samples taken by				
Name (please print)	**Date**	**Time**	**Signature**	**Role (please circle)**
	/ /	:		Doctor
				Nurse

Follow up review completed by				
Name (please print)	**Date**	**Time**	**Signature**	**Role (please circle)**
	/ /	:		Doctor
				Nurse

Participant Trial Number:	

SePSI Trial Case Report Form: Febrile Patient **Version 2.1**

Retrospective Analysis

FORM C

Participant Trial ID:	
Date of Birth:	

Time discharged (00:00)	:
Date discharged (DD/MM/YY)	/ /
Length of Stay, hours	

Isolated in Side room at any time	**Yes**	**No**
Date isolated	/	/
Date removed from isolated	/	/
Reason for isolation		

Blood Investigations (admission result and most abnormal)					
Full Blood Count	**Admission**	**Most abnormal**	**Liver function**	**Admission**	**Most abnormal**
White blood cells			ALT		
Lymphocytes			Bilirubin		
Neutrophils			ALP		
Eosinophils			AST		
Haemoglobin			Albumin		
Platelets			**Other**		
Electrolytes			CRP		
Creatinine			Lactate		
Urea			1.		
Sodium			2.		
Potassium			3.		

Other Investigations	**Y/N**	**Number**	**Date of 1st**	**Results**			
Chest X Ray							
ECG							
Lumbar puncture				**Opening pressure**		**Lymph. (%)**	
				WCC		**Glucose**	
				Poly. (%)		**Protein**	
				PCR & culture			
Imaging							

Other				

Antimicrobial	Dose	Route	Time 1st dose	Date 1st dose	Time last dose	Date last dose	Total duration, hrs
1.							
2.							
3.							
4.							
5.							
6.							
7.							
8.							
9.							

Results from samples taken as part of routine care (document N/A if not taken)

Sample type (number if more than one)	**Specify type, site or tests requested**	**Inpatient result**	**Follow up results**	**Total number of investigations for this illness**
Blood cultures				
Respiratory sample – virology (sputum, BAL etc)				
Respiratory sample – bacterial (sputum, BAL etc)				
Mycobacterial cultures (eg. respiratory, blood, biopsy)				
Nasopharyngeal swab- viral				
Pharyngeal swab- bacterial				
Whole blood (specify tests requested)				
Urine MC&S				
Cerebral spinal (specify tests requested)				
Biopsy material (specify tests requested)				
Stool sample				
1.				
2.				

Serology test	Result	Serology test	Result	Serology test	Result
HIV		EBV IgG		Parovirus IgM	
HBSAg		EBV IgM		Parvovirus IgG	
HBCore Ab		CMV IgG		Mycoplasma	
HCV Ab		CMV IgM		Toxoplasma	
Syphilis		ASOT		HAV	
HEV IgG		HEV IgM		Lyme antibody	
Measles IgG		Mumps		leptospirosis	
Other.....		Other.....		Other.....	
Other.....		Other.....		Other.....	

Result of Unbiased Next Generation Sequencing			
Sample type	**Recruitment Sample**		
	Date taken	**Date NGS performed**	**Result**
Nasopharyngeal swab			
Serum (10mls)			
Whole blood (10mls)			
Urine			
Cerebral spinal fluid			
Biopsy material			
Stool sample			
Sample type	**Follow up Sample**		
Nasopharyngeal swab			
Serum (10mls)			
Whole blood (10mls)			
Urine			
Cerebral spinal fluid			
Biopsy material			
Stool sample			

Participant Trial Number:

SePSI Trial Case Report Form: Febrile Patient **Version 2.1**

Diagnosis made by samples taken as part of standard care (if yes please detail below)	**Yes**	**No**

Diagnosis made by samples processed with unbiased next generation sequencing (if yes please detail below)	**Yes**	**No**

Investigator Final / 'Actionable' Diagnosis and basis for decision

Safety Data	**Yes**	**No**
Admitted to General intensive Care (GICU)		
Admitted to Respiratory High Care		
Died during current admission		
Discharged for end of life care		
Evidence of prolonged Hospital Stay		
Readmission to Hospital within 30 days		
New persistent or significant disability or incapacity		
Evidence of congenital anomaly or birth defect		
Died during follow up period		
If 'yes' to any of the above has a SAE form been completed?		

Retrospective data completed by				
Name (please print)	**Date**	**Time**	**Signature**	**Role (please circle)**
	/ /	:		Study doctor

Appendix D Laboratory protocol

SePSI laborarory Protocol Version 1.0 **26th October 2015**

Contents:

1) Viral nose and Throat swabs (green swabs)
2) Serum (10mls red top tube)
3) Whole blood (10mls purple top tube)
4) Urine (sterile container up to 15mls)
5) Other specimens (eg CSF, stool, histology specimens)

1) Viral nasopharyngeal swabs

a) A single sigma virocult (green) swab should be used to obtain first pharyngeal then nasal sampling
b) The swab should be broken to allow the tip to remain in the viral transport medium
c) Ensure appropriate labelling in indellable ink
   - Patient's Trial number Starting SePSI_____,
   - Date
   - Freeze to at least -70°C in a temperature monitored freezer

2) Serum (10mls red top)

a) Collect 10mls of blood in red top serum bottle via peripheral venopuncture
b) Invert tube a minimum of 5 times to ensure proper mixing with clotting activators- ensure sample taken
c) Allow to clot for a minimum of 60mins and maximum of 4 hours ideally in the upright position
d) Centrifuge at ≤ 1300g for 10minutes at 25°C room temperature
e) Pre-label microtubules with the following information:
   - Patient's Trial number Starting SePSI_____,
   - Date
   - 'SERUM'
   - Number tubules 1 of 3, 2 of 3 etc.
f) Following aseptic techniques transfer serum into the microtubules using a sterile graduated pipette; each tubule should contain a maximum of 2mls serum
g) Freeze to at least -70°C in a temperature monitored freezer

3) Whole Blood (10mls EDTA, purple top)

a) Collect 10mls EDTA blood via peripheral venopuncture
b) Invert the sample 5 times to ensure proper mixing with EDTA to prevent clotting
c) Ensure sample transported to the lab ASAP and within 4 hours
d) Add 5mls of PBS (pH 7.4) to a 15mL conical tube then add 5mL of EDTA blood sample to this, gently mix by capping and inverting the tube several times.
e) Freeze the 5mL remaining whole blood at -80C and continue to process the 5mLs of blood in PBS
f) Density gradient separation: carefully and very slowly add the blood/PBS mix to a 15mL conicle tube pre-loaded with 3mL of an appropriate endotoxin free density gradient medium (Histopaque, Ficoll, Lymphoprep etc) avoiding mixing.

g) Carefully load the conical tubes into the centrifuge buckets without disturbing the layers and centrifuge at 657g for 20mins at room temperature with no centrifuge brake.

h) After centrifuging, carefully remove the tubes from the buckets. The layers should now be visable. Layers from top to bottom of the tube will be plasma (yellow), PBMCs (white disc or halo), Ficol (clear), red blood cells (red).

i) Pre-label 2X 2.0mL microtubules
- Patient's Trial number Starting SePSI_____,
- Date
- 'PLASMA'
- Number tubules 1 , 2 etc.

j) Using graduated sterile transfer pipette and following aseptic technique inside a microbiology safety cabinet, remove plasma without disturbing the other layers, transfer 1-2mLs plasma into the the pre-labelled microtubules. The tubes should contact 1-2mLs plasma

4) Urine

a) Obtain mid stream urine from patient in study 15mLs sterile collection container

b) Ensure appropriate labelling in indellable ink
- Patient's Trial number Starting SePSI_____,
- Date
- 'URINE'

c) Freeze at -80°C

5) Other specimens

Very rarely an additional specimen not mentioned above such as Cerebral spinal fluid, stool or a histology specimen which is felt to be of importance the study will be collected. This will be pre-arranged with the CRF **PRIOR** to sample arrival. Processing of these samples will be minimal.

a) Ensure appropriate labelling

b) Ensure appropriate labelling in indellable ink
- Patient's Trial number Starting SePSI_____,
- Date
- 'SAMPLE TYPE'

d) Freeze at -80°C

Appendix E Case report form healthy volunteers

Participant Trial Number:	

SePSI Trial Case Report Form: Healthy volunteer **Version 2.1**

FORM D

Initial enrolment

You cannot proceed without a 'Yes' to all questions if applicable	Yes	No
Has given fully informed written consent		
Understands that blood borne viruses may be detected as a result of Next generation sequencing of blood samples but that this is not a validated test for HIV, Hepatitis B or Hepatitis C.		
If female, a negative pregnancy test		
Has been well with no symptoms of significant illness (including; fever, chills, sweats, myalgia, arthralgia, malaise, weight loss, cough, chest pain, rhinorrhoea, sore throat, abdominal pain, diarrhoea, dysuria, urinary frequency, haematuria, severe headache, collapse or seizure) in the past 14 days.		
Normally fit and well **WITHOUT** significant medical co-morbidity (including chronic cardiovascular, respiratory, renal, hepatic, or neurological illness, diabetes mellitus or immune compromised as defined by: ! HIV infection with a CD4 less than 200 cells/ml ! Any primary immunodeficiency ! Current or recent (within six months) chemotherapy or radiotherapy for malignancy ! Solid organ transplant recipients on immunosuppressive therapy ! Bone marrow transplant recipients currently receiving immunosuppressive treatment, or who received it within the last 12 months ! Patients with graft vs host disease ! Patients currently receiving high dose systemic corticosteroids (equivalent ≥ 40mg prednisolone per day for ≥ 3 weeks), and for as leas three months after treatment has stopped Patients currently or recently (within three months) on other types of immunosuppressive therapy.		

Participant Trial ID:	
Date of Birth:	
Date of recruitment (dd/mm/yyyy)	
Time recruited (00:00)	
Date follow up call required (7 days post enrolement)	

Demographics		
Sex: (please circle)	**Female**	**Male**
Age:		

1

Participant Trial Number:

SePSI Trial Case Report Form: Healthy volunteer **Version 2.1**

Occupation:	

Ethnicity	**(tick one box only)**
White British	
White Other (please specify)	
Indian	
Pakistani	
Bangladeshi	
Afro-Caribbean	
Black African	
Chinese	
Other- please specify	

Co-morbidities (Tick/List all that apply)
1.
2.
3.
4.

Drug History (regular medications taken for chronic pre-existing co-morbidities)			
Medication	**Dose**	**Route**	**Frequency**
1.			
2.			
3.			
4.			

Study Samples Taken			
Sample type	**Number/volume taken (mls)**	**Date**	**Time (24 hour clock)**
Nasopharyngeal swab		/ /	:
Serum (10mls)		/ /	:
Whole blood (10mls)		/ /	:
PAXgene tube (first 25 participants only)		/ /	:
Urine		/ /	:

Study samples taken by				
Name (please print)	**Date**	**Time**	**Signature**	**Role (please circle)**
	/ /	:		Doctor
				Nurse

CRF Completed by				
Name (please print)	**Date**	**Time**	**Signature**	**Role (please circle)**
	/ /	:		Doctor
				Nurse

Participant Trial Number:	

SePSI Trial Case Report Form: Healthy volunteer **Version 2.1**

FORM E

Follow up Contact 7days post recruitment

Participant Trial ID:	
Date of Birth:	/ /
Date of Review:	/ /

Illness & antimicrobials since recruitment- please tick	**Yes-** if yes please p details below	**No**
Has been well with no symptoms of significant illness (including; fever, chills, sweats, myalgia, arthralgia, malaise, weight loss, cough, chest pain, rhinorrhoea, sore throat, abdominal pain, diarrhoea, dysuria, urinary frequency, haematuria, severe headache, collapse or seizure) in the past 7 days.		
Has received antimicrobials since recruitment?		

Detail of any new symptoms since enrolment (give duration and any relevant details)	**Duration (days, weeks)**
Constitutional (fever, weight loss, chills, sweats,malaise)	
Respiratory (cough, chest pain, rhinorrhoea, sore throat)	
Gastro-intestinal (abdominal pain, diarrhoea, dysuria, urinary frequency, haematuria)	
Cardiovascular (chest pain, palpitations)	
Neurology /ophthalmology (severe headache, collapse or seizure)	
Musculoskeletal (myalgia, arthralgia)	
Genitourinary (dysuria, urinary frequency, haematuria)	

Rashes (please specify sites)	
Other signs / symptoms	

Details of any antimicrobials received since enrolment.						
Antimicrobial	**Dose**	**Route**	**Time first dose**	**Date first dose**	**Time last dose**	**Date last dose**
1.						
2.						
3.						
4.						
5.						
6.						

Follow up data completed by				
Name (please print)	**Date**	**Time**	**Signature**	**Role (please circle)**
	/ /	:		Nurse / Doctor

Participant Trial Number:

SePSI Trial Case Report Form: Healthy volunteer **Version 2.1**

Retrospective Analysis

FORM F

Participant Trial ID:	
Date of Birth:	

Result of Unbiased Next Generation Sequencing			
Sample type	**Recruitment Sample**		
	Date taken	**Date NGS performed**	**Result**
Nasopharyngeal swab			
Serum (10mls)			
Whole blood (10mls)			
Urine			
PAXgene (first 25 only)			

Results from samples processed with unbiased next generation sequencing

Retrospective data completed by				
Name (please print)	**Date**	**Time**	**Signature**	**Role (please circle)**
	/ /	:		Study doctor

List of References

1 Espinosa N, Cañas E, Bernabeu-Wittel M, Martín A, Viciana P, Pachón J. The changing etiology of fever of intermediate duration. Enfermedades Infecciosas y Microbiología Clínica. 2010;28(7):416-420.

2 WHO informal consultation on fever management in peripheral health care settings: A global review of evidence and practice [Internet]. WHO.int. 2022 [cited 13 January 2022]. Available from: https://www.who.int/publications/i/item/9789241506489

3 Crump J, Kirk M. Estimating the Burden of Febrile Illnesses. PLOS Neglected Tropical Diseases. 2015;9(12):e0004040.

4 Mattern S. Galen and his patients. The Lancet. 2011;378(9790):478-479.

5 Science Medicine and Animals- A theory of germs [Internet]. Washington DC: National Academies Press (USA); 2004 [cited 16 January 2022]. Available from: https://doi.org/10.17226/10733

6 Noakes T, Borresen J, Hew-Butler T, Lambert M, Jordaan E. Semmelweis and the aetiology of puerperal sepsis 160 years on: an historical review. Epidemiology and Infection. 2007;136(1):1-9.

7 Crump J. Time for a comprehensive approach to the syndrome of fever in the tropics. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2014;108(2):61-62.

8 Abrahamsen S, Haugen C, Rupali P, Mathai D, Langeland N, Eide G et al. Fever in the tropics: aetiology and case-fatality - a prospective observational study in a tertiary care hospital in South India. BMC Infectious Diseases. 2013;13(1):555.

9 Punjabi N, Taylor W, Murphy G, Purwaningsih S, Picarima H, Sisson J et al. Etiology of Acute, Non-Malaria, Febrile Illnesses in Jayapura, Northeastern Papua, Indonesia. The American Journal of Tropical Medicine and Hygiene. 2012;86(1):46-51.

10 Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019 Reference Life Table. Seattle, United States of America: Institute for Health Metrics and Evaluation (IHME), 2021. DOI https://doi.org/10.6069/1D4Y-YQ37

11 Susilawati T, McBride W. Undiagnosed undifferentiated fever in Far North Queensland, Australia: a retrospective study. International Journal of Infectious Diseases. 2014;27:59-64.

12 Kasper M, Blair P, Touch S, Sokhal B, Yasuda C, Williams M et al. Infectious Etiologies of Acute Febrile Illness among Patients Seeking Health Care in South-Central Cambodia. The American Journal of Tropical Medicine and Hygiene. 2012;86(2):246-253.

13 Le-Viet N, Le V, Chung H, Phan D, Phan Q, Cao T et al. Prospective case-control analysis of the aetiologies of acute undifferentiated fever in Vietnam. Emerging Microbes & Infections. 2019;8(1):339-352.

14 Global priority list of antibiotic-resistant bacteria to guide research, discover and development of new antibiotics. [Internet]. Who.int. 2022 [cited 15 January 2022]. Available from: https://www.who.int/medicines/publications/WHO-PPL-Short_Summary_25Feb-ET_NM_WHO.pdf

15 Morse S, Mazet J, Woolhouse M, Parrish C, Carroll D, Karesh W et al. Prediction and prevention of the next pandemic zoonosis. The Lancet. 2012;380(9857):1956-1965.

16 Susilawati T, Jex A, Cantacessi C, Pearson M, Navarro S, Susianto A et al. Deep sequencing approach for investigating infectious agents causing fever. European Journal of Clinical Microbiology & Infectious Diseases. 2016;35(7):1137-1149.

17 Manock S, de Bravo N, Smalligan R, Russell K, Espín J, Jacobsen K et al. Etiology of Acute Undifferentiated Febrile Illness in the Amazon Basin of Ecuador. The American Journal of Tropical Medicine and Hygiene. 2009;81(1):146-151.

18 Wangrangsimakul T, Althaus T, Mukaka M, Kantipong P, Wuthiekanun V, Chierakul W et al. Causes of acute undifferentiated fever and the utility of biomarkers in Chiangrai, northern Thailand. PLOS Neglected Tropical Diseases. 2018;12(5):e0006477.

19 Bhargava A, Ralph R, Chatterjee B, Bottieau E. Assessment and initial management of acute undifferentiated fever in tropical and subtropical regions. BMJ. 2018;:k4766.

20 Mattar S, Tique V, Miranda J, Montes E, Garzon D. Undifferentiated tropical febrile illness in Cordoba, Colombia: Not everything is dengue. Journal of Infection and Public Health. 2017;10(5):507-512.

21 Suttinont C, Losuwanaluk K, Niwatayakul K, Hoontrakul S, Intaranongpai W, Silpasakorn S et al. Causes of acute, undifferentiated, febrile illness in rural Thailand: results of a prospective observational study. Annals of Tropical Medicine & Parasitology. 2006;100(4):363-370.

22 Mittal G, Ahmad S, Agarwal R. K, Dhar M, Mittal M, Sharma S.Aetiologies of Acute Undifferentiated Febrile illness in Adult Patients – an Experience from a Tertiary Care Hospital in Northern India. Journal of clinical and diagnostic research. 2015;

23 Thompson C, Blacksell S, Paris D, Arjyal A, Karkey A, Dongol S et al. Undifferentiated Febrile Illness in Kathmandu, Nepal. The American Journal of Tropical Medicine and Hygiene. 2015;92(4):875-878.

24 Mueller T, Siv S, Khim N, Kim S, Fleischmann E, Ariey F et al. Acute Undifferentiated Febrile Illness in Rural Cambodia: A 3-Year Prospective Observational Study. PLoS ONE. 2014;9(4):e95868.

25 Bailey M, Trinick T, Dunbar J, Hatch R, Osborne J, Brooks T et al. Undifferentiated Febrile Illnesses Amongst British Troops in Helmand, Afghanistan. Journal of the Royal Army Medical Corps. 2011;157(2):150-155.

26 Mayxay M, Castonguay-Vanier J, Chansamouth V, Dubot-Pérès A, Paris D, Phetsouvanh R et al. Causes of non-malarial fever in Laos: a prospective study. The Lancet Global Health. 2013;1(1):e46-e

27 Bisoffi Z, Buonfrate D. When fever is not malaria. The Lancet Global Health. 2013;1(1):e11-e12.

28 Naing C, Kassim A. Scaling-up attention to nonmalaria acute undifferentiated fever. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2012;106(6):331-332.

29 Crump J, Morrissey A, Nicholson W, Massung R, Stoddard R, Galloway R et al. Etiology of Severe Non-malaria Febrile Illness in Northern Tanzania: A Prospective Cohort Study. PLoS Neglected Tropical Diseases. 2013;7(7):e2324.

30 Prasad N, Murdoch D, Reyburn H, Crump J. Etiology of Severe Febrile Illness in Low- and Middle-Income Countries: A Systematic Review. PLOS ONE. 2015;10(6):e0127962.

31 Geddes A., Unexplained Fever. British Medical Journal, 1974. 4: p. 397-398.

32 Gur H, Aviram R, Or J, Sidi Y. Unexplained fever in the ED: analysis of 139 patients. The American Journal of Emergency Medicine. 2003;21(3):230-235.

33 Petersdorf R, Beeson P. Fever of unexplained origen: report on 100 cases. Medicine. 1961;40(1):1-30.

34 Chrispal A, Boorugu H, Gopinath K, Chandy S, Prakash J, Thomas E et al. Acute undifferentiated febrile illness in adult hospitalized patients: the disease spectrum and diagnostic

predictors – an experience from a tertiary care hospital in South India. Tropical Doctor. 2010;40(4):230-234.

35 Kashinkunti MD, G.S., Dhananjaya M, acute undifferentiated febrile illness clinical spectrum and outcome from a tertiary care teaching hospital of north karnataka. Int J Biol Med Res., 2013. 4(2): p. 3399-4402.

36 Jung H, Chon S, Lee H, Oh W, Lee D. Etiologies of Acute Undifferentiated Fever and Clinical Prediction of Scrub Typhus in a Non-Tropical Endemic Area. The American Journal of Tropical Medicine and Hygiene. 2015;92(2):256-261.

37 Shelke Y, Deotale V, Maraskolhe D. Spectrum of Infections in Acute Febrile Illness in Central India. Indian Journal of Medical Microbiology. 2017;35(4):480-484.

38 Phuong H, de Vries P, Nagelkerke N, Giao P, Hung L, Binh T et al. Acute undifferentiated fever in Binh Thuan province, Vietnam: imprecise clinical diagnosis and irrational pharmaco-therapy. Tropical Medicine and International Health. 2006;11(6):869-879.

39 Mørch K, Manoharan A, Chandy S, Chacko N, Alvarez-Uria G, Patil S et al. Acute undifferentiated fever in India: a multicentre study of aetiology and diagnostic accuracy. BMC Infectious Diseases. 2017;17(1),

40 Leelarasamee A, C.C., Chenchittikul M, Udompanthurat S. , Etiologies of acute undifferentiated febrile illness in Thailand. J Med Assoc Thai, 2004. 87: p. 464–472.

41 Abhilash K, Jeevan J, Mitra S, Paul N, Murugan T, Rangaraj A et al. Acute undifferentiated febrile illness in patients presenting to a Tertiary Care Hospital in South India: clinical spectrum and outcome. Journal of Global Infectious Diseases. 2016;8(4):147.

42 Phommasone K, Althaus T, Souvanthong P, Phakhounthong K, Soyvienvong L, Malapheth P et al. Accuracy of commercially available c-reactive protein rapid tests in the context of undifferentiated fevers in rural Laos. BMC Infectious Diseases. 2015;16(1).

43 Thangarasu S, Natarajan P, Rajavelu P, Rajagopalan A, Seelinger Devey J. A protocol for the emergency department management of acute undifferentiated febrile illness in India. International Journal of Emergency Medicine. 2011;4(1).

44 Chheng K, Carter M, Emary K, Chanpheaktra N, Moore C, Stoesser N et al. A Prospective Study of the Causes of Febrile Illness Requiring Hospitalization in Children in Cambodia. PLoS ONE. 2013;8(4):e60634.

45 Emergency care systems for universal health coverage: ensuring timely care for the acutely ill and injured [Internet]. WHO.int. 2022 [cited 14 January 2022]. Available from: https://www.who.int/publications/i/item/emergency-care-systems-for-universal-health-coverage-ensuring-timely-care-for-the-acutely-ill-and-injured

46 Knott J. Febrile adults presenting to the emergency department: outcomes and markers of serious illness. Emergency Medicine Journal. 2004;21(2):170-174.

47 Goto M, Koyama H, Takahashi O, Fukui T. A Retrospective Review of 226 HospitalizedPatients with Fever. Internal Medicine. 2007;46(1):17-22.

48 Zhiyong Z, Bingjun L, Xiaoju L, Xinjuan F,Ping F,Wenya W. Fever of unknown origin: a report from China of 208 cases. International Journal of Clinical Practice. 2003 Sep;57(7):592-596. PMID: 14529060.

49 Cheung M, Temple L, Khan M. An unexpected cause of deranged liver function: pyogenic liver abscess. JRSM Short Reports. 2011;2(1):1-4.

50 Mazhar M, K.J., Bolger DT Jr. A 23-year-old Man with Leptospirosis and Acute Abdominal Pain. Hawaii J Med Public Health. 2016 Oct;75(10):291-294. PMID: 27738562; PMCID: PMC5056631.

51 Susilawati T. and McBride W., Undiagnosed undifferentiated fever in Far North Queensland, Australia: a retrospective study. International Journal of Infectious Diseases, 2014. 27: p. 59-64.

52 Singh R. A study of Etiological Pattern in an Epidemic of Acute Febrile Illness During Monsoon in a Tertiary Health Care Institute of Uttarakhand, India. JOURNAL OF CLINICAL AND DIAGNOSTIC RESEARCH. 2014;.

53 Rani R, Sundararajan T, Rajesh S, Jeyamurugan T. A Study on Common Etiologies of Acute Febrile Illness Detectable by Microbiological Tests in a Tertiary Care Hospital. International Journal of Current Microbiology and Applied Sciences. 2016;5(7):670-674.

54 Employment in agriculture (% of total employment) (modeled ILO estimate) | Data [Internet]. Data.worldbank.org. 2022 [cited 11 January 2022]. Available from: https://data.worldbank.org/indicator/SL.AGR.EMPL.ZS?end=2019&start=1991&view=map

55 Environment, agriculture and food | GOV.WALES [Internet]. GOV.WALES. 2022 [cited 14 January 2022]. Available from: https://gov.wales/preparing-wales-brexit/environment-agriculture-and-food

56 Martinón-Torres F, Salas A, Rivero-Calle I, Cebey-López M, Pardo-Seco J, Herberg JA, Boeddha NP, et al; EUCLIDS Consortium. Life-threatening infections in children in Europe (the EUCLIDS Project): a prospective cohort study. Lancet Child Adolesc Health. 2018 Jun;2(6):404-414. doi: 10.1016/S2352-4642(18)30113-5. Epub 2018 Apr 28. PMID: 30169282.

57 Hagedoorn NN, Borensztajn D, Nijman RG, Nieboer D, Herberg JA , Balode A, et al. Development and validation of a prediction model for invasive bacterial infections in febrile children at European Emergency Departments: MOFICHE, a prospective observational study; Arch Dis Child 2021;106:641–647.

58 Borensztajn D, Hagedoorn NN, Carrol E, Von Both U,Dewez JE, Emonts M, et al. Characteristics and management of adolescents attending the ED with fever: a prospective multicentre study. BMJ Open 2022;12:e053451. doi:10.1136/ bmjopen-2021-053451

59 Dabrera G, Said B, Kirkbride H, On behalf of the USII Collaborating Group C. Evaluation of the surveillance system for undiagnosed serious infectious illness (USII) in intensive care units, England, 2011 to 2013. Eurosurveillance. 2014;19(46)

60 Marí-Dell'Olmo M, Gotsens M, Palència L, Burström B, Corman D, Costa G, et al., Socioeconomic inequalities in cause-specific mortality in 15 European cities. Journal of Epidemiology and Community Health, 2015. 69(5): p. 432-441.

61 Premaratna R, Rajapakse R, Chandrasena T, Nanayakkara D, Bandara N, Kularatna S et al. Contribution of rickettsioses in Sri Lankan patients with fever who responded to empirical doxycycline treatment. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2010;104(5):368-370.

62 Ahmad S, Dhar M, Mittal G, Bhat N, Shirazi N, Kalra V et al. A comparative hospital-based observational study of mono- and co-infections of malaria, dengue virus and scrub typhus causing acute undifferentiated fever. European Journal of Clinical Microbiology & Infectious Diseases. 2016;35(4):705-711.

63 Pokharel S, Karki M, Acharya B, Marasini B, Arjyal A. Outbreak of acute undifferentiated febrile illness in Kathmandu, Nepal: clinical and epidemiological investigation. BMC Infectious Diseases. 2020;20(1).

64 Limper M, Gerstenbluth I, Duits A, Van Gorp E. One-year epidemiology of febrile diseases on the Emergency Department of a Caribbean island: the Curaçao-experience. International Journal of Infectious Diseases. 2010;14:e201-e202.

65 D'Acremont V, Kilowoko M, Kyungu E, Philipina S, Sangu W, Kahama-Maro J et al. Beyond Malaria — Causes of Fever in Outpatient Tanzanian Children. New England Journal of Medicine. 2014;370(9):809-817.

66 Hospital based surveillance of Acute Febrile Illness in India [Internet]. Manipal. 2022 [cited 14 January 2022]. Available from: https://manipal.edu/dvr/research/projects/AFIstudy.html

67 Murdoch D, Basnyat B, Keenan A, Dull P, Archibald L, Woods C et al. The eitiology of febrile illness in adults presenting to Patan hospital in Kathmandu, Nepal The American Journal of Tropical Medicine and Hygiene. 2004;70(6):670-675.

68 Radhi S, Patel S. Aetiology of Febrile Illnesses Presenting to a District Hospital. Pediatric Infectious Diseases: Open Access. 2017;02(02).

69 Garnacho-Montero J, Garcia-Garmendia J, Barrero-Almodovar A, Jimenez-Jimenez F, Perez-Paredes C, Ortiz-Leyba C. Impact of adequate empirical antibiotic therapy on the outcome of patients admitted to the intensive care unit with sepsis. Critical Care Medicine. 2003;31(12):2742-2751.

70 Nishiguchi S, Nishino K, Kitagawa I, Tokuda Y. Inappropriate use of antibiotics in primary care for patients with infective endocarditis. Journal of Infection and Chemotherapy. 2020;26(6):640-642.

71 Roberts M, Lever A. An analysis of imported infections over a 5-year period at a teaching hospital in the United Kingdom. Travel Medicine and Infectious Disease. 2003;1(4):227-230.

72 Ahmed F, Lindley M, Allred N, Weinbaum C, Grohskopf L. Effect of Influenza Vaccination of Healthcare Personnel on Morbidity and Mortality Among Patients: Systematic Review and Grading of Evidence. Clinical Infectious Diseases. 2013;58(1):50-57.

73 Macsween K, Higgins C, McAulay K, Williams H, Harrison N, Swerdlow A et al. Infectious Mononucleosis in University Students in the United Kingdom: Evaluation of the Clinical Features and Consequences of the Disease. Clinical Infectious Diseases. 2010;50(5):699-706.

74 Sepkowitz K. Tuberculosis as the cause of fever of unknown origin: A review. International Journal of Infectious Diseases. 1997;2(1):47-51.

75 Lekhbal A, Chaker K, Halily S, Abada R, Rouadi S, Roubal M et al. Treatment of cervical lymph node tuberculosis: When surgery should be performed? A retrospective cohort study. Annals of Medicine and Surgery. 2020;55:159-163.

76 Iqbal M, Subhan A, Aslam A, Frequency of Tuberculosis in Cervical Lymphadenopathy. Journal of Surgery Pakistan (International), 2010. 15(2).

77 Andrade-Castellanos C, Paz-Velarde B, Carreón-Bautista E, Pozos-Ochoa L. Fever, night sweats, weight loss and lymphadenopathy. More than meets the eye. Enfermedades Infecciosas y Microbiología Clínica. 2020;38(8):392-393.

78 Global action plan on antimicrobial resistance [Internet]. Who.int. 2022 [cited 14 January 2022]. Available from: https://www.who.int/publications/i/item/9789241509763

79 UK 5-year action plan for antimicrobial resistance 2019 to 2024 [Internet]. GOV.UK. 2022[cited 14 January 2022]. Available from: https://www.gov.uk/government/publications/uk-5-year-action-plan-for-antimicrobial-resistance-2019-to-2024

80 No time to Wait: Securing the future from drug-resistant infections [Internet]. Who.int. 2022[cited 14 January 2022]. Available from: https://www.who.int/publications/i/item/no-time-to-wait-securing-the-future-from-drug-resistant-infections

81 Shah S, Parry M, Salter S, Pangan A. A Therapeutic Challenge Due to Infection with an Unusual New Delhi Metallo-Beta-Lactamase-Producing Pseudomonas aeruginosa in Pregnancy. Clinical Microbiology Newsletter. 2020;:192-194.

82 Goff D, File T. The risk of prescribing antibiotics “just-in-case” there is infection. Seminars in Colon and Rectal Surgery. 2018;29(1):44-48.

83 Wang C, Jerng J, Cheng K, Lee L, Yu C, Hsueh P et al. Pandrug-resistant Pseudomonas aeruginosa among hospitalised patients: clinical features, risk-factors and outcomes. Clinical Microbiology and Infection. 2007;13(2):iii.

84 Zhang D, Micek S, Kollef M. Time to Appropriate Antibiotic Therapy Is an Independent Determinant of Postinfection ICU and Hospital Lengths of Stay in Patients With Sepsis*. Critical Care Medicine. 2015;43(10):2133-2140.

85 Bint A, Burtt I. Adverse Antibiotic Drug Interactions. Drugs. 1980;20(1):57-68.

86 Legendre D, Muzny C, Marshall G, Swiatlo E. Antibiotic Hypersensitivity Reactions and Approaches to Desensitization. Clinical Infectious Diseases. 2013;58(8):1140-1148.

87 Crew P, Rhodes N, O'Donnell J, Miglis C, Gilbert E, Zembower T et al. Correlation between hospital-level antibiotic consumption and incident health care facility-onset Clostridium difficile infection. American Journal of Infection Control. 2018;46(3):270-275.

88 Brown K, Valenta K., Fisman D, Simor A, Daneman N., Hospital Ward Antibiotic Prescribing and the Risks of Clostridium difficile Infection. . JAMA Intern Med. , 2015. 175(4): p. 626-633.

89 Dethlefsen L, Huse S, Sogin M, Relman D. The Pervasive Effects of an Antibiotic on the Human Gut Microbiota, as Revealed by Deep 16S rRNA Sequencing. PLoS Biology. 2008;6:e280.

90 Schwartz T, Kohnen W, Jansen B, Obst U. Detection of antibiotic-resistant bacteria and their resistance genes in wastewater, surface water, and drinking water biofilms. FEMS Microbiology Ecology. 2003;43(3):325-335.

91 Joshi R, Colford J, Kalantri S, Reingold A. Nonmalarial Acute Undifferentiated Fever in a Rural Hospital in Central India: Diagnostic Uncertainty and Overtreatment with Antimalarial Agents. The American Journal of Tropical Medicine and Hygiene. 2008;78(3):393-399.

92 Langford B, So M, Raybardhan S, Leung V, Soucy J, Westwood D et al. Antibiotic prescribing in patients with COVID-19: rapid review and meta-analysis. Clinical Microbiology and Infection. 2021;27(4):520-531.

93 Lansbury L, Lim B, Baskaran V, Lim W. Co-Infections in People with COVID-19: A Systematic Review and Meta-Analysis. SSRN Electronic Journal. 2020;

94 Murchison C. A Treatise on the Continued Fevers of Great Britain. Longmans, Green and Company;.

95 Murchison C. The prevalence of continued fevers in the metropolis and other parts of the United Kingdom during the year 1858. The Lancet. 1859;73(1857):337-338.

96 Madhav N, O.B., Gallivan M, et al. Pandemics: Risks, Impacts, and Mitigation. In: Jamison DT, Gelband H, Horton S, et al, Pandemics: Risks, Impacts, and Mitigation, in Disease Control Priorities: Improving Health and Reducing Poverty. 3rd edition. 2017, Washington (DC): The International Bank for Reconstruction and Development / The World Bank;. p. 315-347.

97 Wacharapluesadee S, Tan C, Maneeorn P, Duengkae P, Zhu F, Joyjinda Y et al. Evidence for SARS-CoV-2 related coronaviruses circulating in bats and pangolins in Southeast Asia. Nature Communications. 2021;12(1).

98 Taylor L, Latham S, woolhouse M. Risk factors for human disease emergence. Philosophical Transactions of the Royal Society of London Series B: Biological Sciences. 2001;356(1411):983-989.

99 European Centre for Disease Prevention and Control. HIV/AIDs surveillance in Europe 2019. 2019 p. 1-95

100 Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet. 2020;395(10223):497-506.

101 Petridou C, Simpson A, Charlett A, Lyall H, Dhesi Z, Aarons E. Zika virus infection in travellers returning to the United Kingdom during the period of the outbreak in the Americas (2016-17): A retrospective analysis. Travel Medicine and Infectious Disease. 2019;29:21-27.

102 Feldmann H, Geisbert T. Ebola haemorrhagic fever. The Lancet. 2011;377(9768):849-862.

103 Barry M, Traoré F, Sako F, Kpamy D, Bah E, Poncin M et al. Ebola outbreak in Conakry, Guinea: Epidemiological, clinical, and outcome features. Médecine et Maladies Infectieuses. 2014;44(11-12):491-494.

104 Xu B, Low S, Tan R, Vasanwala F. A case series of atypical presentation of Zika Virus infection in Singapore. BMC Infectious Diseases. 2016;16(1)

105 Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. New England Journal of Medicine. 2020;382(13):1199-1207.

106 Li Y, Peng S, Li L, Wang Q, Ping W, Zhang N et al. Clinical and Transmission Characteristics of Covid-19 — A Retrospective Study of 25 Cases from a Single Thoracic Surgery Department. Current Medical Science. 2020;40(2):295-300.

107 Nadeau S, Vaughan T, Scire J, Huisman J, Stadler T. The origin and early spread of SARS-CoV-2 in Europe. Proceedings of the National Academy of Sciences. 2021;118(9):e2012008118.

108 Lecuit M, Eloit M. The potential of whole genome NGS for infectious disease diagnosis. Expert Review of Molecular Diagnostics. 2015;15(12):1517-1519.

109 Overview | Sepsis: recognition, diagnosis and early management | Guidance | NICE [Internet]. Nice.org.uk. 2022 [cited 14 January 2022]. Available from: https://www.nice.org.uk/guidance/ng51

110 Chang E, Kao K, Tsai M, Yang C, Huang Y, Liu C et al. Occult Klebsiella pneumoniae bacteremia at emergency department: A single center experience. Journal of Microbiology, Immunology and Infection. 2015;48(6):684-691.

111 Wheeler H. Dark ground microscopy and treponemal serological tests in the diagnosis of early syphilis. Sexually Transmitted Infections. 2004;80(5):411-414.

112 Heymans R, van der Helm J, de Vries H, Fennema H, Coutinho R, Bruisten S. Clinical Value of Treponema pallidum Real-Time PCR for Diagnosis of Syphilis. Journal of Clinical Microbiology. 2010;48(2):497-502.

113 UK SMI B 26: investigation of fluids from normally sterile sites [Internet]. GOV.UK. 2022 [cited 14 January 2022]. Available from: https://www.gov.uk/government/publications/smi-b-26-investigation-of-fluids-from-normally-sterile-sites

114 UK SMI B 27: investigation of cerebrospinal fluid [Internet]. GOV.UK. 2022 [cited 14 January 2022]. Available from: https://www.gov.uk/government/publications/smi-b-27-investigation-of-cerebrospinal-fluid

115 UK SMI B 37: investigation of blood cultures (for organisms other than Mycobacterium species) [Internet]. GOV.UK. 2022 [cited 14 January 2022]. Available from: https://www.gov.uk/government/publications/smi-b-37-investigation-of-blood-cultures-for-organisms-other-than-mycobacterium-species

116 Goto M, Al-Hasan M. Overall burden of bloodstream infection and nosocomial bloodstream infection in North America and Europe. Clinical Microbiology and Infection. 2013;19(6):501-509.

117 Thigpen M, Whitney C, Messonnier N, Zell E, Lynfield R, Hadler J et al. Bacterial Meningitis in the United States, 1998–2007. New England Journal of Medicine. 2011;364(21):2016-2025.

118 Long, B., A. Koyfman, and M. Gottlieb, Evaluation and Management of Septic Arthritis and its Mimics in the Emergency Department. Western Journal of Emergency Medicine, 2019. 20(2): p. 331-341.

119 Dever J, Sheikh M. Review article: spontaneous bacterial peritonitis - bacteriology, diagnosis, treatment, risk factors and prevention. Alimentary Pharmacology & Therapeutics. 2015;41:1116-1131.

120 Köser C, Holden M, Ellington M, Cartwright E, Brown N, Ogilvy-Stuart A et al. Rapid Whole-Genome Sequencing for Investigation of a Neonatal MRSA Outbreak. New England Journal of Medicine. 2012;366:2267-2275.

121 Yagupsky P, Nolte F. Quantitative aspects of septicemia. Clinical Microbiology Reviews. 1990;3(3):269-279.

122 Sautter R, Bills A, Lang D, Ruschell G, Heiter B, Bourbeau P. Effects of Delayed-Entry Conditions on the Recovery and Detection of Microorganisms from BacT/ALERT and BACTEC Blood Culture Bottles. Journal of Clinical Microbiology. 2006;44(4):1245-1249.

123 Linsenmeyer K, Gupta K, Strymish J, Dhanani M, Brecher S, Breu A. Culture if spikes? Indications and yield of blood cultures in hospitalized medical patients. Journal of Hospital Medicine. 2016;11(5):336-340.

124 Jannes G, De Vos D. A Review of Current and Future Molecular Diagnostic Tests for Use in the Microbiology Laboratory, in Diagnostic Bacteriology Protocols, Methods Mol Bio 2006;345:1-21. doi: 10.1385/1-59745-143-6:1.

125 Jordon J, Durso M, Real-Time Polymerase Chain Reaction for Detecting Bacterial DNA Directly from Blood of Neonates Being Evaluated for Sepsis. Journal of Molecular Diagnostics, 2005. 7(5).

126 Amar C, East C, Gray J, Iturriza-Gomara M, Maclure E, McLauchlin J. Detection by PCR of eight groups of enteric pathogens in 4,627 faecal samples: re-examination of the English case-control Infectious Intestinal Disease Study (1993–1996). European Journal of Clinical Microbiology & Infectious Diseases. 2007;26(5):311-323.

127 Houghton R, Ellis J, Galiano M, Clark T, Wyllie S. Haemagglutinin and neuraminidase sequencing delineate nosocomial influenza outbreaks with accuracy equivalent to whole genome sequencing. Journal of Infection. 2017;74(4):377-384.

128 Neoh H, Tan X, Sapri H, Tan T. Pulsed-field gel electrophoresis (PFGE): A review of the “gold standard” for bacteria typing and current alternatives. Infection, Genetics and Evolution. 2019;74:103935.

129 Laakso S, Kirveskari J, Tissari P, Mäki M. Evaluation of High-Throughput PCR and Microarray-Based Assay in Conjunction with Automated DNA Extraction Instruments for Diagnosis of Sepsis. PLoS ONE. 2011;6:e26655.

130 Standards for Microbiology Investigations Investigation of hepatitis B infection [Internet]. 4th ed. 2021 [cited 16 January 2022]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1010975/V_4i6.1.pdf

131 UK Standards for Microbiology Investigations Screening for hepatitis C infection [Internet]. 2017 [cited 16 January 2022]. Available from:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/634739/V_5i7.pdf

132 UK Standards for Microbiology Investigations HIV screening and confirmation. 11th ed. 2017. [cited 16 January 2022]. Available from:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/581545/V_11i4.1.pdf

133 UK Standards for Microbiology Investigations Syphilis serology [Internet]. 44th ed. 2016 [cited 16 January 2022]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/539597/V_44i2.1.pdf

134 Eldin C, Parola P, Raoult D. Limitations of diagnostic tests for bacterial infections. Médecine et Maladies Infectieuses. 2019;49(2):98-101.

135 Rajput R, Denniston A, Murray P. False Negative Toxoplasma Serology in an Immunocompromised Patient with PCR Positive Ocular Toxoplasmosis. Ocular Immunology and Inflammation. 2017;26(8):1200-1202.

136 Pierce V, Elkan M, Leet M, McGowan K, Hodinka R. Comparison of the Idaho Technology FilmArray System to Real-Time PCR for Detection of Respiratory Pathogens in Children. Journal of Clinical Microbiology. 2011;50(2):364-371.

137 Buss S, Leber A, Chapin K, Fey P, Bankowski M, Jones M et al. Multicenter Evaluation of the BioFire FilmArray Gastrointestinal Panel for Etiologic Diagnosis of Infectious Gastroenteritis. Journal of Clinical Microbiology. 2015;53(3):915-925.

138 Leveque N, Van Haecke A, Renois F, Boutolleau D, Talmud D, Andreoletti L. Rapid Virological Diagnosis of Central Nervous System Infections by Use of a Multiplex Reverse Transcription-PCR DNA Microarray. Journal of Clinical Microbiology. 2011;4;:3874-3879.

139 Brendish N, Schiff H, Clark T. Point-of-care testing for respiratory viruses in adults: The current landscape and future potential. Journal of Infection. 2015;71(5):501-510.

140 Liverpool Covid-19 Community Testing Pilot [Internet]. 1st ed. Liverpool; 2020 [cited 16 January 2022]. Available from: https://www.liverpool.ac.uk/media/livacuk/research/Liverpool,Community,Testing,Pilot,Interim,Evaluation.pdf

141 Rivett L, Sridhar S, Sparkes D, Routledge M, Jones NK, Forrest et al. Screening of healthcare workers for SARS-CoV-2 highlights the role of asymptomatic carriage in COVID-19 transmission. Elife. 2020 May 11;9:e58728. doi: 10.7554/eLife.58728. PMID: 32392129; PMCID: PMC7314537.

142 Lee LYW, R.S., Pang M, Charlett A, Anderson C, Hughes GJ, Barnard M et al. SARS-CoV-2 infectivity by viral load, S gene variants and demographic factors and the utility of lateral flow devices to prevent transmission. Clin Infect Dis, 2021.

143 Peto T; UK COVID-19 Lateral Flow Oversight Team. COVID-19: Rapid antigen detection for SARS-CoV-2 by lateral flow assay: A national systematic evaluation of sensitivity and specificity for mass-testing. EClinicalMedicine. 2021 Jun;36:100924. doi: 10.1016/j.eclinm.2021.100924. Epub 2021 May 30. PMID: 34101770; PMCID: PMC8164528.

144 Moshe M, Daunt A, Flower B, Simmons B, Brown J, Frise R et al. SARS-CoV-2 lateral flow assays for possible use in national covid-19 seroprevalence surveys (React 2): diagnostic accuracy study. BMJ. 2021;:n423.

145 Koczula K, Gallotta A. Lateral flow assays. Essays in Biochemistry. 2016;60(1):111-120.

146 Parisi M, Soldini L, Di Perri G, Tiberi S, Lazzarin A, Lillo FB. Offer of rapid testing and alternative biological samples as practical tools to implement HIV screening programs. New Microbiol. 2009 Oct;32(4):391-6. PMID: 20128446.

147 Kim J, Cao X, Finkelstein J, Cárdenas W, Erickson D, Mehta S. A two-colour multiplexed lateral flow immunoassay system to differentially detect human malaria species on a single test line. Malaria Journal. 2019;18(1).

148 Pai M, Denkinger CM, Kik SV, Rangaka MX, Zwerling A, Oxlade O, et al. Gamma interferon release assays for detection of Mycobacterium tuberculosis infection. Clin Microbiol Rev. 2014 Jan;27(1):3-20. doi: 10.1128/CMR.00034-13. PMID: 24396134; PMCID: PMC3910908.

149 Andersen P, Munk ME, Pollock JM, Doherty TM. 2000. Specific immune-based diagnosis of tuberculosis. Lancet 356:1099–1104. http: //dx.doi.org/10.1016/S0140-6736(00)02742-2.

150 Schmidt K, Mwaigwisya S, Crossman L, Doumith M, Munroe D, Pires C et al. Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. Journal of Antimicrobial Chemotherapy. 2016;72(1):104-114.

151 Doan T, Wilson M, Crawford E, Chow E, Khan L, Knopp K et al. Illuminating uveitis: metagenomic deep sequencing identifies common and rare pathogens. Genome Medicine. 2016;8(1)

152 Sanger F, Coulson A. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. Journal of Molecular Biology. 1975;94(3):441-448.

153 Sanger F, Air G, Barrell B, Brown N, Coulson A, Fiddes J et al. Nucleotide sequence of bacteriophage ϕX174 DNA. Nature. 1977;265 (5596):687-695.

154 Sanger F, Coulson A , Hong G, Hill D, Petersen G, Nucleotide sequence of bacteriophage lambda DNA. Journal of Molecular Biology, 1982. 162(4): p. 729-773.

155 Chee M, Bankier A, Beck S, Bohni R, Brown C, Cerny R, et al. Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. Curr Top Microbiol Immunol. 1990;154:125-69. doi: 10.1007/978-3-642-74980-3_6. PMID: 2161319.

156 Goebel S, Perkus M, Davis S, Winslow J, Paoletti E., The complete DNA sequence of vaccinia virus. Virology., 1990. 179(1): p. 247-66

157 Fleischmann R, Adams M, White O, Clayton R, Kirkness E, Kerlavage A, et al. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science. 1995 Jul 28;269(5223):496-512. doi: 10.1126/science.7542800. PMID: 7542800.

158 Pettersson, E., J. Lundeberg, and A. Ahmadian, Generations of sequencing technologies. Genomics, 2009. 93(2): p. 105-11.

159 Quail M, Swerdlow H, Turner D. Improved Protocols for the Illumina Genome Analyzer Sequencing System. Current Protocols in Human Genetics. 2009;62(1).

160 Tyler A, Mataseje L, Urfano C, Schmidt L, Antonation K, Mulvey M et al. Evaluation of Oxford Nanopore's MinION Sequencing Device for Microbial Whole Genome Sequencing Applications. Scientific Reports. 2018;8(1).

161 NCBI Virus. 2021 [cited 2021 8th July]; Available from: https://www.ncbi.nlm.nih.gov/labs/virus/vssi/ -/virus?SeqType_s=Nucleotide&VirusLineage_ss=SARS-CoV-2,%20taxid:2697049.

162 Walker T, Kohl T, Omar S, Hedge J, Del Ojo Elias C, Bradley P et al. Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. The Lancet Infectious Diseases. 2015;15(10):1193-1202.

163 Gordon N, Price J, Cole K, Everitt R, Morgan M, Finney J et al. Prediction of Staphylococcus aureus Antimicrobial Resistance by Whole-Genome Sequencing. Journal of Clinical Microbiology. 2014;52(4):1182-1191.

164 Azarian T, Cook R, Johnson J, Guzman N, McCarter Y, Gomez N et al. Whole-Genome Sequencing for Outbreak Investigations of Methicillin-Resistant Staphylococcus aureus in the Neonatal Intensive Care Unit: Time for Routine Practice. Infection Control & Hospital Epidemiology. 2015;36(7):777-785.

165 Harris S, Cartwright E, Török M, Holden M, Brown N, Ogilvy-Stuart A et al. Whole-genome sequencing for analysis of an outbreak of meticillin-resistant Staphylococcus aureus: a descriptive study. The Lancet Infectious Diseases. 2013;13(2):130-136.

166 Gilchrist C, Turner S, Riley M, Petri W, Hewlett E. Whole-Genome Sequencing in Outbreak Analysis. Clinical Microbiology Reviews. 2015;28(3):541-563.

167 Kundu S, Lockwood J, Depledge D, Chaudhry Y, Aston A, Rao K et al. Next-Generation Whole Genome Sequencing Identifies the Direction of Norovirus Transmission in Linked Patients. Clinical Infectious Diseases. 2013;57(3):407-414.

168 Penedos A, Myers R, Hadef B, Aladin F, Brown K. Assessment of the Utility of Whole Genome Sequencing of Measles Virus in the Characterisation of Outbreaks. PLOS ONE. 2015;10:e0143081.

169 Walker T, Ip C, Harrell R, Evans J, Kapatai G, Dedicoat M et al. Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: a retrospective observational study. The Lancet Infectious Diseases. 2013;13(2):137-146.

170 Grard G, Fair J, Lee D, Slikas E, Steffen I, Muyembe J et al. Correction: A Novel Rhabdovirus Associated with Acute Hemorrhagic Fever in Central Africa. PLOS Pathogens. 2016;12(3):e1005503.

171 Wise E, Márquez S, Mellors J, Paz V, Atkinson B, Gutierrez B et al. Oropouche virus cases identified in Ecuador using an optimised qRT-PCR informed by metagenomic sequencing. PLOS Neglected Tropical Diseases. 2020;14(1):e0007897.

172 Fresia P, Antelo V, Salazar C, Giménez M, D'Alessandro B, Afshinnekoo E et al. Urban metagenomics uncover antibiotic resistance reservoirs in coastal beach and sewage waters. Microbiome. 2019;7(1).

173 Hendriksen R, Munk P, Njage P, van Bunnik B, McNally L, Lukjancenko O et al. Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. Nature Communications. 2019;10(1).

174 Quick J, Ashton P, Calus S, Chatt C, Gossain S, Hawker J et al. Rapid draft sequencing and real-time nanopore sequencing in a hospital outbreak of Salmonella. Genome Biology. 2015;16(1).

175 Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, et al. Real-time, portable genome sequencing for Ebola surveillance. Nature. 2016 Feb 11;530(7589):228-232. doi: 10.1038/nature16996. Epub 2016 Feb 3. PMID: 26840485; PMCID: PMC4817224.

176 Ocheretina O, Shen L, Escuyer V, Mabou M, Royal-Mardi G, Collins S et al. Whole Genome Sequencing Investigation of a Tuberculosis Outbreak in Port-au-Prince, Haiti Caused by a Strain with a "Low-Level" rpoB Mutation L511P – Insights into a Mechanism of Resistance Escalation. PLOS ONE. 2015;10(6):e0129207.

177 Price J, Golubchik T, Cole K, Wilson D, Crook D, Thwaites G et al. Whole-Genome Sequencing Shows That Patient-to-Patient Transmission Rarely Accounts for Acquisition of Staphylococcus aureus in an Intensive Care Unit. Clinical Infectious Diseases. 2013;58(5):609-618.

178 Jolley K, Hill D, Bratcher H, Harrison O, Feavers I, Parkhill J et al. Resolution of a Meningococcal Disease Outbreak from Whole-Genome Sequence Data with Rapid Web-Based Analysis Methods. Journal of Clinical Microbiology. 2012;50(9):3046-3053.

179 Faria N, Azevedo R, Kraemer M, Souza R, Cunha M, Hill S et al. Zika virus in the Americas: Early epidemiological and genetic findings. Science. 2016;352(6283):345-349.

180 Peto L, Rodger G, Carter D, Osman K, Yavuz M, Johnson K et al. Diagnosis of SARS-CoV-2 Infection with LamPORE, a High-Throughput Platform Combining Loop-Mediated Isothermal Amplification and Nanopore Sequencing. Journal of Clinical Microbiology. 2021;59(6).

181 Greninger A, Naccache S, Federman S, Yu G, Mbala P, Bres V et al. Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. Genome Medicine. 2015;7(1).

182 Dahmane A, van Griensven J, Van Herp M, Van den Bergh R, Nzomukunda Y, Prior J et al Constraints in the diagnosis and treatment of Lassa fever and the effect on mortality in hospitalized children and women with obstetric conditions in a rural district hospital in Sierra Leone. Trans R Soc Trop Med Hyg 2014; 108: 126-132

183 Andersen KG, Shapiro BJ, Matranga CB, Sealfon R, Lin AE, Moses LM et al Clinical Sequencing Uncovers Origins and Evolution of Lassa Virus. Cell. 2015 Aug 13;162(4):738-50. doi: 10.1016/j.cell.2015.07.020. PMID: 26276630; PMCID: PMC4537774.

184 Olschläger S, Lelke M, Emmerich P, et al. Improved detection of Lassa virus by reverse transcription-PCR targeting the 5' region of S RNA. Journal of Clinical Microbiology. 2010 Jun;48(6):2009-2013. DOI: 10.1128/jcm.02351-09. PMID: 20351210; PMCID: PMC2884523.

185 Kafetzopoulou LE, Pullan ST, Lemey P, et al. Metagenomic sequencing at the epicenter of the Nigeria 2018 Lassa fever outbreak. Science (New York, N.Y.). 2019 Jan;363(6422):74-77. DOI: 10.1126/science.aau9343. PMID: 30606844; PMCID: PMC6855379.

186 Duan H, Li X, Mei A, Li P, Liu Y, Li X et al. The diagnostic value of metagenomic next generation sequencing in infectious diseases. BMC Infectious Diseases. 2021;21(1).

187 Schildgen V, Mai S, Khalfaoui S, Lüsebrink J, Pieper M, Tillmann R et al. Pneumocystis jirovecii Can Be Productively Cultured in Differentiated CuFi-8 Airway Cells. mBio. 2014;5(3).

188 Munoz P, Bouza E, Cuenca-Estrella M, Eiros J, Perez M, Sanchez-Somolinos M et al. Saccharomyces cerevisiae Fungemia: An Emerging Infectious Disease. Clinical Infectious Diseases. 2005;40:1625-1634.

189 Parize P, Muth E, Richaud C, Gratigny M, Pilmis B, Lamamy A et al. Untargeted next-generation sequencing-based first-line diagnosis of infection in immunocompromised adults: a multicentre, blinded, prospective study. Clinical Microbiology and Infection. 2017;23(8):574.e1-574.e6.

190 Wilson MR, Naccache SN, Samayoa E, Biagtan M, Bashire H, Guixia Y et al. Actionable diagnosis of neuroleptospirosis by next generation sequencing. N Engl J Med 2014 370(25): 2408-2417. Doi:10.1056/NEJMoa1401268

191 Naccache S, Peggs K, Mattes F, Phadke R, Garson J, Grant P et al. Diagnosis of Neuroinvasive Astrovirus Infection in an Immunocompromised Adult With Encephalitis by Unbiased Next-Generation Sequencing. Clinical Infectious Diseases. 2015;60(6):919-923.

192 Wang C, Li A, Shi Q, Yu Z. Metagenomic next-generation sequencing clinches diagnosis of leishmaniasis. The Lancet. 2021;397(10280):1213.

193 Dedkov V, Lukashev A, Deviatkin A, Kuleshov K, Safonova M, Poleshchuk E et al. Retrospective diagnosis of two rabies cases in humans by high throughput sequencing. Journal of Clinical Virology. 2016;78:74-81.

194 Wu X, Yan G, Han S, Ye Y, Cheng X, Gong H et al. Diagnosing Balamuthia mandrillaris encephalitis via next-generation sequencing in a 13-year-old girl. Emerging Microbes & Infections. 2020;9(1):1379-1387.

195 Yozwiak N, Skewes-Cox P, Stenglein M, Balmaseda A, Harris E, DeRisi J. Virus Identification in Unknown Tropical Febrile Illness Cases Using Deep Sequencing. PLoS Neglected Tropical Diseases. 2012;6(2):e1485.

196 Gu W, Miller S, Chiu C. Clinical Metagenomic Next-Generation Sequencing for Pathogen Detection. Annual Review of Pathology: Mechanisms of Disease. 2019;14(1):319-338.

197 Laboratory-Based SARS-CoV-2 Viral Detection tests 2021 21st June 2021 [cited 2021 7th July]; Available from: https://www.gov.uk/government/publications/how-tests-and-testing-kits-for-coronavirus-covid-19-work/target-product-profile-laboratory-based-sars-cov-2-viral-detection-tests.

198 Govender K, Street T, Sanderson N, Eyre D. Metagenomic Sequencing as a Pathogen-Agnostic Clinical Diagnostic Tool for Infectious Diseases: a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies. Journal of Clinical Microbiology. 2021;59(9).

199 Charalampous T, Kay G, Richardson H, Aydin A, Baldan R, Jeanes C et al. Nanopore metagenomics enables rapid clinical diagnosis of bacterial lower respiratory infection. Nature Biotechnology. 2019;37(7):783-792.

200 Perlejewski K, Bukowska-Ośko I, Rydzanicz M, Pawełczyk A, Caraballo Cortès K, et al. Next-generation sequencing in the diagnosis of viral encephalitis: sensitivity and clinical limitations. Scientific Reports. 2020;10(1).

201 Bharucha T, Oeser C, Balloux F, Brown J, Carbo E, Charlett A et al. STROBE-metagenomics: a STROBE extension statement to guide the reporting of metagenomics studies. The Lancet Infectious Diseases. 2020;20(10):e251-e260.

202 Southampton University Hospital NHS Foundation Trust Website]. Available from: https://www.uhs.nhs.uk/AboutTheTrust/AboutTheTrust.aspx. 2021 [cited 2021 30th April 2021];

203 Associated British Ports. 2021 [cited 2021 30th April 2021]; Available from: https://www.abports.co.uk/locations/southampton/.

204 Green book, Chapter 6, contraindications and special considerations [cited 2022 10th November 2022]; from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/655225/Greenbook_chapter_6.pdf

205 Wain J, Diep TS , Ho VA, Walsh AM, Tuyet Hoa TN, Parry C, White NJ, Quantitation of Bacteria in Blood of Typhoid Fever Patients and Relationship between Counts and Clinical Features, Transmissibility, and Antibiotic Resistance. J Clin Microbiol. 1998 Jun; 36(6): 1683-1687

206 O'Grady J, Wain, J, Mwaigwisya S, Kay G. WO2018109454 – Method for nuclei acid depletion. 21.06.2018 [cited 2021 28.11.2021]; Available from: https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2018109454.

207 The Approved List of biological agentsAdvisory Committee on Dangerous Pathogens [Internet]. 4th ed. 2021 [cited 16 January 2022]. Available from: https://www.hse.gov.uk/pubns/misc208.pdf

208 Thoendel M, Jeraldo P, Greenwood-Quaintance K, Yao J, Chia N, Hanssen A et al. Comparison of microbial DNA enrichment tools for metagenomic whole genome sequencing. Journal of Microbiological Methods. 2016;127:141-145.

209 Hasan M, Rawat A, Tang P, Jithesh P, Thomas E, Tan R et al. Depletion of Human DNA in Spiked Clinical Specimens for Improvement of Sensitivity of Pathogen Detection by Next-Generation Sequencing. Journal of Clinical Microbiology. 2016;54(4):919-927.

210 McCann C, Jordan J. Evaluation of MolYsis™ Complete5 DNA extraction method for detecting Staphylococcus aureus DNA from whole blood in a sepsis model using PCR/pyrosequencing. Journal of Microbiological Methods. 2014;99:1-7.

211 Loonen A, Bos M, van Meerbergen B, Neerken S, Catsburg A, Dobbelaer I et al. Comparison of Pathogen DNA Isolation Methods from Large Volumes of Whole Blood to Improve Molecular Diagnosis of Bloodstream Infections. PLoS ONE. 2013;8(8):e72349.

212 Trung NT, Hien TTT, Huyen TTT, Quyen DT, Van Son T, Hoan PQ et al. Enrichment of bacterial DNA for the diagnosis of blood stream infections. BMC Infectious Diseases 2016 16:235, DOI 10.1186/s12879-016-1568-1

213 Fiore E, Van Tyne D, Gilmore M. Pathogenicity of Enterococci. Microbiology Spectrum. 2019;7(4).

214 Prakash S., Dental abscess: A microbiological review. Dental Research Journal, 2013. 10(5).

215 Wong Y, Cheung C, Ting W, Chak W. Intra-abdominal abscess with Prevotella species and Streptococcus anginosus co-infection in a peritoneal dialysis patient. Nephrology. 2019;25(3):273-274.

216 Lei W, Chang W, Shih S, Liu C, Shih C. Pyogenic Liver Abscess with Prevotella Species and Fusobacterium necrophorum as Causative Pathogens in an Immunocompetent Patient. Journal of the Formosan Medical Association. 2009;108(3):253-257.

217 Tong S, Davis J, Eichenberger E, Holland T, Fowler V. Staphylococcus aureus Infections: Epidemiology, Pathophysiology, Clinical Manifestations, and Management. Clinical Microbiology Reviews. 2015;28(3):603-661.

218 D'Argenio V, Casaburi G, Precone V, Salvatore F. Comparative metagenomic analysis of human gut microbiome composition using two different bioinformatic pipelines. BioMed research International. Vol 2014 DOI 10.1155/2014/325340

219 Glassing A, Dowd S, Galandiuk S, Davis B, Chiodini R. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. Gut Pathogens. 2016;8(1).

220 Sato M, Kuroda M, Kasai M, Matsui H, Fukuyama T, Katano H et al. Acute encephalopathy in an immunocompromised boy with astrovirus-MLB1 infection detected by next generation sequencing. Journal of Clinical Virology. 2016;78:66-70.

221 Wang Q, Li J, Ji J, Yang L, Chen L, Zhou R et al. A case of Naegleria fowleri related primary amoebic meningoencephalitis in China diagnosed by next-generation sequencing. BMC Infectious Diseases. 2018;18(1).

222 Albariño C, Foltzer M, Towner J, Rowe L, Campbell S, Jaramillo C et al. Novel Paramyxovirus Associated with Severe Acute Febrile Disease, South Sudan and Uganda, 2012. Emerging Infectious Diseases. 2014;20(2):211-216.

223 Wilson M, Sample H, Zorn K, Naccache S, Miller S, Chiu C. 2563. Clinical Metagenomic Next-Generation Sequencing for Diagnosis of Meningitis and Encephalitis. N Engl J Med 2019 Jun 13;380(24):2327-2340

224 Zhou X, Wu H, Ruan Q, Jiang N, Chen X, Shen Y et al. Clinical Evaluation of Diagnosis Efficacy of Active Mycobacterium tuberculosis Complex Infection via Metagenomic Next-Generation Sequencing of Direct Clinical Samples. Frontiers in Cellular and Infection Microbiology. 2019;9.

225 Yu G, Zhao W, Shen Y, Zhu P, Zheng H. Metagenomic next generation sequencing for the diagnosis of tuberculosis meningitis: A systematic review and meta-analysis. PLOS ONE. 2020;15(12):e0243161.

226 Marais S, Thwaites G, Schoeman J, Török M, Misra U, Prasad K et al. Tuberculous meningitis: a uniform case definition for use in clinical research. The Lancet Infectious Diseases. 2010;10:803-812.

227 Greninger A, Naccache S, Federman S, Yu G, Mbala P, Bres V et al. Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. Genome Medicine. 2015;7(1).

228 Vijayvargiya P, Jeraldo P, Thoendel M, Greenwood-Quaintance K, Esquer Garrigos Z, Sohail M et al. Application of metagenomic shotgun sequencing to detect vector-borne pathogens in clinical blood samples. PLOS ONE. 2019;14(10):e0222915.

229 Hedberg S, Ziegler I, Olcen P, Hallstrom B, Alm E, Krause J et al. Total metagenomics of bacterial DNA in whole blood from sepsis patients. Presentation presented at; 2022.

230 Salter S, Cox M, Turek E, Calus S, Cookson W, Moffatt M et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biology. 2014;12(1).

231 Chiu C, Miller S. Clinical metagenomics. Nature Reviews Genetics. 2019;20(6):341-355.

232 Fong W, Rockett R, Timms V, Sintchenko V. Optimization of sample preparation for culture-independent sequencing of Bordetella pertussis. Microbial Genomics. 2020;6(3).

233 Yang J, Yang F, Ren L, Xiong Z, Wu Z, Dong J et al. Unbiased Parallel Detection of Viral Pathogens in Clinical Samples by Use of a Metagenomic Approach. Journal of Clinical Microbiology. 2011;49(10):3463-3469.

234 Grumaz S, Stevens P, Grumaz C, Decker S, Weigand M, Hofer S et al. Next-generation sequencing diagnostics of bacteremia in septic patients. Genome Medicine. 2016;8(1).

235 Abhilash K, Jeevan J, Mitra S, Paul N, Murugan T, Rangaraj A et al. Acute undifferentiated febrile illness in patients presenting to a Tertiary Care Hospital in South India: clinical spectrum and outcome. Journal of Global Infectious Diseases. 2016;8(4):147.

236 Caserta M, Mock D, Dewhurst S. Human Herpesvirus 6. Clinical Infectious Diseases. 2001;33(6):829-833.

237 Akan H, İzbırak G, Gürol Y, Sarıkaya S, Gündüz T, Yılmaz G et al. Rotavirus and adenovirus frequency among patients with acute gastroenteritis and their relationship to clinical parameters: a retrospective study in Turkey. Asia Pacific Family Medicine. 2009;8(1):8.

238 Borensztajn DM, Hagedoorn NN, Rivero Calle I, Maconochie IK, von Both U, Carrol ED, et al.(2021) Variation in hospital admission in febrile children evaluated at the Emergency Department (ED) in Europe: PERFORM, a multicentre prospective observational study. PLoS ONE 16(1):e0244810. https://doi.org/10.1371/journal.pone.0244810

239 Leigh S, Grant A, Murray N,Faragher B, Desai H, Dolan S et al The cost of diagnostic uncertainty: a prospective economic analysis of febrile children attending an NHS emergency department BMC Med 2019 Mar 6;17(1):48. doi: 10.1186/s12916-019-1275-z.

240 Marquet M, Zollkau J, Pastuschek J, Viehweger A, Schleuβner E, Makarewicz O. Evaluation of microbiome enrichment and host DNA depletion in human vaginal samples using Oxford Nanopore's adaptive sequencing. Seientific Reports 2022 12:4000 DOI: 10.1038/s41598-022-08003-8

241 Morfopoulou S, Mee E, Connaughton S, Brown J, Gilmour K, Chong W et al. Deep sequencing reveals persistence of cell-associated mumps vaccine virus in chronic encephalitis. Acta Neuropathologica. 2016;133(1):139-147.

242 Weekly national Influenza and COVID-19 surveillance report Week 51report (up to week 50data) 23December2021 [Internet]. UK Health Security Agency; 2021 [cited 16 January 2022]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1043583/Weekly_Flu_and_COVID-19_report_w51.pdf

243 Bramugy J, Valente M, Ajanovic S, Baerenbold O, Hopkins H, Mabey D, et al. Fighting adult mortality through etiology of fever studies: Description of high mortality in an adult inpatient population in Mozambique, [poster] The American Society of Tropical Medicine and Hygiene 2021.

244 Hopkins H, Bassat Q, Chandler C, Crump J, Feasey N, Ferrand R et al. Febrile Illness Evaluation in a Broad Range of Endemicities (FIEBRE): protocol for a multisite prospective observational study of the causes of fever in Africa and Asia. BMJ Open. 2020;10(7):e035632.

245 Agency, U.H.S., HIV testing, new diagnoses, outcomes and quality of care for people accessing HIV services: 2021 report. 2021.

246 AIDSinfo, U., Country factsheets Malawi 2020 HIV and AIDS estimates. 2020.

247 Flygare S, Simmon K, Miller C, Qiao Y, Kennedy B, Di Sera T et al. Taxonomer: an interactive metagenomics analysis portal for universal pathogen detection and host mRNA expression profiling. Genome Biology. 2016;17(1).

248 Thair S, Seng H, Hollemon D, Hong D, Blauwkamp T, Kertesz M et al. The SEP-SEQ Trial: Clinical Validation of the Karius Plasma Next-Generation Sequencing Test for Pathogen Detection in Sepsis. Open Forum Infectious Diseases. 2017;4(suppl_1):S735-S735.

249 Brown, J.R., T. Bharucha, and J. Breuer, Encephalitis diagnosis using metagenomics: application of next generation sequencing for undiagnosed cases. Journal of Infection, 2018. 76(3): p. 225-240.

250 Bucciol G, M.L., Payne K, et al. , Patient with X-Linked Agammaglobulinemia. Journal of Clinical Immunology. Journal of Clinical Immunology, 2018.

251 NIH.Reporter.Gov. 2022 [cited 2022 15th Jan 2022]; Available from: https://reporter.nih.gov/search/IGVQ_nvguUq9X20LFEXVag/projects.