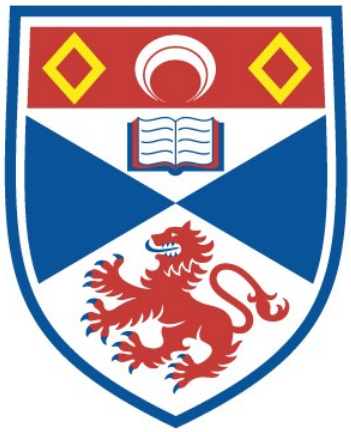
Whole genome sequencing service - investigating healthcare associated infections

Benjamin John Parcell

A thesis submitted for the degree of MD
at the
University of St Andrews



2021

Thesis declarations

Candidate's declaration

I, Benjamin John Parcell, do hereby certify that this thesis, submitted for the degree of MD, which is approximately 49,954 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree. I confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in November 2016.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date 18/02/2021 Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of MD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

Date Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Benjamin John Parcell, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

Date 18/02/2021 Signature of candidate

Date Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Benjamin John Parcell, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date 18/02/2021 Signature of candidate

Thesis abstract

Advances in the field of whole genome sequencing (WGS) have resulted in lowered costs, increased capacity and improved reproducibility of results. WGS now has the potential to revolutionise the investigation and management of healthcare-associated infection (HAI) outbreaks replacing conventional typing systems. The main objective of this work was to establish a WGS service for the investigation of suspected HAI outbreaks that could confirm or refute outbreaks in real-time in the National Health Service (NHS). WGS results were compared to conventional typing results and the practical barriers and clinical benefits associated with implementing this technology in a clinical environment were identified. Over a five year period a WGS service was used to investigate twenty one suspected outbreaks. The challenges of establishing a WGS service fell into six main areas: infrastructure; performance and quality assessment of data and processing; pipelines and management of reference databases; when to use WGS; clinical interpretation of results and finally when to use increased WGS discriminatory power in outbreak investigations. The clinical benefits of translating genomics into clinical practice comprised five themes: WGS can provide results with greater granularity than routine typing methods; genomic analysis can enhance the detection of "alert organisms'"; WGS can replace the need for multiple tests allowing streamlining of clinical microbiology services; genomic analysis can be used to rule out outbreaks and therefore minimise disruption to healthcare services; WGS can be utilised to investigate new resistance mechanisms. Identifying these practical barriers and clinical benefits informed the development of a clinical decision aid to assist staff on how best to utilise a WGS service. Implementing WGS as a standard of care in real-time was found to be a major advance in day-to-day IPC practice which is particularly relevant in view of the global threat we face with increasing antimicrobial resistance (AMR) and limited treatment options.

General acknowledgments

Firstly, I would like to thank my supervisors Professor Matthew Holden and Professor Stephen Gillespie at the University of St Andrews. I spent 3 months in St Andrews as a Visiting Scholar working with the team at the Infection and Global Health Division during the final year of my microbiology training, particularly looking at outbreak investigation using whole genome sequencing (WGS). From this, further opportunities arose to be involved with an application for funding from SIRN-CSO Scottish HAI Research Consortium Funding to form the Scottish Healthcare Associated Infection Prevention Institute (SHAIPI). My interest in research and in sequencing grew and I was encouraged every step of the way by Professor Holden and Professor Gillespie as I embarked on an MD. I am privileged to have worked with such inspirational supervisors, and have valued the advice and support they have given and the wisdom and experience they have shared. Their support has also been invaluable as I undertake an NHS Research Scotland (NRS) Fellowship.

I am incredibly grateful to Dr Katarina Oravcova for all the time she gave in which she provided me with training in practical procedures such as DNA extraction and running the Illumina MiSeq. I would like to thank Dr Kerry Pettigrew for all the laboratory work she has undertaken on the Illumina Miseq and all her support assisting me in the interpretation of sequencing data and her teaching on how to use SRST2 and ARIBA software. I would like to thank Professor Matthew Holden and Miguel Pinheiro for their time in which they undertook bioinformatics. I would particularly like to thank Professor Matthew Holden for interpreting sequencing data and for all of the teaching he gave me on bioinformatics. I would like to thank the staff at the Infection and Global Health Division, University of St Andrews for being so welcoming, inclusive, and inviting me to carry out more research in collaboration with them.

I would also like to thank Jane Turton from the National Reference Laboratory (Antimicrobial Resistance and Healthcare Associated Infections Reference Unit (AMRHAI), Public Health England, Colindale) for carrying out pulsed-field gel electrophoresis (PFGE) and variable-number tandem repeat (VNTR) typing, for her support in the interpretation of these techniques, and for allowing me to use their images in my thesis. Additionally, her

contributions to the publication '*Pseudomonas aeruginosa* intensive care unit outbreak: winnowing of transmissions with molecular and genomic typing' were vital.

I am very grateful for the funding I received from NHS Tayside and NHS Grampian to support this work. I would also like to extend my thanks to colleagues from NHS Tayside and NHS Grampian departments of Medical Microbiology, Infection Prevention and Control Team (IPCT), and Health Protection Team (HPT) for their invaluable assistance and contributions. In particular, I would like to thank Carolyn Sinclair for her assistance in producing statistical process control (SPC) c charts, Jennifer Lee for her assistance in producing timelines and graphs showing the measures instigated for borderline oxacillin resistant *Staphylococcus aureus* (BORSA), and Wai-Lum Sung Graphic Designer at the University of Aberdeen for his assistance in graphics development.

Finally, I wish to thank my wife Katherine and my children Charlie, Helen and Henry for their unending patience as I have attempted to balance my early consultant career research interests and pursuing my MD with my clinical work and home life. Since my MD first began we have relocated, welcomed Henry to our family and more recently dealt with the pressures of the pandemic as we shielded, home schooled and worked from home and I could not have completed this work without their support and love.

Funding

This work was supported by the Wellcome Trust ISSF award [grant number 097831/Z/11/Z]; and the Chief Scientist Office through the Scottish Infection Research Network [SIRN10].

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ABHR	Alcohol based hand rub
ACT	Artemis Comparison Tool
AFLP	Amplified fragment length polymorphism
AMR	Antimicrobial resistance
AMRHAI	Antimicrobial Resistance & Healthcare Associated Infections Reference Unit
AP-PCR	Arbitrarily primed PCR
AST	Antibiotic sensitivity testing
bp	Base pair
BC	Blood culture
BLAST	Basic local alignment search tool
BMA	British Medical Association
BORSA	Borderline oxacillin resistant *Staphylococcus aureus*
BSAC	British Society for Antimicrobial Chemotherapy
BSI	Bloodstream infection
CCD	Charged couple device
CDC	Centers for Disease Control and Prevention
CDI	*Clostridiodes difficile* infection
CF	Cystic fibrosis
CHEF	Contour clamped homogenous electric field
CLIMB-BIG-DATA	Cloud Infrastructure for Big Data Microbial Bioinformatics
COG-UK	COVID-19 Genomics UK
cPAL	Combinatorial probe-anchor ligation
cPAS	Combinatorial probe- anchor synthesis
CPE	Carbapenamase-producing *Enterobacterales*
CPHM	Consultant in Public Health Medicine
DEPC	Diethylpyrocarbonate
DH	Department of Health
DNA	Deoxyribonucleic acid
DoD	Department of Defence
DRE	Double repetitive element

ECCMID	European Congress of Clinical Microbiology & Infectious Diseases
ECDC	European Centre for Disease Prevention and Control
ENA	European Nucleotide Archive
ERIC	Enterobacterial repetitive intergenic consensus sequences
ESBL	Extended-spectrum β-lactamase
ETA	Endotracheal aspirate
EUCAST	European Committee on Antimicrobial Susceptibility Testing
EVD	Ebola virus disease
FDA	Food and Drug Administration
FDA-ARGOS	FDA dAtabase for Regulatory-grade microbial Sequences
FFP3	Filtering Face Piece 3
FIS	Federation of Infection Societies
GAS	Group A *Streptococcus*
GB	Gigabytes
GI	Gastro-intestinal
GMC	General Medical Council
GP	General Practice
GPs	General practitioners
HAI	Healthcare-associated infections
HAP	Hospital acquired pneumonia
HCCA	α-Cyano-4-hydroxycinnamic acid
HCWs	Healthcare workers
HIIA	Hospital Infection Incident Assessment
HiRECCs	High-risk enterococcal clonal complexes
HIRUs	High-infection risk units
HIS	Healthcare Infection Society
HPS	Health Protection Scotland
HPTs	Health Protection Teams
HTS	High-throughput sequencing
ICUs	Intensive care units
IMT	Incident Management Team
IPC	Infection prevention and control
IPCD	Infection prevention and control doctor

IPCN	Infection prevention and control nurse
IPCT	Infection prevention and control team
IS	Insertional sequences
KPC	*Klebsiella pneumoniae* carbapenemase
KPC-KP	*Klebsiella pneumoniae* carbapenemase (KPC)-producing *K. pneumoniae*
KWS	Kauffmann White Scheme
LIMS	Laboratory information management system
LIS	Laboratory information system
MALDI-TOF MS	Matrix-assisted laser desorption ionization-time of flight mass spectrometry
MDR	Multidrug resistance
MDRO	Multidrug resistance organisms
MGEs	Mobile Genetic Elements
MDR-TB	Multidrug resistant *Mycobacterium tuberulcosis*
MHDUs	Medical high dependency units
MIC	Minimum inhibitory concentration
MLEE	Multilocus enzyme electrophoresis
MLST	Multilocus sequence typing
MLVA	Multilocus VNTR analysis
MRC	Medical Research Council
MRSA	Meticillin resistant *Staphylococcus aureus*
NCBI	National Center for Biotechnology Information
NDM	New Delhi metallo-β-lactamase
NGS	Next generation sequencing
NHS	National Health Service
NICUs	Neonatal ICUs
NRS	NHS Research Scotland
NVS	Nutritionally variant streptococci
PAG	Problem Assessment Group
PFGE	Pulse-field gel electrophoresis
PGRS	Polymorphic guanine/cytosine-rich repetitive sequences
PHE	Public Health England
PPA	Plasmid profile analysis

PPE	Personal protective equipment
RAPD	Random amplification of polymorphic DNA
RDU	Renal dialysis unit
REA	Restriction endonuclease analysis
REALPHY	Reference sequence Alignment based Phylogeny builder
REP	Repetitive extragenic palindromic sequences
rep-PCR	Repetitive element PCR
RFLPs	Restriction fragment length polymorphisms
RNA	Ribonucleic acid
SBL	Sequencing by ligation
SCBU	Special baby care unit
SCN	Senior charge nurse
SHAIPI	Scottish Healthcare Associated Infection Prevention Institute
SHDUs	Surgical high dependency units
SICPs	Standard Infection Control Precautions
SLST	Single locus sequence typing
SMiRL	Scottish Microbiology Reference Laboratories, Glasgow
SMRT	Single molecule real-time
SMS	Single molecule sequencing
SNP	Single nucleotide polymorphism
SPC	Statistical process control
SSI	Surgical site infections
SSR	Simple sequence repeats
ST	Sequence type
SVR	Short variable region
TATs	Turn-around times
TB	Tuberculosis
TBPs	Transmission based precautions
UK	United Kingdom
USA	United States of America
VHFs	Viral haemorrhagic fevers
VNTR	Variable-number tandem repeat
VRE	Vancomycin-resistant enterococci
VREfm	Vancomycin-resistant *E. faecium*

VSEfm	Vancomycin-susceptible *E. faecium*
WGS	Whole genome sequencing
WHBs	Wash-hand basins
WHO	World Health Organisation
WRAS	Water regulations advisory scheme
XDR TB	Extensively drug-resistant TB
ZMW	Zero mode waveguides

1 Introduction

1.1 Evolution of molecular epidemiology

The delivery of healthcare should not cause harm to patients. This key principle attributed to Hippocrates in the fourth century B.C has been adopted by medicals schools and accreditation bodies such as the United Kingdom (UK) General Medical Council (GMC).(1) Despite this, prevention of healthcare-associated infections (HAI) continues to be one of the greatest challenges to the successful delivery of healthcare. The concept of germ theory in 1546 in which Girolamo Fracastoro proposed that minute particles are spread by direct contact, through the air and by clothing was a turning point in infection prevention and control (IPC).(2) Since then healthcare providers and researchers have strived to prevent the transmission of infection making efforts to understand disease by describing different infections with varying incidence and severity.(3) As a result of this, the term epidemiology was borne now defined as: *"the study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the control of health problems"*.(4) Over time, various typing schemes based on serology or other phenotypic markers were developed to classify microorganisms and disease.(3) Using epidemiology and typing methods, researchers identified that some infections were due to certain serotypes of bacteria and with this, the term "infectious disease epidemiology" came into being.(3)(5) This branch of epidemiology consists of quantitative and descriptive assessments of disease occurrence, reservoirs, patterns and modes of transmission and biological factors related to pathogens and hosts that influence transmission.(5) This field has evolved rapidly in parallel with new molecular diagnostic developments in the clinical microbiology laboratory resulting in the formation of another subdivision of epidemiology namely "molecular epidemiology of infectious diseases".(6) Levin *et al.* stated in 1999 that the goals of this are to "identify the micro- parasites (viruses, bacteria, fungi, and protozoa) responsible for infectious diseases and determine their physical sources, their biological (phylogenetic) relationships, and their routes of transmission and those of the genes (and accessory elements) responsible for their virulence, vaccine-relevant antigens, and drug resistance".(3)

Each day the clinical microbiology laboratory plays a critical role in the detection, surveillance, and prevention of microorganism transmission in the healthcare environment. Information such as organism identification, antibiotic susceptibility and epidemiologic typing is elicited from cultures taken from patients.(7) This vital resource can be actively used as part of surveillance by the infection prevention and control team (IPCT) to refute or confirm clonality of organisms thereby aiding in their quest to contain the single clones of bacteria and their mobile genetic elements for antimicrobial resistance spreading within a population.(8) At present determining whether transmission has occurred is a major challenge for IPCTs as current typing methods do not always give results that have sufficient granularity or robustness to unequivocally define strains. In addition, there is often an inbuilt delay in receiving typing results as isolates are often sent to reference laboratories for typing and epidemiological data is not always available to teams to establish links between patients and the environment. There is great potential to optimise clinical microbiology laboratory processes to enable early detection, enhanced surveillance, and refutation of outbreaks to produce actionable results. Developments in whole genome sequencing (WGS) have resulted in increased capacity, lowered costs, improvement in speed and reproducibility of results and so in recent years WGS has emerged as the ultimate typing tool.(9)(10)(11) WGS can be used to pinpoint and track bacteria to a greater degree than traditional typing methods and in some cases it can be used in real-time to inform the stepping down of IPC measures in hospital.(12) WGS has been applied to the investigation of a wide variety of outbreaks for example a meningococcal serogroup B (MenB) outbreak, *Legionella* outbreaks and it has been found to be a useful tool to delineate outbreaks of *Mycobacterium tuberculosis*.(13)(14)(15) Potentially WGS technology has the capacity to allow outbreaks to be routinely detected in real-time, enabling rapid implementation of targeted IPC measures. WGS may also be used to rule out transmission and prevent the need for beds to close, minimising disruption to clinical services. There is, however, limited data on how best to establish a WGS service for real-time outbreak investigation and as yet it has not been fully applied operationally throughout clinical microbiology laboratories or reference laboratories in the UK to meet these clinical needs.

1.2 Occurrence and distribution of HAI

Purpose built hospitals have existed in Greece, Egypt, India, and Palestine since 500 B.C. and the Charaka-Samhita, a Sanskrit medical text describes the earliest recommendations of hospital hygiene advising that hospital workers should be known for their good behaviour, clean habits, and purity. It mentions the advantages of consulting engineers to construct hospitals and provides insight into the benefits of using ventilation to reduce smoke, dust, harmful sounds, scents and tastes.(16) These IPC strategies are recognised today to be effective methods to prevent and control HAI.(17) However, European hospitals were slow to adopt these practices with the need for hospitals only being fully recognised by religious orders in the 12th century.(18) They remained unsafe with frequent outbreaks and patients dying in large numbers due to small pox, "hospital fever" (louse –borne typhus), plague, typhoid and dysentery. During this time, patient placement and isolation were not carried out and multiple patients occupying single beds resulted in the spread of infection. Postoperative mortality rates ranged between 60%-80% with surgery being carried out by barbers/surgeons using no asepsis or anaesthesia. This resulted in patients commonly dying from post surgery hospital (streptococcal) gangrene.(18) Despite the evident advances in healthcare the problem of HAI remains as relevant now as ever and it is imperative that IPC is accepted to be a high-level priority for patient safety. Presently HAIs comprise any infection arising from care given in a health-care facility or hospital that was not incubating or present on admission.(19) This term also covers infections that patients have acquired in hospitals that may not have been apparent whilst patients were admitted and as part of this, infections which are diagnosed after discharge. Additionally the term also encompasses infection acquired by staff whilst at work.(19) The terms "hospital acquired" or "nosocomial" infection are becoming outdated as it is appreciated that the delivery of care has changed with more reliance in the UK on the National Health Service (NHS) providing care in the community in ambulatory settings such as outpatient, day patient settings, general practice (GP) and at home in partnership with social care.(20)

The Scottish National HAI and Antimicrobial Prescribing Point Prevalence Survey 2016 found the prevalence of HAI in acute Scottish hospitals to be 4.5%. This equates to 55,500 infections every year or approximately 1 in 22 acute adult inpatients having at least 1 HAI.(21) This survey also identified that the prevalence of HAI in non-acute hospitals to be

3.2%. These results are similar to findings from the largest European point prevalence survey which identified one in 18 patients to have an HAI.(22) In this survey HAI prevalence varied in relation to location for instance primary hospitals also known as district hospitals had the lowest HAI prevalence of 4.8%. For secondary hospitals HAI prevalence was 5.0%, tertiary hospitals 7.2% and for specialised hospitals it was 6.0%. In this survey patients in Intensive Care Units (ICUs) had the highest HAI prevalence with 19.5% of patients experiencing at least one HAI. For patients, HAIs may result in pain, disability, and psychological effects from isolation measures. Additionally HAIs are associated with higher rates of mortality and inpatient costs and it has been estimated there are approximately 37,000 deaths per year directly due to HAIs in Europe.(23)(24). HAIs result in 16 million extra days of hospital stay in Europe and patients may experience loss of their own earnings due to increased length of stay.(25) Considerable additional funds are required for the management of HAIs and in Scotland the inpatient cost of HAI is estimated to be approximately £137 million a year.(21) HAI outbreaks are a significant risk to patient safety, as well as being costly and time consuming to investigate requiring additional staff time and resources such as additional staffing, cleaning resources, personal protective equipment **(**PPE**),** and antibiotic use. Hospital services may experience disruption due to ward closures, cancellation of procedures/surgeries or clinics and staff exclusion resulting in anxiety for patients, their relatives, or the wider community. Patients may also have to undergo additional screening tests in the investigation of outbreaks. These can include nasal, perineal, throat and rectal swabs taken in order to determine the extent of transmission and risk of infection. HAIs can impact on employees resulting in reduced staff morale, poor public image and increased scrutiny and inspections by the Department of Health (DH) Improvement Teams when performance targets are not met.(26) At least 20% of HAI are considered to be potentially avoidable.(27) Surveillance and rapid identification of patients is critical to allow timely implementation of measures to prevent the spread of these infections and reduce the impact to patients, families and staff.

Point prevalence studies have shown that similar types of HAIs predominate across the world. The largest European point prevalence survey most frequently reported HAIs included respiratory tract infections (pneumonia 19.4% and lower respiratory tract 4.1%), surgical site infection (SSI) (19.6%) urinary tract infections (UTI) (19.0%) bloodstream infections (BSI) (10.7%) and gastro-intestinal infections (GI) (7.7%), with *Clostridiodes difficile* infections (CDI) representing 48% of the gastro-intestinal infections (3.6% of all

HAIs).(28) The European Centre for Disease Prevention and Control (ECDC) point prevalence survey found *Escherichia coli* to be the most frequently isolated microorganism from HAIs present in 15.9% of infections. This was followed by *Staphylococcus aureus* (12.3%), *Enterococcus* species (9.6%), *Pseudomonas aeruginosa* (8.9%), *Klebsiella species* (8.7%), Coagulase-negative staphylococci (7.5%), *Candida* species (6.1%), *Clostridium difficile* (5.4%), *Enterobacter* species (4.2%), *Proteus* species (3.8%), *Acinetobacter* species (3.6%), *Serratia* species 1.1%, *Stenotrophomonas maltophilia* (1.0%) and *Aspergillus* species (0.4%).(28) Antimicrobial susceptibility testing (AST) results were available for 85% of HAIs and a significant number of multidrug resistance organisms (MDROs) were detected with 41.2% of *S. aureus* found to be Meticillin resistant *Staphylococcus aureus* (MRSA). Vancomycin-resistant enterococci (VRE) were detected in 10.2% of *Enterococcus* species. It was also identified that 33.4% of all *Enterobacterales* were resistant to third-generation cephalosporins and this was highest in *Klebsiella pneumoniae* isolates. Carbapenem resistance was found to occur in 7.6% of all *Enterobacterales* (highest in *K. pneumoniae*), 31.8% of *P. aeruginosa* isolates and in 81.2% of *Acinetobacter baumanni*.(28) HAIs due to MDROs are particularly concerning are they are associated with an increase in both patient mortality and readmissions, compared to those due to susceptible strains.(29)

1.3 Risk factors for HAI related to hosts

Risk factors for HAI depend on the type of healthcare patients receive and where they receive care. The commonest risk factors described by the World Health Organisation (WHO) independently associated with HAI can be seen in Table 1.1.(19)

Table 1.1. Risk factors independently associated with HAI

Risk factors independently associated with HAI
Age >65 years.
Admission as an emergency and to the intensive care unit (ICU).
Admission duration more than seven days.
Placement of a central venous catheter.
Indwelling urinary catheter.
Endotracheal tube.
Undergoing surgery.
Trauma-induced immunosuppression.
Neutropenia.
A rapidly or ultimately fatal disease (according to the McCabe-Jackson classification).
Impaired functional or coma status.
Additional determinants in low- and middle-income countries
Malnutrition.
Age < 1 year.
Low birth weight.
Parenteral nutrition.
Two or more underlying diseases.

Various factors can contribute to HAI and these can be related to the patient's condition, healthcare interventions, and the environment. (30) Risk factors include immunosuppression due to disease, patient demographics such as increased age and body mass index (BMI), use of interventions, surgical procedures and invasive devices made of prosthetic materials for example urinary catheters, lines, drains or prosthetic implants.(19)(30)(31) These devices breach skin and mucous membranes and provide a portal of entry to the body. They are also commonly made of material that encourages biofilm formation making infection more difficult to treat due to reduced antibiotic penetration. HAI burden is more severe in high-risk populations such as patients admitted to high dependency settings e.g. adult intensive care units (ICUs), neonatal ICUs (NICUs), medical or surgical high dependency units (MHDUs, SHDUs) and units providing the care for patients with burns, transplants or neurological conditions requiring surgery. The proportion of infected patients in ICU has been reported to be as high as 51%.(32) Patients in high dependency settings are susceptible to infection for

several reasons. Immune responses are often diminished by the stress and metabolic effects of existing disease. Phagocytosis can be impaired in critically ill patients.(33) This patient group may have also recently undergone anaesthetic and surgical procedures. Many barriers to infection provided by the innate immunity are breached because of the need for intravenous catheters, invasive monitoring, urinary catheterisation, artificial ventilation, and procedures such as dialysis. In the neonatal period the consequences of HAI can be catastrophic impacting on survival and neurodevelopmental outcomes.(34) The majority of these infections affect preterm infants, those with low-birth-weight, or those who have undergone surgery, with HAI rates in NICU ranging from 6.0% to 50.0% per admission.(35)(36) This population of patients is particularly vulnerable to acquiring HAI due to a variety of different factors. Neonates are likely to undergo invasive procedures involving devices such as mechanical ventilation and central venous catheters. They may require parenteral nutrition and H2 blocker/proton pump inhibitors all of which can increase the risk of infection.(37)(38) Neonates additionally have impaired host defence mechanisms with limited protective endogenous skin and mucosal flora, along with reduced skin barrier function putting them at risk of colonization and infection from the hands of healthcare workers (HCWs) and environmental microorganisms.(36)(37) Outbreaks of HAI are a significant threat to neonates and after analysing world wide databases of HAI outbreaks it was found that the average mortality in NICU outbreaks is 6.4%.(39) Outbreaks are also associated with increased healthcare costs and staffing issues.(34) In a recent survey of the United Kingdom 12.2% of NICUs had closed due to IPC issues in the previous year and 14.1% had current IPC issues.(40) These closures impact on the delivery of healthcare services and patients may have to be transferred to units further afield impacting on patients and families. The investigation and management of outbreaks can also cause disruption for staff, units and visitors as targeted IPC measures are often required such as isolation, additional contact precautions and screening. Lack of financial support, inadequate training in IPC, lack of staff and equipment or supplies have not been shown to be independent risk factors but are barriers to optimal IPC practices.(19) The environment that patients are treated in can increase the risk of HAI acquisition and guidance on the general design for healthcare buildings recommends that “healthcare facilities should provide a therapeutic environment in which the overall design of the building contributes to the process of healing and reduces the risk of HAIs rather than simply being a place where treatment takes place”.(41) The British Medical Association (BMA) also advise that consideration should be given to the design of clinical areas in new healthcare buildings, with a view to introducing single occupancy rooms instead

of multi-bed wards.(42). Despite this, environmental factors due to the nature of healthcare facilities can increase the risk of HAI for vulnerable patients. Many hospitals have been built before the development of these guidelines and some patient areas require major repair. Recent experience has demonstrated that new buildings are not exempt from problems either.(43)The Scottish Centre for Reducing Infection and Risk in the Healthcare Environment is a new national body that was formed in 2019 to have oversight into the building and design of major NHS infrastructures and for the production of guidance and policy in relation to outbreaks.(44) Air-conditioning and water systems in particular can become contaminated contributing to the development of HAIs. The BMA also recognise that standards for HAI prevention are at risk of being compromised when hospitals are under pressure during times of increased patient admissions. This can result in inappropriately low staff levels, reduced hand hygiene and a reduction in the recommended space between patient beds.(42) Figure 1.1 illustrates the factors involved in the development of infection (adapted from Harris A, 2008).(45)

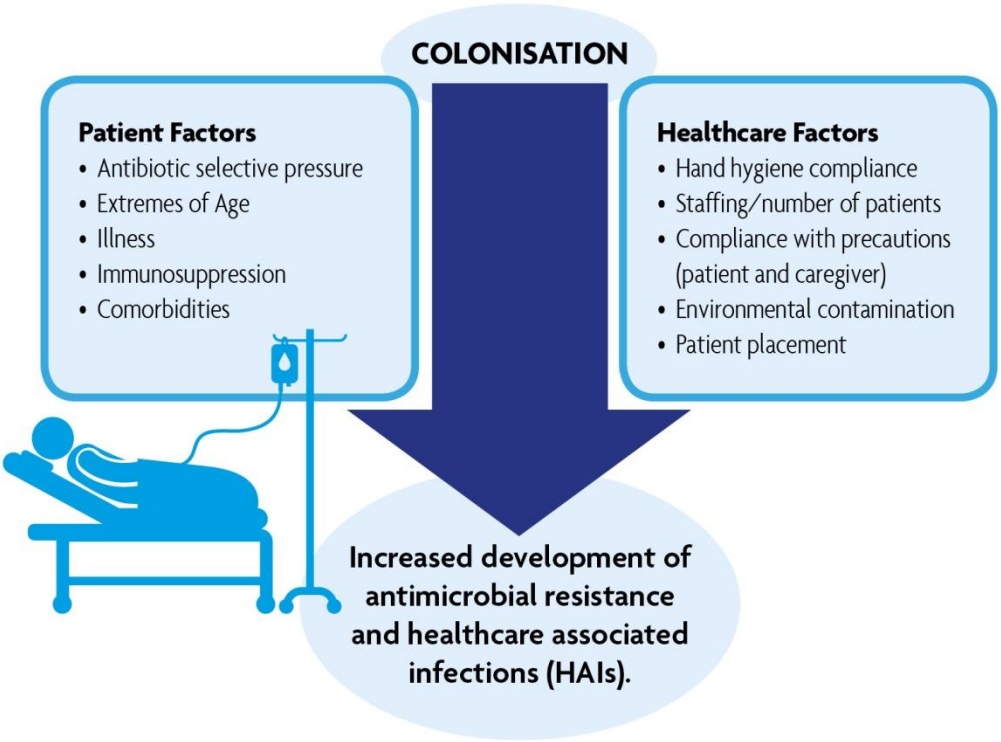


Figure 1.1. Overview of the development of infection in a patient accessing healthcare. Patient and healthcare factors are shown to highlight the risk of the development of infection and antimicrobial resistance in a colonised patient (adapted from Harris A, 2008).

1.4 Modes of transmission of HAI

As described previously "infectious disease epidemiology" involves quantitative and descriptive assessments of the patterns and modes of transmission of disease. To prevent outbreaks it is essential to have an understanding of this.(5) The origin of the term "transmission" can be traced back to its use in Latin in the 4th century C.E. "*Trans*" pertained to "across" and the verb "*mittere*" expressed "to send".(46) It was used in medicine in the Renaissance period by Girolamo Fracastoro in his book written in 1546 on contagion entitled "On Contagion and Contagious Diseases (*De contagione et contagiosis morbis*)". In this book he described his observation that there were two types of fever: one caused from within and another due to external factors such as the air, germs, or vapours. He proposed that the latter was "transmitted" from one person to another (*febris in nobis primo pestilens sit et ab uno in alium transmittitur*).(46) This was an important step in the understanding of disease propagation and a similar observation was made by the Public Heath Physician John Snow who investigated a cholera outbreak in London in 1854. He recorded cholera deaths and water consumption and identified that infection had been transmitted from a pump in Broad Street.(5)As a result of this, the local council removed the pump handle. A quantitative understanding of transmission is an essential part of "infectious disease epidemiology" that can be used to inform evidence-based prevention and control measures to contain transmission of microorganisms.(5) HAI may occur as a result of transmission of organisms from exogenous or endogenous sources. Exogenous sources include the environment, equipment used by HCWs, and also HCWs and visitors. Endogenous sources include the patient's own microflora.(47) A full appreciation of the chain of infection is paramount so that HAIs can be rapidly identified and prevented. Transmission can be stopped by breaking one or more or the links in the chain of infection. Figure 1.2 shows an overview of the chain of infection (image based on text by Damani, 2012).(47)

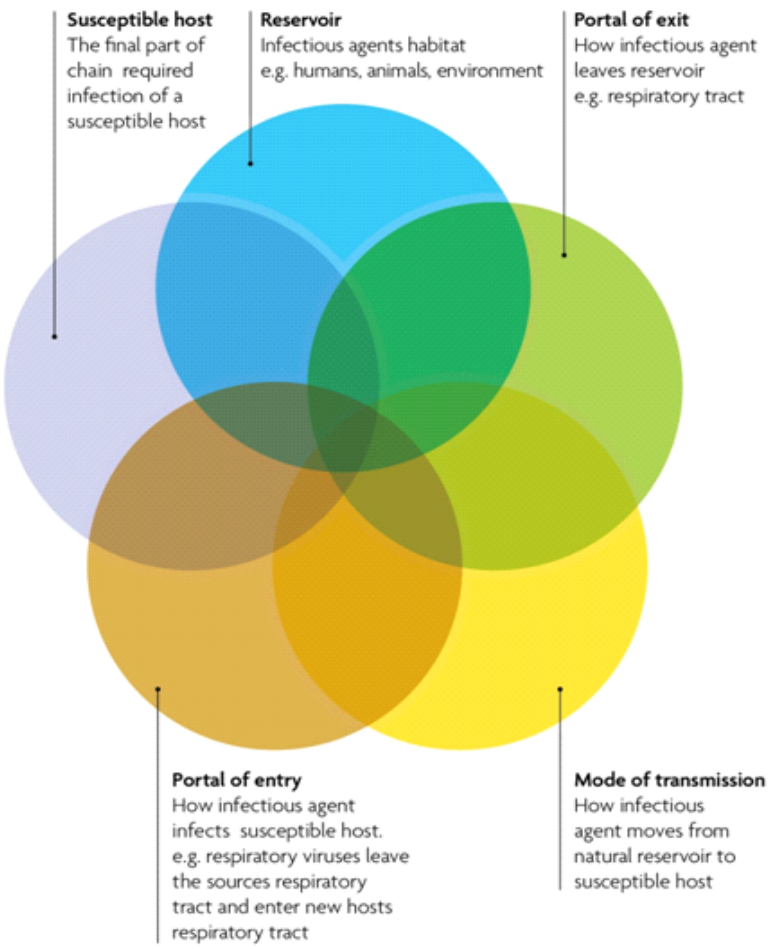


Figure 1.2. Overview of the chain of infection. This figure demonstrates the concept of how infections are acquired. Pathogens survive in their own host/reservoir and leave via a portal of exit. They then follow a mode of transmission and portal of entry to infect a host (image based on text by Damani, 2012).

Modes of transmission include direct contact e.g. skin-to-skin contact, contact, droplet spread e.g. produced by sneezing, coughing and indirect transmission i.e. transfer of an infectious agent from reservoir to a host by air particles, inanimate objects (vehicles), or animate intermediaries (vectors).(48) The most common causes leading to HAIs and outbreaks include poor hand hygiene, contaminated medical devices and failure to comply with local policies, procedures and guidelines.(49) It is therefore essential that organisations ensure that the core components of IPC such as standard infection control precautions (SICPs) and transmission based precautions (TBPs) are in place and adhered to.
SICPs apply to staff caring for every patient and can be seen in Table 1.2.

Table 1.2. Overview of standard infection control precautions (SICPS) to prevent transmission

SICPs
Patient placement
Hand hygiene
Respiratory and cough hygiene
Correct use of PPE
Safe management of care equipment
Safe management of care environment
Safe management of linen
Safe management of blood and body spillages
Disposal of waste in including sharps
Occupational safety; prevention and exposure management (including sharps)

SICPs, standard infection control precautions; PPE, personal protective equipment.

SICPs are basic measures used to reduce the chance of transmission of organisms.(50) They include patient placement in which patients should be assessed for risk of infection. Patients should be treated in a side room if they have any respiratory symptoms, diarrhoea, vomiting, unexplained rash, fever, colonisation with multidrug resistance organisms (MDRO) or prior hospitalisation abroad. Other SICPS include hand hygiene, respiratory and cough hygiene, PPE, safe management of care equipment and environment, safe management of lines, blood and blood fluid spillages care, disposal of waste (including sharps) and occupational safety. These measures should always be used to prevent risk of transmission of microorganisms for all patients.(50) There are times when the application of SICPs may not be enough to reduce the risk of transmission of specific infectious organisms. Transmission based precautions (TBPs) should be used in instances in which patients are incubating an infection, colonised with an infectious agent or have symptoms of infection. TBPs include contact, droplet or airborne precautions.(51) The most common route of cross-infection transmission is by direct contact with patients or from touching the immediate care environment around the patient and care equipment. Contact precautions should be carried out when there is likely to be direct contact with patients and when HCWs are likely to touch patient's immediate environment. If a patient has diarrhoea the risk of transmission by the faecal oral route

increases. When patients have diarrhoea, oozing wound or body fluids gloves and apron should be used. For patients with fever, cough/ influenza-like illness or vomiting a side room is preferred. Organisms may be transmitted and spread by droplets (>5μm) from a person's respiratory tract landing onto another person's mucosal surface or conjunctivae. Droplet precautions include face masks and staff are advised to wear fluid resistant surgical mask, gloves, and apron in these instances.(51) For patients with severe respiratory illness e.g. high risk for multidrug resistant *Mycobacterium tuberulcosis* (MDR-TB) or measles-like rash staff should wear Filtering Face Piece 3 (FFP3) mask, gloves, apron and the patient should be nursed in side room with negative pressure ventilation. Infections can also be spread by aerosols (≤5μm) from a person's respiratory tract onto a mucosal surface or conjunctivae of another individual. Airborne precautions include FFP3 respirators.(51). It is considered that outbreaks should be the exception when SICPs and TBPs are in place, unfortunately the development of HAIs and outbreaks are still the commonest adverse event during the delivery of healthcare.

1.5 Surveillance of infection and outbreaks

Most HAIs become apparent 48 hours or more after a patient has been admitted. This period of time is thought to be the typical incubation period.(19) In some circumstances patients or healthcare staff may recognise an HAI outbreak and identify the cause. This may occur during outbreaks that occur over a short period of time and involve pathogens that present with obvious symptoms such as vomiting or diarrhoea for example norovirus. It is also possible that outbreaks may not be identified by staff or patients and they can remain undetected for a variety of reasons. In some situations, HAIs occur as part of an endemic ongoing trend over a long period of time. Clustered organisms can arise sufficiently too far apart in time or across different locations. In some instances, IPC measures may not be put in place to halt transmission as there may be a breakdown in communication of results for example patients may be discharged home or to other healthcare areas while samples are being processed with outstanding microbiology results. Transmission may not be identified as it may not necessarily result in infection but colonisation where the individual is asymptomatic. Staff or the patient may never identify that transmission has taken place unless further screening is carried out or the patient submits clinical samples in which the organism is later identified. Even in this circumstance, staff would need a way of recognising

that the organisms are similar to the original transmitted organism and microbiology and IPC surveillance is required to identify this.

Surveillance of HAIs and rapid identification of outbreaks are a strategy to limit the spread of infection.(52)(8) Surveillance is defined as: *"ongoing systematic collection, analysis and interpretation of health data essential to planning, implementation and evaluation of public health practice"*.(19) It is a key part of improving the quality of patient care and there is evidence that surveillance can result in reduced rates of infection. Various types of surveillance exist and an overview of these can be seen in Table 1.3.(19)(53)

Table 1.3. Types of surveillance to monitor transmission

HAI surveillance type	Overview
Incidence surveillance	Systematic, regular data collection in a population over a period of time.
Prevalence surveys	Records number of specified events in a specific population at a point in time or over a specified period of time.
Passive surveillance	Commonest form of surveillance using routine data from automated patient records, e.g. laboratory-based surveillance/ data from patient records. Low sensitivity can lead to misclassification and underreporting as criteria for diagnosis may be not easily available. However, it is less demanding and could be carried out in settings that are unable to carry out active surveillance due to lack of resources.
Active surveillance	Should be conducted by trained staff looking for evidence to meet standardized HAI diagnostic criteria. Higher specificity and sensitivity than passive surveillance
Prospective surveillance	Gold standard for the collection of reliable and timely information. Monitors pre-selected indicators in patients, according to specific protocols. Can be extended to the post-discharge period. More resource intensive and time consuming than retrospective surveillance
Retrospective surveillance	Relies on collecting routine data post discharge therefore some information may be missing which may make fulfilling diagnostic criteria harder.

HAI, healthcare associated infection.

Surveillance can be passive or active, and can involve virtual surveillance such as the application of mathematical models to detect outbreaks or pattern identification and data mining. (8) In order for surveillance to be effective it is essential that there is ongoing education for healthcare staff. The Centers for Disease Control and Prevention (CDC) first developed standard definitions for nosocomial infections and surveillance methods in the

1970s' and these were used by infection control committees to influence practice. (54) The commonest type of surveillance is passive surveillance in which data is pulled from patient records e.g. laboratory results. Active surveillance has higher sensitivity and specificity and is conducted by trained staff using many data sources in order to meet standardised diagnostic HAI criteria. The World Health Organisation (WHO) recommends that that prospective surveillance is the gold standard as data is more likely to be reliable and timely although it can be more expensive compared to retrospective surveillance.(19) The microbiology laboratory plays a vital role in providing surveillance information. In the UK microbiology laboratories take part in voluntary reporting to Public Health England (PHE) as part of disease surveillance in England. Clinical microbiologists were first encouraged to report identification and antibiotic sensitivity results from blood culture isolates to LabBase2 (a national database) in 1989 so that national trends in resistance could be generated. There are also surveillance systems in place which involve sentinel laboratories that collect bacterial isolates which then send them to national reference laboratories. The British Society for Antimicrobial Chemotherapy (BSAC) sponsored two programmes in which respiratory isolates and bacteraemia isolates were collected.(55) HAI prevalence across Europe is collected by the ECDC.(22) Other national systems include the National Healthcare Safety Network (NHSN) developed by the CDC which provides infection rates which can be used by healthcare staff and national healthcare planners to set targets and evaluate their work. In the United States of America (USA) this is the most widely used HAI tracking system.(56) Over the last 20 years surveillance has evolved and now concentrates on problem categories such as high risk areas for example ICUs, infections that can be prevented e.g. intravascular device related infections and emerging pathogens/resistance. At national levels surveillance can be used to monitor trends over time, detect outbreaks and emerging threats, investigate the impact of interventions, inform strategic planning, policies on IPC and prescribing and drive improvements.(31)(49) It can be used to allocate resources, determine best practice, facilitate benchmarking and identify HAI trends for instance *S.aureus* bloodstream infections, *C. difficile* infection and SSIs.(31)(49) This has enabled targets across NHS boards to be set that aim to improve patient care. Local surveillance programmes allow the IPCTs to detect local outbreaks at the earliest opportunity. They are required to follow the nationally agreed minimum list of alert organism/conditions.(57) These organisms can be particularly significant in certain clinical areas e.g. high-infection risk units (HIRUs) in which patients are most vulnerable to infection for example burns units, transplant, cancer care, ICUs, renal departments and cystic fibrosis (CF) units.(57) At a local level baseline endemic rates can be

established through alert organism surveillance revealing trends in infection. This information aids in incident recognition and can be used as an early warning system when abnormal levels are identified triggering action from IPCTs, Health Protection Teams (HPTs) and clinical teams to assist staff to review or change factors locally.(57) Examples can be seen in Figure 1.3.

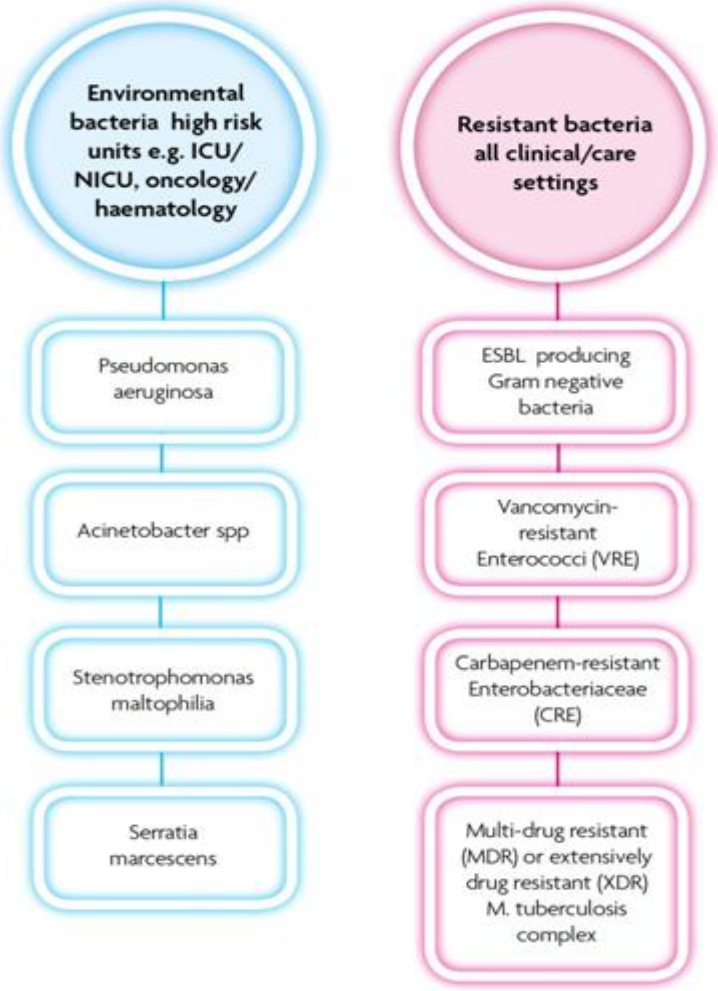


Figure 1.3. Figure demonstrating alert organisms and clinical locations used to track transmission. Alert organisms are pathogens that are particularly transmissible and virulent. In certain clinical settings patients are more prone to risk of infection by certain pathogens (adapted from Appendix 13 – NHSScotland list of alert organisms / conditions. 2017;(April). Available from: http://www.nipcm.hps.scot.nhs.uk/appendices/appendix-13-mandatory-nhsscotland-alert-organismcondition-list/)

Microbiologists, Infection specialists, and Infection Prevention and Control Doctors (IPCDs) may carry out laboratory bench rounds reviewing alert organisms and authorising results as

part of this. The aim is to recognise organisms and antibiograms that are similar in different patient samples suggesting there is an outbreak. Microbiology laboratories have computer reporting systems called laboratory information management systems (LIMS), also referred to as a laboratory information system (LIS). In addition to storing data such as sample type, date, location, organism's identification, and susceptibility profile these computer systems can be used to produce microbiology reports and audit turn-around times and quality. This data can be used to identify alert organisms or clusters of similar organisms in specific sample types or patients from any location.(58) These systems can be automated and set up so that reports are created to detect trends of infection and resistance for daily review. They may also be linked to ICNet infection surveillance software which alerts the IPCT to potential outbreaks.(59) Automated outbreak detection tools which are statistically-based have been found to be beneficial in terms of detecting outbreaks and streamlining work for IPCTs.(60) Novel applications have been developed in which routine IPC data such as ward admission, date of samples and pathogen genomic information can be visualised. An example is Shiny - a web application for statistical software R.(61) Other platforms using R software have also been made available such as OutbreakTools which was developed for storing and visualising data.(62) Microreact software is also freely available and can be utilised to interactively upload dendrograms and to provide phylogenetic snapshots of population diversity in geographic locations (63)(64)

Local surveillance systems should be set to have a trigger/ threshold that prompts IPC action and sending of bacterial isolates for typing at reference laboratories. Triggers can be due to outbreaks but in some instances may be due to natural variation in alert organism incidence.(65) They may also be a single event as the background level of a specific alert organism may be zero for examples *Legionella* in any clinical location. Prevalence of alert organisms varies with patient population and healthcare risk factors and therefore the same trigger cannot be set for all clinical areas. Benefits of setting triggers includes that they signal problems at an early stage so that staff can manage outbreaks before organisms spread further. Statistical Process Control (SPC) c charts are used by IPCTs and can be used to demonstrate triggers showing chronological data and natural or unnatural variation. They apply statistical theory to quality control and calculations are made from the clinical areas previous data to produce three lines.(65) The centre line shows the average number of alert organisms per month. The trigger line (trigger) represents a warning limit and if results meet or go over the trigger, action should be taken. The upper control limit is the limit of natural

variation and any results above this is unnatural variation and out of statistical control.(65) An example of an SPC chart is shown in Figure 1.4.

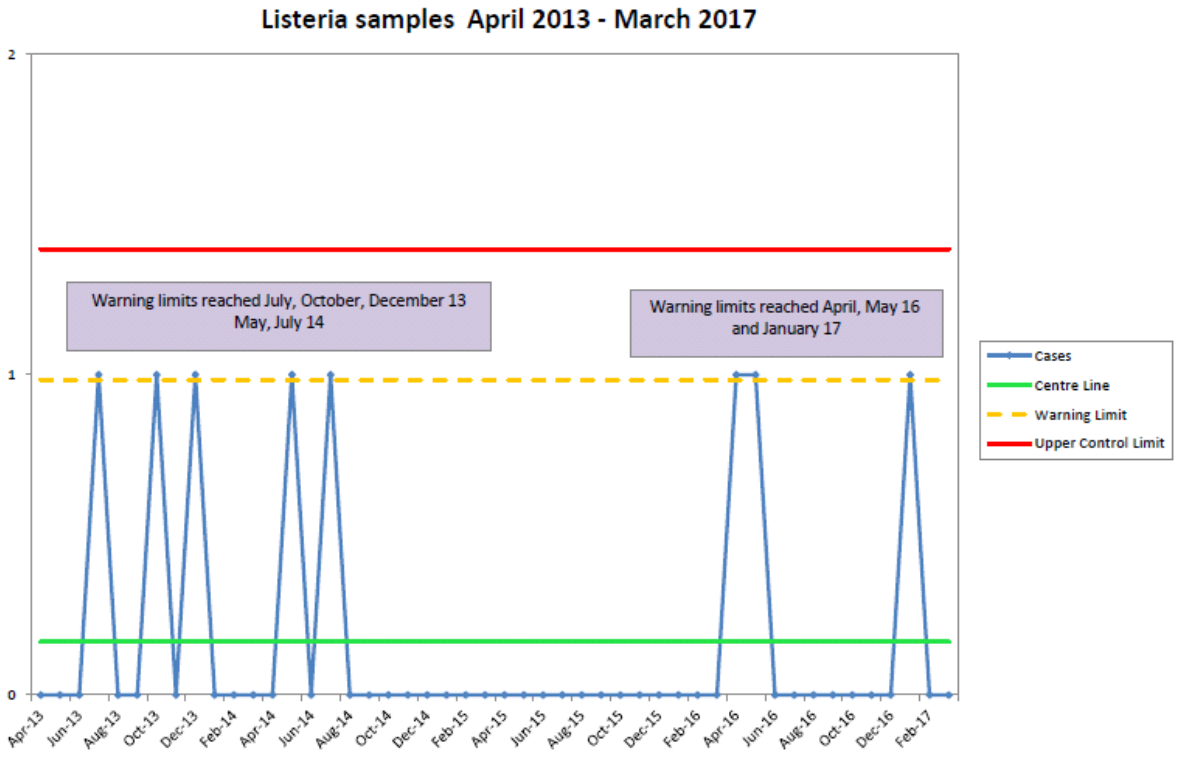


Figure 1.4. Example of a Statistical Process Control (SPC) c chart displaying number of *Listeria* cases from April 2013 to March 2017. Produced by NHS Grampian IPC department using data from ICNet outbreak surveillance software.

1.6 Investigation & management of outbreaks

The application of "infectious disease epidemiology" often involves retrospective investigation of transmission events in outbreak investigation, and epidemiological tracing.(66)(67) In these events transmission routes are assessed by taking contact histories, defining risk factors, creating case definitions. and determining key events such as exposure.(66)(67) Published examples of the investigation of outbreaks in NHS Tayside and NHS Grampian include a *Neisseria meningitidis* outbreak among Scouts returning from the world Scout Jamboree, *P.aeruginosa* outbreak in ICU, *Bacillus anthracis* infection among persons who inject drugs, prison and community outbreak of severe respiratory infection due to adenovirus type 14p1 and an astrovirus type 5 gastroenteritis outbreaks in a residential elderly care home.(68)(69)(70)(71)

An early response is crucial to ensure that effective and targeted measures are in place to halt transmission. The CDC defines an epidemic is an "*increase, often sudden, in the number of cases of a disease above what is normally expected in that population in that area*".(48) They define that an outbreak: *"carried the same definition of epidemic, but is often used for a more limited geographic area*"; a cluster is an *"aggregation of cases grouped in place and time that are suspected to be greater than the number expected, even though the expected number may not be known"* and a pandemic is defined as *"an epidemic that has spread over several countries or continents, usually affecting a large number of people"*.(48) Health Protection Scotland (HPS) advises IPCTs to assume there is a HAI outbreak when there are "*two or more linked cases with the same infectious agent associated with the same healthcare setting over a specified time period*" or when there is a "*higher than expected number of cases of HAI in a given healthcare area over a specified time period*".(72) In these situations a trigger may have been breached on an SPC chart.(65) HPS define an exceptional infection episode as *"a single case of any serious illness which has major implications for others (patients, staff and/or visitors), the organisation or wider public health e.g. infectious diseases of high consequence such as VHF or XDR-TB"*.(72)(73) The approach to the investigation of a suspected outbreak may differ depending on various circumstances. In some instances it may be necessary to hold meetings referred to as Problem Assessment Group (PAG) or Incident Management Team (IMT) meetings.(72)(74) In NHS healthcare settings this will usually be chaired by an IPCD. If there are implications for the wider community, for example a *M. tuberculosis* outbreak the Consultant in Public Health Medicine (CPHM) may chair the meeting.(73) These meetings allow planning of IPC measures and discussion of any further testing required to identify patients. The membership of meetings will vary depending on the nature of the incident. Table 1.4 lists members of such a group.(75)

Table 1.4. Staff attending PAGs/IMTs

Key members	Representatives who may also be included if appropriate
Consultant Microbiologist	Environmental Health
Consultant in Public Health	Medical/Nursing staff from the affected ward
Infectious Diseases Consultant and/or Paediatrician	Occupational Health
Consultant or GP (community hospitals) in charge of the index patient	Medical/Nursing staff from the affected wards
Infection Prevention and Control Nurse	Pharmacy Team Lead
Infection Control Manager	Staff side
Representative from the Management team	Estates and facilities
Lead Nurse for the clinical area involved	Catering
Corporate Communications Officer	Any other department specified by the CPHM or IPCD

PAGs, problem assessment groups; IMTs, incident management team meetings; CPHM, Consultant in Public Health Medicine; IPCD, infection prevention and control doctor.

Actions following identification of a suspected outbreak can be seen in Figure 1.5 .(75)(74)(72)(52)(58)

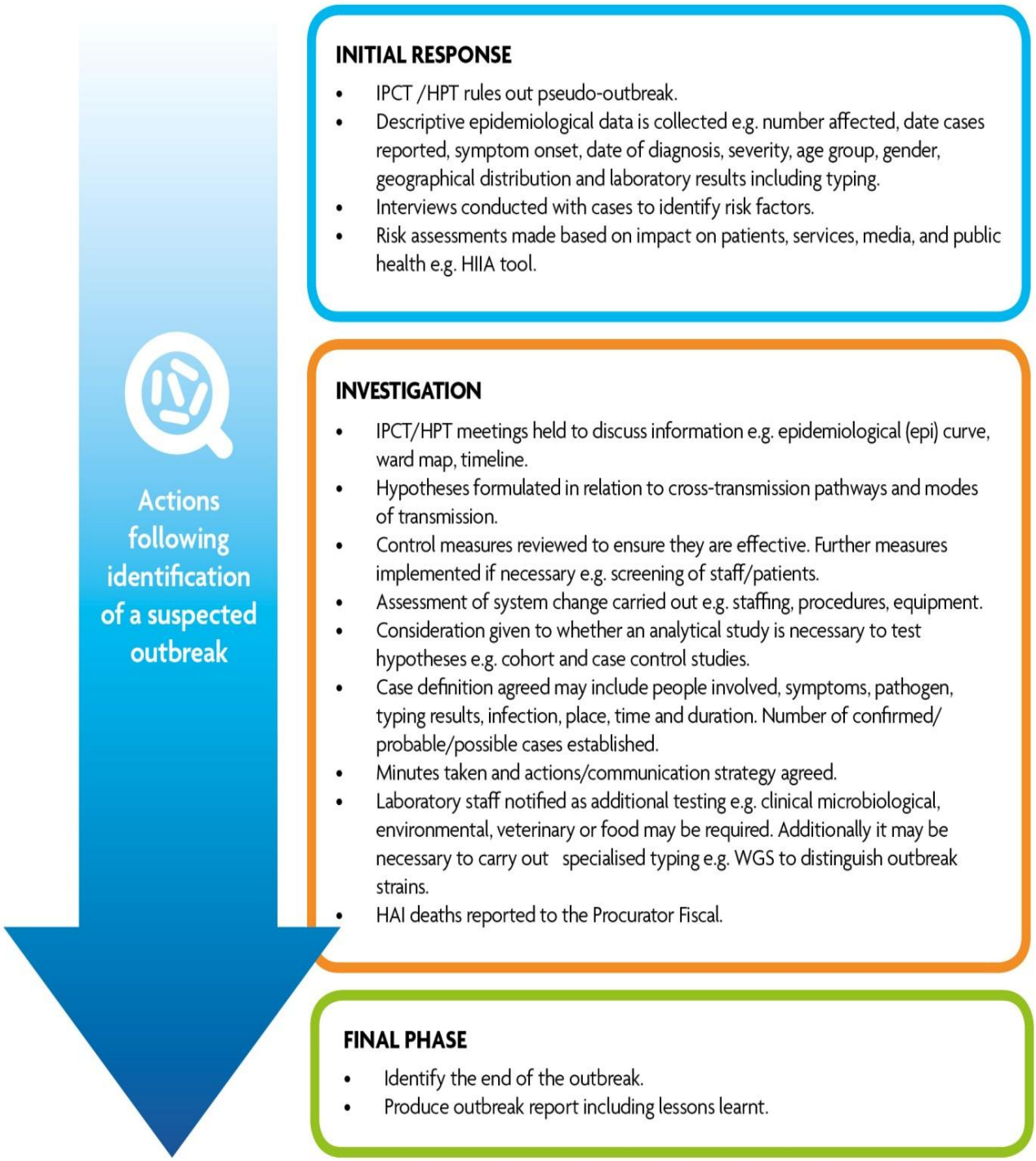


Figure 1.5. Chart showing infection prevention and control and public health actions taken following identification of a suspected outbreak at various stages. These include the initial response to the outbreak, investigation and final phase in which the end of the outbreak is identified.

Following collection of infectious disease epidemiological data, the investigators will also carry out a risk assessment of the situation. An impact assessment tool called the Hospital Infection Incident Assessment (HIIA) Tool is completed to assess impact on patients,

services, public health, and public anxiety. This will inform whether a press statement should be released.(76) This can be classified as green (manage within NHS board, inform Public Health consultant), amber (report to Scottish Government Health Directorate ask HPS for support if required) or red (report to Scottish Government Health Directorate, report to HPS, issue press statement). As part of the outbreak investigation the IPCT will agree a case definition and gather epidemiological information e.g. incidence rates and exposed populations. They will produce a line listing of patients, time line and/or an epidemic curve.(75) The group will consider implementing interventions to reduce the risk of transmission and development of illness e.g. antibiotics, immunisation, prophylaxis. Measures such as contact/droplet precautions may be instigated along with interventions such as special cleaning/disinfection to eliminate contamination of the environment. They may place restriction on visiting healthcare institutions and exclude individuals from work. In addition food preparation or retail premises may be closed if there is suspicion of inappropriate preparation of food.(75) Healthcare services such as wards may be closed to admissions if there is risk that further patients may acquire infection.

Epidemiological data is often accompanied by pathogen typing data.(67) Isolates from patients and/or food and environmental samples will be sent for typing at reference laboratories. Outbreak management hinges on Microbiology services providing results in good time. Pathogen typing data should be detailed enough to allow the discrimination of dissimilar strains detected from hosts therefore allowing transmissions events to be ruled out. It has been reported that "much greater levels of confidence may be achieved where pathogen sequence data are available".(67) Determining whether transmission has occurred is a major challenge as there may be issues establishing the epidemiological links between patients and the environment for instance, data such as which bay or bed a patient may have been in or which hospital they have been transferred from or too is not always recorded or accessible. Additionally, Reference laboratory typing results do not always have sufficient granularity or robustness to unequivocally define strains and confirm that transmission has taken place. These results are often not available until after the outbreak is over. This may be due to the long turnaround time of typing results due to multiple testing, batching of tests and transport times for samples. This can be detrimental to the outbreak investigation as a rapid response is essential to implement measures to prevent further transmission. WGS now has the capacity to allow outbreaks to be detected in real-time, enabling rapid implementation of targeted IPC measures. It has been used previously to investigate outbreaks of *M. tuberculosis, E. coli*

0104, MRSA, *N. meningitidis* and, *Legionella* but it has not it has not yet been fully applied operationally throughout clinical microbiology laboratories or reference laboratories in the UK.(15)(77)(78)(13)(14) A centralised service at Public Health England (PHE) Colindale delivers WGS for pathogens such as *E. coli*, *Shigella*, *Listeria*, *Campylobacter*, *S. aureus*, *Salmonella* and *Mycobacteria* and was one of the first laboratories to receive accreditation for this.(79) The Scottish Microbiology Reference Laboratories, Glasgow (SMiRL) established sequencing for *Salmonella* and *Shigella* in October 2017 and uses this method to identify somatic and flagellar antigen genes and multilocus sequence typing (MLST) sequence types.(80) Using this data serotype can be inferred. This data is important nationally and single nucleotide polymorphism (SNP) based genetic cluster analysis can also carried out to inform epidemiological investigations, such as the UK-wide *Salmonella enterica* serotype Enteritidis 25-SNP cluster t25.12 outbreak and national surveillance. (81) As a response to the coronavirus (COVID-19) pandemic sequencing was introduced nationally in collaboration with COVID-19 Genomics UK (COG-UK) Consortium.(82) As part of this, a combined phylogenetic and epidemiological approach was undertaken using Oxford Nanopore and Illumina MiSeq technology in Scotland. In the first wave of the pandemic, 1,314 SARS-CoV-2 viral genomes were sequenced and it was identified that SARS-CoV-2 was introduced into Scotland on at least 283 occasions during February and March in 2020. In combination with epidemiological data it was found that early introductions of SARS-CoV-2 mainly originated from Italy and Spain and that community transmission was occurring 3 weeks prior to the implementation of control measures which informed the advice that travel restrictions or quarantine measures should have been introduced earlier to control the transmission of SARS-CoV-2.(83)

1.7 Overview of typing methods

1.7.1 The role of the microbiology laboratory in outbreak detection

The advances in “infectious disease epidemiology” have been driven by the development of new techniques in the microbiology laboratory which can be divided into methods used for identification of pathogens or methods used to fingerprint deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) (Table 1.5).(6)

Table 1.5. Typing methods and examples of epidemiological studies

Applications	Method	Technique	Example of a study using this technique
Identification of pathogens	Conventional	Culture/isolating bacteria from animals	An investigation of a *Legionella* outbreak in 1976.(84)
		ELISA	Identification of a carrier by using Vi enzyme-linked immunosorbent assay serology in an outbreak of Typhoid fever on an Indian reservation in 1983.(85)
		EIA	Study describing a specific, sensitive and quantitative enzyme linked immunosorbent assay for human immunoglobulin G antibodies to Anthrax toxin protective antigen in 2001.(86)
		Monoclonal antibody typing	Study comparing molecular and antibody typing during the investigation of an outbreak of Legionnaires' disease in 2002.(87)
	Nucleic acid based	DNA hybridization for known genes	Outbreak of *Salmonella typhimurium* infection traced to contaminated chocolate and caused by a strain lacking the 60-megadalton virulence plasmid.(88)
		Direct sequencing of one or more regions	MLST plus sequencing of the *flaA* short variable region (SVR) for investigation of outbreaks of gastroenteritis caused by *Campylobacter jejuni f*rom 1981-1998.(89)
		MLST	Multi-country outbreak of *Listeria monocytogenes* investigated by PCR and MLST, 2017.(90)
	PCR based	PCR Assay	Detection of adenovirus outbreak at a municipal swimming pool by nested PCR amplification in 1998.(91)
		LCR	A pseudo-outbreak of *Chlamydia trachomatis* in a state residential facility in 1998.(92)
	Protein based	Western blot or immunoblotting	Evaluation of a western blot test in an outbreak of acute pulmonary histoplasmosis in 1999.(93)
		MALDI-TOF MS	Biotyping of multidrug-resistant *K. pneumoniae* clinical isolates from France and Algeria using MALDITOF MS 2008-2011.(94)
Fingerprinting	Conventional	Serotype	Study describing DNA fingerprinting and serotyping of *Campylobacter jejuni* isolates from epidemic outbreaks.(95)
		Antibiotic susceptibilities	A retrospective cohort study investigating the use of automated detection of infectious disease outbreaks in hospitals. Published in 2010.(96)
	Nucleic acid based	Plasmid profiles	An analysis of plasmid pattern in paediatric ICU outbreaks of nosocomial infection due to *Enterobacter cloacae* published in 1991.(93)
		RFLP	An epidemiological analysis of a MRSA outbreak using RFLP of genomic DNA.(97)
		PFGE	Use of PFGE to investigate an outbreak of *Serratia marcescens*. Published 1997.(98)
		RNA electrophoresis (electropherotyping)	Paper describing RNA-electrophoresis as a typing method for nosocomial rotavirus infection in a special-care baby unit.(99)
		rRNA gene restriction patterns (ribotyping)	Study examining rRNA gene restriction patterns as an epidemiological marker in nosocomial outbreaks of *S. aureus* infections. Published in 1993.(100)
		Direct sequencing of one or more regions	Paper describing routine WGS for outbreak investigations of *S. aureus* in a national reference centre in 2018.(101)
		MLST	A comparative analysis of core genome MLST and SNP typing within a European *Salmonella* serovar Enteritidis outbreak. Written in 2018.(102)
	PCR based	Amplification of a single target specific	Paper describing the development of a genomics-based PCR assay for the detection of *Mycoplasma pneumoniae* in

		to a pathogen	a large outbreak in New York State in 2001.(103)
		ERIC*	A study describing the use of this technique in an outbreak of carbapenem-resistant *A .baumannii* carrying the carbapenemase OXA-23 in ICU of the eastern Heilongjiang Province, China in 2019.(104)
		REP*	Paper describing a comparison of a REP-based PCR method and clinical and microbiological methods for determining strain sources in patients of nosocomial *A. baumannii* bacteraemia. Published in 2002.(105).
		DRE*	DRE PCR method for subtyping *M. tuberculosis* in clinical isolates. Published in 1995.(106)
		BOX*	BOX–PCR–based DNA analysis of non-serotypeable *Streptococcus pneumoniae* implicated in outbreaks of conjunctivitis. Published in 1997. (107)
		IS*	DNA fingerprinting of *Vibrio cholerae* strains with a novel insertion sequence element. Published in 1996.(108)
		PGRS*	Use of PGRS for the identification of a W variant outbreak of *M. tuberculosis* via population-based molecular epidemiology .Published in 1999.(109)
		RAPD	A paper in which RAPD PCR was used for the comparison of *Mycobacterium abscessus* strains from nosocomial outbreaks. Published in 1997.(110)
		AP-PCR	Study looking at the role of AP- PCR in identifying the source of an outbreak of Legionnaires' disease in 1997.(111)
		AFLP	An evaluation of AFLP as a tool for molecular subtyping of enterohemorrhagic *E. coli* O157:H7 isolates. Published in 1999.(112)
	Protein based	MLEE	A paper employing this technique for the analysis of rapidly-growing mycobacteria. Published in 2016.(113)
	Gene expression	Reverse transcriptase PCR	Study describing rapid diagnosis of Ebola hemorrhagic fever by reverse transcription-PCR in an outbreak setting and assessment of patient viral load as a predictor of outcome. Published in 2004.(114)
		Microarray technologies	Study examining high density microarray analysis to gain new insights into genetic footprints of *Listeria monocytogenes* strains involved in listeriosis outbreaks. Published in 2012.(115)

ELISA, Enzyme-linked immunosorbant assay; EIA, enzyme immunosorbant assay; MLST, multilocus sequence typing; PCR, polymerase chain reaction; LCR, ligase chain reaction; MALDI-TOF MS, matrix-assisted laser desorption ionization–time of flight; RFLP, restriction fragment length polymorphism, PFGE, pulsed field gel electrophoresis; MLST, multilocus sequence typing; ERIC, enterobacterial repetitive intergenic consensus sequences; REP, Repetitive extragenic palindromic sequences; DRE double repetitive element; IS, insertional sequences; PGRS, polymorphic guanine/cytosine-rich repetitive sequences; RAPD, random primers (randomly amplified polymorphic DNA; AP-PCR, arbitrary primed PCR; AFLP, amplified fragment length polymorphism; MLEE, multilocus enzyme electrophoresis**, *** Targeting known repetitive sequences.

1.7.2 Phenotypic typing methods

Identification of the pathogen is one of the most important first steps in the investigation of outbreaks.(116) The development of agar was one of the first laboratory techniques that made "infectious disease epidemiology" possible leading to development of phenotypic typing methods. Culture was first applied by Walther Hesse a physician whilst he worked as a student with Robert Koch from 1881-1882 revolutionising the "germ theory of disease" and the approach by public health to the identification, control and prevention of infectious disease.(6) Phenotypic methods such as growth characteristics, morphology, biochemical properties, antibiotic susceptibility, serologic properties and functional/ physiological properties can be used to define outbreak strains.(58)(117) Table 1.6 shows an overview of the phenotypic characteristics of bacteria (adapted from Riley LW, 2018).(116)

Table 1.6. Phenotypic characteristics of bacteria

Characteristics		Example of microorganism
Growth behaviour	Nutrient requirement	Nutritionally variant streptococci (NVS) requiring either pyridoxal or L-cysteine for growth.
	Growth in presence of absence of oxygen Requirement of light or darkness	*Rhodopseudomonas palustris* - photosynthetic, anaerobic, red bacterium. Also grows aerobically in the dark.(118)
	Temperature	*Streptococcus thermophilus* Grows at 50 ° C.
	Slow or rapid growth	Rapid growing mycobacteria (*Mycobacterium. chelonae*, *abscessus or fortuitum*)
Morphology	Colony colour, shape, smell, staining pattern e.g., Gram stain, shape observed under microscope e.g. rod coccus, spirochete	*Rhodococcus equi* salmon-pink colonies *Streptococcus* millleri- butterscotch smell due to diacetyl.(118) *Staphylococcus* -Gram positive clusters
Biochemical properties	Fermentation, gas production, substrate utilization, mass spectrometry profile	*Streptococcus thermophilus* ferments glucose and lactose, producing lactic acid; used in the production of yoghurt. *Acetobacter aceti* oxidises ethanol to carbon dioxide and water. (118)
Serologic properties	Recognition by antibodies to O polysaccharide (O antigen), flagella protein H (H antigen), capsular polysaccharide (K antigen)	Salmonella
Functional/ physiological	Virulence factors (cell attachment, invasion, cytotoxicity, toxin production, pathogen-associated molecular patterns, disease production in an animal model, antimicrobial susceptibility, host effector molecules, (antibodies, reactive oxygen, nitrogen intermediates, antimicrobial peptides, granzymes/phages, host subversion and evasion mechanisms, infectious inoculum dose	*Corynebacterium diphtheria* production of black colonies on Hoyles tellurite and toxin testing using ELEK.

If a phenotype is rare and epidemiology suggests patients are involved in transmission it may be enough to convince the IPCTs there has been spread e.g. *E. coli* 0157.(58) However, isolates can share phenotypic markers but be genotypically different.(58) Bacterial characteristics are not always stable and may be affected by environmental selective pressures and resistance can be altered due to selective pressure of antibiotic use. Antigenic traits can also be altered by random mutation.(58) Phenotypic testing methods can involve high costs in labour and materials and may be restricted to certain microorganisms e.g. antisera for *Salmonella* serotyping.(119) Individual phenotypic methods will be discussed in further detail in the following paragraphs.

1.7.3 Biotyping

A biological profile of an organism is termed a biogram.(117) Determining the relatedness of bacteria on the basis of their biograms is termed biotyping.(117) Biotyping is a phenotypic method that is widely used and differentiates stains by looking at their biochemical reactions, morphology, and growth. It is now performed in laboratories using automated systems such as Vitek 2 and Matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS).(119) Biograms are not stable and can be affected by variation in gene expression, random mutations, loss and gain of plasmids and technical manipulation. Biograms can also be limited as a typing technique where biochemical diversity is uncommon for example enterococci.(119)

1.7.4 Antibiotic susceptibility patterns

To detect outbreaks microbiology staff will look at the antibiotic susceptibility profiles (antibiograms) of bacteria grown from specimens and compare them. In practice many Microbiologists may regard microorganisms not to be related if they differ by more than 2 antibiotic susceptibilities. Antibiograms are different to resitograms which involves susceptibility profiling of bacteria to dyes and heavy metals. The benefits of antimicrobial susceptibility testing include low cost, ease of use and availability as they are already required to inform antibiotic treatment for patients. Susceptibility testing can be performed by using automated broth microdilution or disk diffusion techniques.(119) Microdilution testing is used more commonly as it can be automated and provides quantitative measures of

the minimum inhibitory concentration (MIC) – the lowest concentration of antibiotic that inhibits growth of bacteria.(120) Both methods are standardised to be reproducible. The antibiotic susceptibility pattern is recognised to lack sensitivity and specificity as a typing method as bacteria that are not genetically and epidemiologically related can have the same susceptibility pattern and *vice versa*.(119) Antibiograms have become less sensitive as a method to detect outbreaks as multidrug resistance (MDR) has increased.(58) Antibiograms are not constant and can change due to selective pressure from antimicrobial therapy. In some instances antibiograms can change due to chromosomal point mutations for example fluroquinolones or from loss or gain of extrachromosomal DNA in the form of mobile genetic elements e.g. transposons and plasmids.(117)(121) Unrelated isolates can have identical resistance profiles and in some instances this may reflect acquisition of a plasmid by multiple species i.e. a "plasmid outbreak".(121)

1.7.5 Serotyping

Serotyping involves utilising a collection of antibodies to identify certain antigens on the surface of bacteria.(119) It can be used for taxonomic grouping of a range of bacterial pathogens for example *Streptococcus pneumoniae*, *Legionella*, *Salmonella* and *Shigella*.(119) Direct antibody-antigen agglutination tests can be used and as part of this a suspension of bacterial cells is mixed with different antibodies. The serotype can be determined by the agglutination profile. *E. coli* use to be traditionally typed using this method examining antibody reactions to O- polysaccharide antigens, flagellar H-antigens, and capsular K-antigens.(122) In the past, the Kauffmann White Scheme (KWS) was considered to be the gold standard for serotyping of *Salmonella*.(123) In the investigation of *S. pneumoniae* the Quellung test can be performed in which antibodies bind to capsular antigen and the capsule swells which can be seen using microscopy.(119) Serotyping is limited as it is considered to be labour-intensive, costly, and time consuming. Some pathogens cannot be typed this way and there is a lack of standardization of methods across laboratories.(122)

1.7.6 Bacteriophage

Bacteriophages (phages) are viruses that can infect and lyse bacteria.(120) Typing involves using sets of phage's to assess the pattern of resistance or susceptibility of bacteria to phages.(119) Resistance depends on whether there are receptors on the cells surface that the phage can bind to. Bacteriophage typing has been used in the past for *S.aureus* and *Salmonella, P. aeruginosa* and *L. monocytogenes.*(119)(124)(58) Serotyping and bacteriophage typing can both be labour intensive and it has become increasingly more difficult to meet quality standards expected by accreditation bodies using these techniques for instance stocks of control strains and phages need to be stored and maintained.(124)(58)

1.7.7 Bacteriocin

Bacteria can also be typed by looking at their susceptibility or inhibition of growth on agar to substances such as proteins produced by other bacteria (bacteriocins) (119) Bacteriocins are lethal to other members of the same species and in some instances other species. The majority of methods utilise the principle that a strain (indicator) will be unable to grow on the surface of an agar plate that has grown a bacteriocin producer strain.(120) Bacteriocin and phage typing systems cannot be applied to all bacterial species.(125) Howard *et al.* describes this technique as being simple and powerful and the simplest methods only require an isolate collection that can be tested against each other for bacteriocin sensitivity or production. In some instances it could identify a unique strain in an outbreak situation but Howard *et al.* warns that the information gained from this technique is less valuable than routine typing methods.(120)

1.8 Molecular typing techniques in epidemiological studies

When describing molecular epidemiology, Foxman *et al.* defined "molecular" typing as *"the use of the techniques of molecular biology to characterize nucleic acid- or amino acid-based content"*.(126) One of the first papers to describe the application of molecular typing to the investigation of HAI outbreaks of infection was published in 1978.(127) Samples were taken from burn unit patients that isolated tobramycin-resistant *K. pneumoniae* and *Enterobacter cloacae* and the isolates were typed using restriction endonuclease digestion patterns and

DNA-DN hybridization studies.(127) Since then, molecular typing has been a fast moving field with developments in techniques such as pulsed field gel electrophoresis (PFGE), plasmid analysis, polymerase chain reaction (PCR) and restriction based methods.(119) There are seven main epidemiology problems that molecular strain typing can be applied to investigate. Figure 1.6 shows an overview of these (adapted from Riley LW, 2004).(6)

Figure 1.6. Applications of molecular strain typing methods. This figure shows the many uses of typing in addition to identifying transmission dynamics (adapted from Riley LW, 2004).

Molecular typing has become a routine in outbreak investigations as it provides more sensitive and specific measurements compared to phenotypic typing. (126) A bibliographic search (carried out 19/9/2019) of Pubmed using the keywords “molecular epidemiology” found 118, 818 articles. Figure 1.7 shows the proportion of citations in PubMed using the search term “Molecular epidemiology” by year from 1975-2018 (created using PubMed by Year: http://esperr.github.io/pubmed-by-year). It can be seen that there has been a dramatic increase in publications involving molecular epidemiology since 1989 although reports were beginning to reduce from 2013-2017.

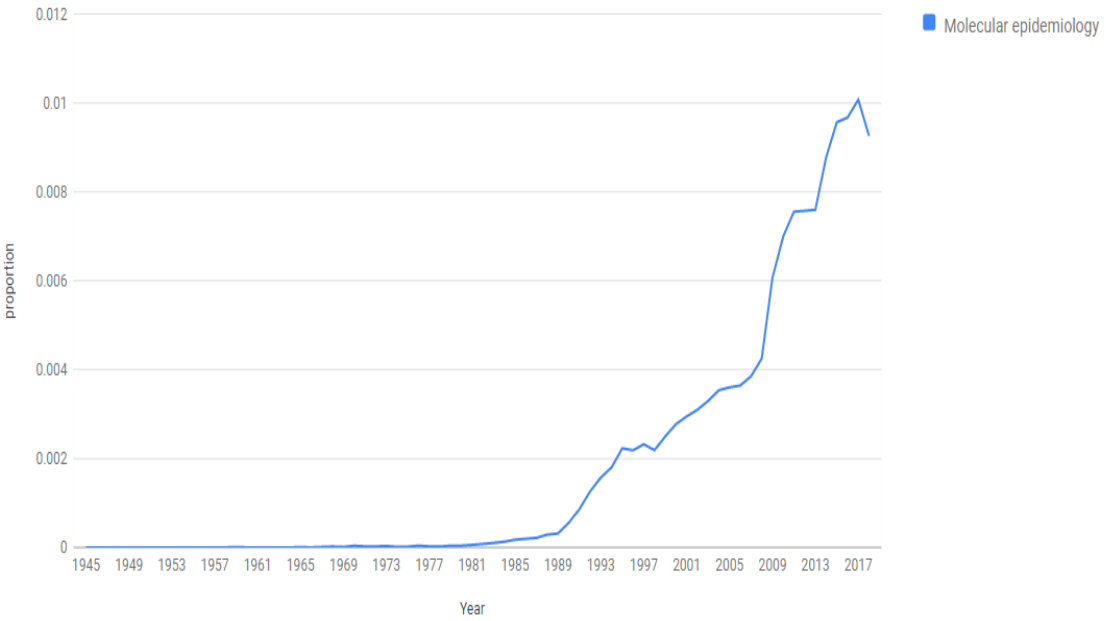


Figure 1.7. Graph showing the proportion of citations in PubMed using the search term “Molecular epidemiology” by year from 1975-2018 (created using PubMed by Year: http://esperr.github.io/pubmed-by-year).

A similar search was carried out to define the proportion of citations in PubMed using various molecular typing techniques. The term “outbreak” was searched (19/09/2019) from 1975-2018 (using PubMed by Year: http://esperr.github.io/pubmed-by-year) and Figure 1.8 was created. From 1981 -1990 it can be seen that plasmid profile analysis (PPA) was a common molecular fingerprinting technique however it became an outdated method after this and replaced by other methods of typing. (126) A rapid rise in PCR was observed from 1990 to 2011. This subsequently decreased whilst publications describing the use of WGS increased.

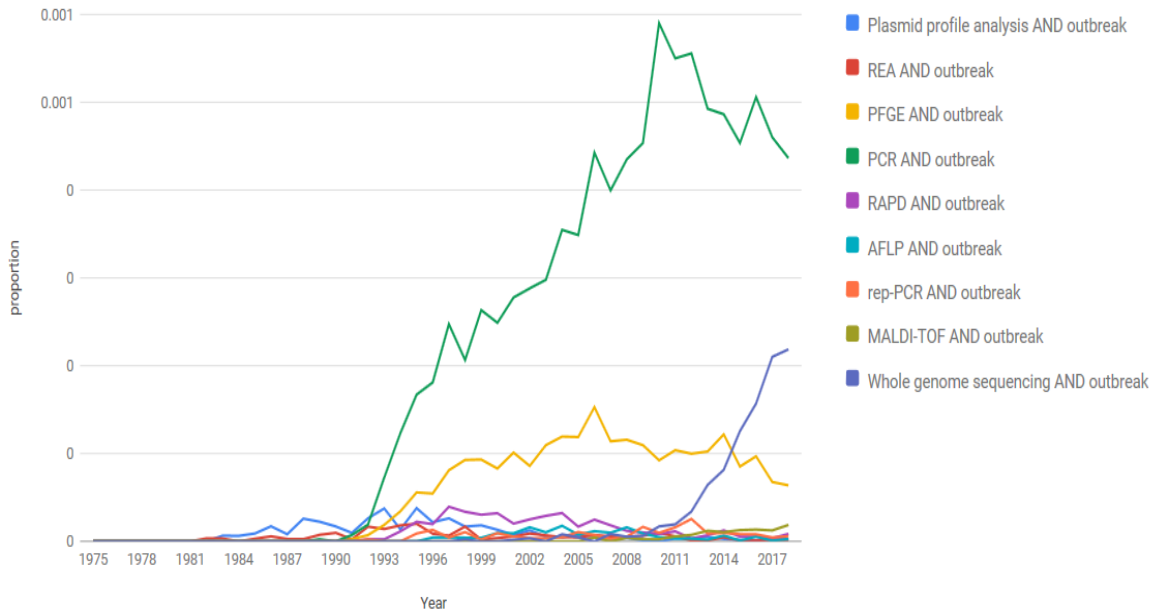


Figure 1.8. Proportion of citations in PubMed describing various outbreak molecular typing techniques from 1975-2018. The term "outbreak" was searched along with various separate typing techniques (19/09/2019) from 1975-2018 (using PubMed by Year: http://esperr.github.io/pubmed-by-year)

Molecular typing schemes are required to have methodologies that are standardized, sensitive specific and objective.(119) Typing systems can be categorised by their reproducibility typeability, discriminatory power, cost, along with how easy the methodology is to perform in terms of technical ease and interpretation of results.(119) Typeability is the capacity of a method to type (give an unambiguous result) to isolates. Reproducibility represents a techniques capability to give the same result when an isolate is re-tested and poor reproducibility can occur due to technical variation or biologic variation. Discriminatory power represents a methods capacity to distinguish isolates that are epidemiologically unrelated. Cost can be considered in terms of materials, equipment and staff time for carrying out the technique and interpreting results.(119) Table 1.7 shows an overview of discriminatory power, reproducibility, repeatability, cost and time for molecular typing techniques (adapted from Foxman *et al*. 2005).(126)

Table 1.7. Overview of discriminatory power, repeatability, reproducibility, and cost for molecular typing techniques

Typing technique	Discriminatory power	Repeatability	Reproducibility	Cost
WGS	H	H	H	H
Microarrays (comparative hybridisation against array containing gene sequence)	H	M-H	M-H	H
Sequencing of one or more genetic region	M-H	H	H	M-H
MLST	M-H	H	H	M-H
Binary typing (presence/absence of selected genes/alleles across genome	M-H	H	Potentially H	M
PFGE	M-H (depends on bands and type/number of enzymes selected)	M-H (depends on species)	M-H	H
Amplification of single target gene of a pathogen	M-H	H	M-H	L-M
AFLP	M-H	H	M-H	L-M
Automated ribotyping	M	H	H	H
Ribosomal RNA gel electrophoresis	M	H	H	L
ERIC, REP, BOX,	L-M	M	L	L-M
RAPD, AP-PCR	L-M	L	L	L-M
Restriction endonuclease on a single amplified product	L-M	H	H	L-M
Plasmid profiles	L	H	M	L

H, High; M, Medium; L, Low; ERIC, enterobacterial repetitive intergenic consensus sequences; REP, repetitive extragenic palindromic sequences; RAPD, random amplification of polymorphic DNA; AP-PCR, arbitrarily primed PCR.

Molecular methods can usually be applied to a variety of species compared to phenotypic methods which tend to be restricted to certain species.(119) One of the first papers focusing on the cost effectiveness of introducing molecular typing for outbreak investigations described how infections per 1,000 patient days fell by 13 per cent (p=0.002) and patients hospitalised with HAI reduced by 23 per cent (p=0.000006) after its introduction. Savings of $5.00 for each dollar spent on the $400,000 per year programme were made. (7) Cost savings can be made by halting the spread of infection and conversely by ruling out outbreaks preventing the need for costly investigations and the utilisation of resources such as increased cleaning and use of PPE. An overview of the procedural steps required for common molecular typing techniques is shown in Figure 1.9 (adapted from Singh A *et al.*, 2006).(119) The rest of this chapter will describe molecular methods in detail.

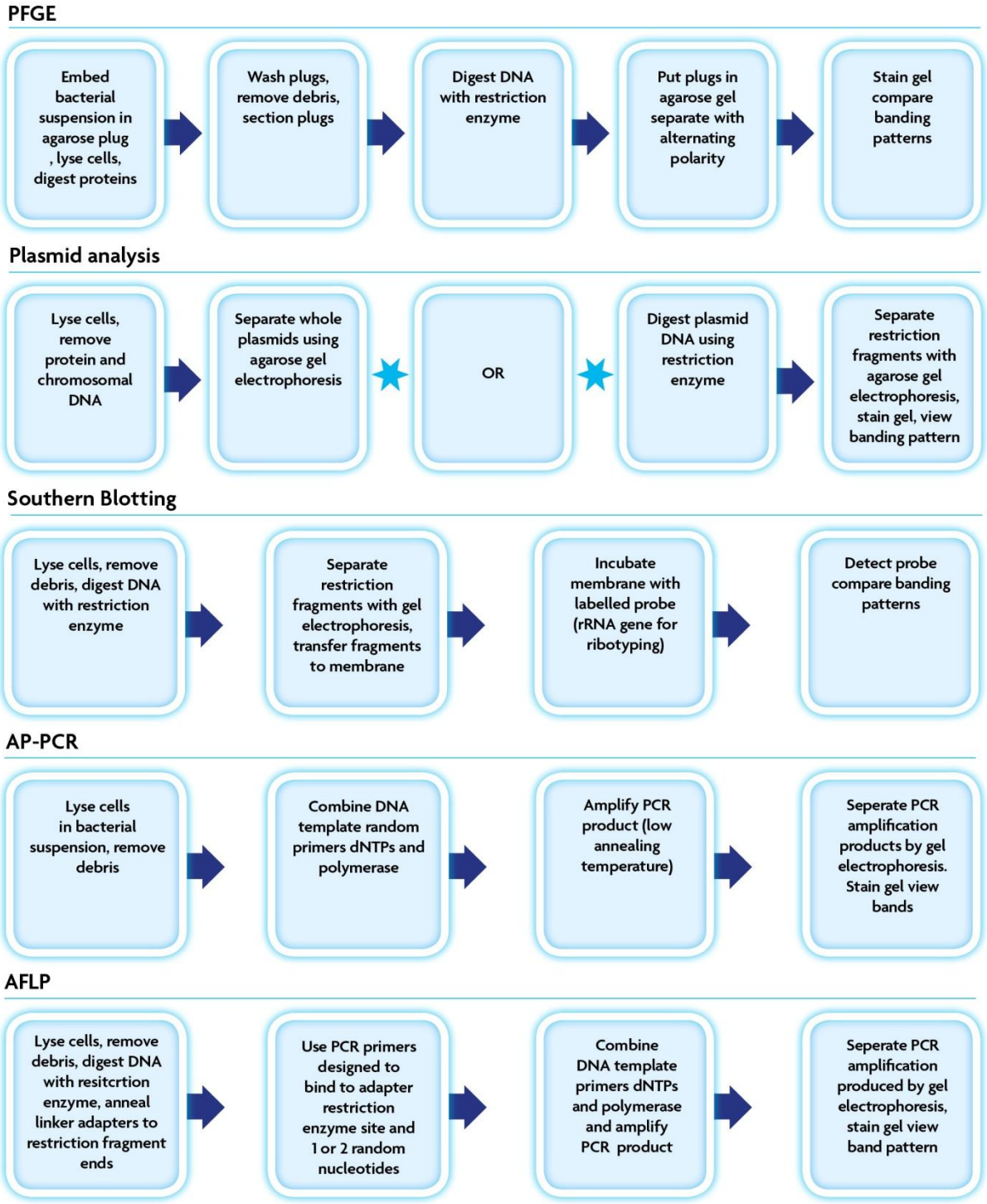


Figure 1.9. Overview of procedural steps required for common molecular typing techniques (adapted from Singh A *et al.*, 2006).

1.8.1 Plasmid profile analysis (PPA)

Plasmid profile analysis (PPA) also known as plasmid fingerprinting was one of the first genotypic methods to be used in outbreaks.(58)(121) Woodford *et al.* described that its use was most popular in the 1980s.(128) Plasmids are self-replicating extra-chromosomal molecules of DNA elements capable of autonomous replication that may be transferrable and are found in the prokaryote cytoplasm.(119)(120) Plasmids can code for resistance to antibiotics, drugs or heavy metals or for toxin or siderophore production. Other plasmids that do not carry these are called cryptic plasmids.(120) For PPA plasmids can by identified by a simple cell lysis procedure using detergents at high pH and subsequent removal and separation of chromosomal and plasmid DNA with agarose gel electrophoresis. The agarose gel is then stained and photographed.(121) After this, the number and size of plasmids can be compared to inform strain identification.(121)(119). PPA can be used to identify outbreak strains and follow spread of resistance genes. If plasmids are unique to a strain this technique can be used to determine that an organism is causing the epidemic as long as the plasmid has been stable for some time and there is no environmental stress. However it should be noted that plasmids can spread rapidly amongst strains and different species.(119) Plasmids can spread by conjugation and it may be that the plasmid is epidemic rather than the bacterial strain.(58) PPA has been used in the investigation of various outbreaks and plasmid profiling has been shown to have better discriminatory power compared to antibiograms and phage typing during *Salmonella* outbreaks.(58) Advantages of PPA include its application to the investigation of various microorganisms, low turnaround time, high reproducibility and low costs. Multiple cultures can also be processed at once.(58) Disadvantages to PPA include the possibility of plasmids being deleted and consequently epidemiologically related isolates may contain different plasmid profiles.(99) PPA is limited by plasmids having certain molecular forms for example supercoiled (closed circle) nicked (open circle) and linear and each type appears different by gel electrophoresis.(58) The discriminatory power of plasmid profiling is decreased with fewer plasmids and PPA is therefore not always the best method for determining the delineation of strain relatedness.(119)

1.8.2 Restriction endonuclease analysis (REA)

Restriction endonuclease analysis (REA) was one of the first restriction fragment length polymorphism (RFLP) methods to be developed.(129)(130) This method uses an enzyme that cuts doubled stranded DNA at specific sites within plasmids or chromosomal DNA by recognising specific nucleotide sequences.(131) Fragments of DNA are then separated by size in agarose gel electrophoresis which result in a restriction endonuclease profiles called RFLPs.(119)(58) The gel is stained with dyes such as ethidium bromide and photographed under ultraviolet light.(58) Despite REA being reproducible with high discrimination power it was replaced by PCR ribotyping for *C. difficile* typing as REA methodology and interpretation were demanding and the data could not be shared easily between laboratories.(130)

1.8.3 Restriction endonuclease analysis (REA) of plasmid DNA

For bacteria that only contain a single large plasmid (size range 100 to 150 kb) an extra step involving restriction endonuclease digestion was added to improve the discriminatory power of the agarose gel electrophoresis.(121) This produces a restriction endonuclease prolife which assists in detecting small differences. If plasmids are the same size with different fragment patterns on REA they could be considered to be different strains. This approach using REA of plasmid DNA has been used in the investigation of VRE and MRSA outbreaks.(58) This technique was rarely used to type Gram negative rods as large plasmids can produce many restriction fragments making it difficult to interpret results. Previously this technique was more commonly used for staphylococci as plasmids are usually <50 kb. It has been reported that it was simple to carry out and inexpensive.(121)

1.8.4 Restriction fragment length polymorphism (RFLP) analysis

RFLP produced by restriction endonucleases can be used for *Mycobacterium* typing during outbreaks and to determine reinfection.(58) Restriction endonuclease analysis is reproducible and accurate at identifying the relatedness of microbial strains. Downsides include that RFLP is expensive, laborious and time consuming and a limitation of using this technique for *Mycobacterium* is that 6 weeks or more of isolate growth is required in order to have enough DNA for analysis. This was overcome by combining RFLP with PCR. In this method the

DNA fragments were put onto a nylon or nitrocellulose membrane (southern blotting).(119). The DNA was then hybridized on the membrane with radioactive or chemically labelled DNA or RNA (probe) binding to its complementary nucleic acid sequences. The various fragments produced were called RFLPs and were recognised as they had different sizes and numbers.(121)

1.8.5 Southern Blotting

One of the commonest targets of southern blotting is the rRNA gene. Targeting this gene is called ribotyping.(119) Ribotyping is one of the commonest typing methods that use chromosomal DNA and ribosomal RNA probes. Practically all bacteria can be typed using probes that detect DNA sequences as this give rise to rRNA loci by using single rRNA probes. This is due to all bacteria having one or more chromosomal rRNA operons in their chromosome and these sequences are high conserved.(121) The discriminatory power of ribotyping is less than PFGE but an advantage is that it can become an automated process.(119) To assist in making results easier to interpret a combination of RFLP and ribotyping can be carried out. As part of this, the genome is digested and a second step is added in involving an rRNA gene-complementary probe to fragments of the genome.(129) Restriction fragment length polymorphism using IS6110 a DNA insertion element was recently the gold standard for typing *M. tuberculosis*.(121) IS6110 is present in nearly all *M. tuberculosis* isolates. Insertion element DNA can move independently inserting into different locations in plasmids and chromosomes however it does not carry antibiotic resistance genes or genes for pathogenesis. As it is mobile is can insert into different locations numerous times and this varies from strain to strain. If isolates had five or more copies of IS6110 it can be discriminatory and identify clustered patients of tuberculosis.(119)

1.8.6 Pulsed-field gel electrophoresis (PFGE)

Pulse field gel electrophoresis (PFGE) was developed to allow large DNA fragments to be separated.(117) PFGE was first described in 1984.(121) There are several types of PFGE and all utilise the principle of using an electric field.(120) It has been the gold standard typing method for many organisms and outbreaks.(132) It has been used to investigate at least 40 pathogens including Gram positive and negative organisms.(121) All species are

potentially typeable by PFGE but in some circumstances DNA can degrade during cell lysis.(121). The bacterial genome is digested using a restriction enzyme that will generate 10 to 30 restriction fragments which range from 10 to 800kb.(121) PFGE separates large DNA fragments by periodically changing the direction of the electric field in which DNA is being separated. There are two approaches to PFGE. In conventional electrophoresis an electric current is applied to the gel in a single direction. Contour clamped homogenous electric field (CHEF) uses a complex electrophoresis chamber and multiple electrodes switches the electric field at 120 degree angles. Field inversion gel electrophoresis uses a conventional electrophoresis chamber and the electric field is inverted periodically by 180 degrees. The electric fields created by the pulses allow the DNA to move in the gel resulting in a pattern of distinct bands.(121) PFGE is one of the most reproducible and highly discriminatory techniques but disadvantages of this technique include high cost of equipment and a delay of 1-3 days to prepare suitable DNA. Outbreak strains are usually indistinguishable but random genetic events such as point mutations or insertions and deletions of DNA caused by plasmids bacteriophages or insertion sequences can occur potentially altering DNA patterns.(133) Criteria for interpreting PFGE developed by Tenover *et al.* can be seen in Table 1.8 (adapted from Tenover FC *et al.*, 1995).(134).

Table 1.8. Criteria for interpreting PFGE patterns

Typical no. of fragment differences compared with outbreak pattern	No. of genetic differences compared with outbreak strain	Category	Epidemiologic interpretation
0	0	Indistinguishable	Isolate is part of the outbreak
2-3	1	Closely related	Isolate is probably part of the outbreak
4-6	2	Possibly related	Isolate is possibly part of the outbreak
Greater or equal to 7	Greater or equal to 3	Different	Isolate is not part of the outbreak

As part of Tenovers banding pattern criteria difference of three fragments may be due to a single genetic event and these would be considered to be highly related isolates which could represent epidemiologically related subtypes of the same strain. Differences of four to six restriction fragments could be due to two genetic events and are less likely to by epidemiologically related. In this event it may not be clear how related the organisms are. They may be related if collected over a long time period for example 3 to 6 months but it may

be that they are actually unrelated.(134)(133) Software is available for analysis of PFGE patterns for example BioNumerics.(133) This can be used store data and read patterns over multiple gels. This software assumes strains are identical if they have 100% similarity and are clonally related if they have more than 80% similarity. BioNumerics algorithms can be used to produce dendrograms that show strain lineages, evolution and ancestral relationships.(133)

1.8.7 Polymerase chain reaction (PCR)

PCR has been used for directly detecting pathogens in clinical samples for a number of years and can also be used as a typing tool.(121) Large quantities of nucleic acid sequence are synthesized and the procedure requires template DNA or RNA from the organism to be present. If RNA is used a reverse transcriptase step is required.(121) Two oligonuceotide primers are used, these flank sequences on the template DNA so that it can be amplified. The primers mark the starting point for the heat stable DNA polymerases to work adding bases to produce a complementary strand to the template. PCR takes approximately 3 hours to carry out 30 cycles.(133) Each amplification cycle involves a heat denaturation phase where double stranded DNA melts to single stands. There is an annealing phase where primers bind to target sequences in the single strands and an extension phase where DNA synthesis occurs as the copy number of DNA is doubled.(133) After 30 cycles a single copy of template DNA can be amplified to 1 billion copies.(121)

1.8.8 Random amplification of polymorphic DNA (RAPD)

Random amplification of polymorphic DNA (RAPD) involves parallel amplification of fragments by primers (short arbitrary sequences of approximately 10 bases). These primers target random genomic sequences initiating DNA synthesis at a low annealing temperature by producing amplicons.(132)(121) This can be used as a typing technique as the position and number of primer binding sites will be different for bacterial strains. RAPD amplicons are analysed using agarose gel electrophoresis or DNA sequencing. RAPD is less discriminatory than PFGE but is inexpensive, rapid, and technically undemanding. It has a low reproducibility as changes in reagents such as pH, ionic strength of buffers, temperature and equipment can affect results.(132) Amplifications of the same strain can be different and the best results are produced when isolates are tested together in a single amplification reaction

therefore comparing results from different centres can be problematic.(121) Arbitrarily primed PCR (AP-PCR) is a variant of RAPD and differs as AP-PCR involves amplification which is carried out in three parts.(132) It has been used for *C. difficile* and *S. aureus* typing.

1.8.9 Amplified fragment length polymorphism (AFLP)

Vos *et al.* first described amplified fragment length polymorphism (AFLP) in 1995.(135) It is a PCR-based fingerprinting technique that also uses restriction enzyme digestion.(136) AFLP utilises the advantages of RFLP and PCR to produce reproducible and easy to read profiles. It is sensitive for detecting DNA sequence polymorphisms.(136) DNA is cut by two restriction endonucleases generating a large number of fragments. Adapter sequences are linked to the ends of restriction fragments extending the length of the sequences. PCR primers hybridize to the adapter sequence, restriction site sequence and nucleotide of the unknown template sequence. Adding a randomly selected nucleotide to the primer reduces fragments that are amplified by a factor of four. After this PCR gel electrophoresis is used to separate the products and programmes such as BioNumerics can be used to analyse banding patterns.(119)

1.8.10 Repetitive element polymerase chain reaction (rep-PCR)

Repetitive DNA sequences can be scattered throughout the genome of many bacteria. PCR primers can be created to amplify DNA between repetitive elements that are in close proximity.(119). Repetitive element PCR (rep-PCR) utilises primers that can hybridise to non-coding intergenic repetitive sequences that are positioned throughout the genome. The DNA positioned between repetitive elements is amplified by PCR producing multiple amplicons. The amplicon sizes are then electrophoretically processed and banding patterns given. Results can be produced quickly and cheaply. It can be highly discriminatory but can lack reproducibility due to changes in reagents and gel electrophoresis. Some examples of repeat sequences that can be used for rep-PCR typing are repetitive extragenic palindromic (REP), enterobacterial repetitive intergenic consensus (ERIC) and 'BOX' sequences. DiversiLAb system (biomerieux) is a semi-automated method based on rep-PCR technology.(132) Rep- PCR has been used for *S. aureus, E. coli* and enterococci typing.(119)

1.8.11 Variable-number tandem repeat (VNTR) typing

The genomes of bacteria can contain regions that have nucleotide repeats as part of non-coding and coding DNA sequences. When repeats are adjacent to each other and numbers of these are different at the same locus these regions are called variable number tandem repeat (VNTR) loci and these can be used for typing.(132) Multilocus VNTR analysis (MLVA) determines tandem repeat sequences by amplifying VNTR loci by multiplex PCR and analysing amplicons on agarose gels. It can be cheap, easy and fast to carry out.(132) Fluorescently labelled PCR primers are made to amplify the repeat region and PCR products can be separated by an automated sequencer by using fluorescently labelled primers.(119)

1.8.12 Multilocus sequence typing (MLST)

Multilocus sequence typing (MLST) is another common typing technique which involves sequencing regions of 400-500bp from multiple (usually seven genes). It is highly reproducible, standardized and easily comparable between laboratories and central MLST databases but is also considered by some to lack discriminatory power for epidemiological surveillance.(137)(138) MLST has been used for *S. pneumonia, S. aureus*, Group A *Streptococcus* (GAS), *Enterococcus faecium, Haemophilus influenzae, N. meningitidis,* and *Campylobacter jejuni* typing. It can be carried out directly on clinical samples and can type bacteria when they have failed to grow.(138)

1.8.13 Single locus sequence typing (SLST)

Single locus sequence typing (SLST) is based on comparing sequence variations in a single target gene.(132) *S. aureus* protein A gene (*spa*)-typing is one of the commonest SLST typing methods used. As part of this sequencing of the polymorphic X region of protein A gene is carried out. It is cost-effective, easy to use, has fast turnaround time and is reproducible but in some instances can misidentify some types as a result of recombination however Koreen *et al.* reports that this is very rare and *spa* typing is useful for characterising genetic microvariation in outbreak investigations and macrovariation in *S. aureus* phylogenetic and population based analysis.(132)(139)

The M protein gene (*emm)* encodes the M protein which is present on the surface of GAS cells.(140) It is one of the main factors for discriminating different GAS.(128) There are at least 100 GAS M serotypes.(140) Typing based on this is the commonest method for investigating GAS isolates.(132) Traditionally M-protein serological typing was the main method which was replaced by *emm* typing method which involves sequence analysis of the hypervariable region at the 5'end of the *emm* gene.(140)(132)(128)(141) Over 200 *emm* types have been identified using this method.(132) Methods include reverse dot blot or reverse line blot hybridization of specific probes with *emm* genes.(128) For full discrimination it is advised that additional methods such as PFGE or MLST should be used alongside *emm* typing.(132)

1.8.14 Microsatellite analysis/simple sequence repeats (SSR)

Microsatellite analysis or simple sequence repeats (SSR) typing is based on PCR amplification.(142) SSR consists of repeating sequences of 2–6 base pairs of DNA.(143) Microsatellite typing is more commonly used for fungal strain typing.(144)

1.8.15 Matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS)

Matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) can be used as a molecular typing technique utilising unique protein profiles to identify bacteria.(129) As part of this technique protein spectrums are created and compared to reference databases to identify the bacteria. MALDI-TOF MS methodology is rapid, sensitive, and has low costs in terms of labour and equipment. One limitation of this method is that the spectral database does not contain peptide mass fingerprints of all bacteria and therefore new isolates will not always be classified by this technique.(145) MALDI-TOF MS has been applied to the investigation of a variety of outbreaks. It was used alongside WGS in the investigation of VRE and the authors found that MALDITOF-MS was a faster typing technique at distinguishing outbreak strains but lacked the level of discrimination that WGS gave. The authors therefore concluded that IPCTs should use this information with caution as there can be variability in protein expression and this means that the phenotype does not always correlate with the genotype.(146) Another study compared simple sequence repeats

(SSRs) polymerase chain reaction (PCR) amplification and MALDI-TOF MS to type isolates of *Candida parapsilosis* isolated from neonate blood cultures. They found both of the methods to be rapid and highly discriminatory.(142) A further study reported that their results of MALDI-TOF MS typing matched perfectly with those obtained from molecular identification (partial *rpoB* gene sequencing) in an outbreak of *Corynebacterium pseudodiphtheriticum* infection in CF patients.(147) MALDI-TOF MS has also been used in an outbreak of *S. pneumoniae* conjunctivitis.(148) In another outbreak it was used to study PVL-positive non-multiresistant MRSA in a neonatal ICU in Australia. In this outbreak researchers compared MALDI-TOF MS and SNP-plus-binary gene typing. They found that both methods were able to identify strains that were related.(149) Recently the IR Biotyper® has been marketed as a spectroscopic system for analysis of isolates. It contains a high performance FT-IR (Fourier Transform Infrared) spectrometer and software for automatic spectra measurements and production of dendrograms.(150)

1.9 Whole Genome Sequencing (WGS)

1.9.1 Applications of WGS in clinical microbiology

In the field of microbiology advances in WGS began with the first bacterial genome (*H.influenzae*) being sequenced in 1995.(151) 'First generation sequencing' techniques marked significant scientific breakthroughs but are now considered by many to be expensive and slow.(152) Competition between companies to improve WGS technology has resulted in an expansion of "next generation sequencing" (NGS) techniques.(153) Examples of these include sequencing by synthesis (Illumina), pyrosequencing (Roche 454), sequencing by ligation (SOLiD platform, Life Technologies), ion semi conduction (Ion Torrent platforms Life Technologies) and single molecule real-time sequencing (Pacific Biosciences). Technology continues to emerge with one of the most recent being nanopore sequencing (Oxford Nanopore).(153) These individual platforms will be discussed in further detail later in this chapter. WGS can be applied to all microorganisms (bacteria, fungus, viruses and parasites) and used to analyse their entire genomes.(154) Figure 1.10 shows the many applications of WGS in microbiology.

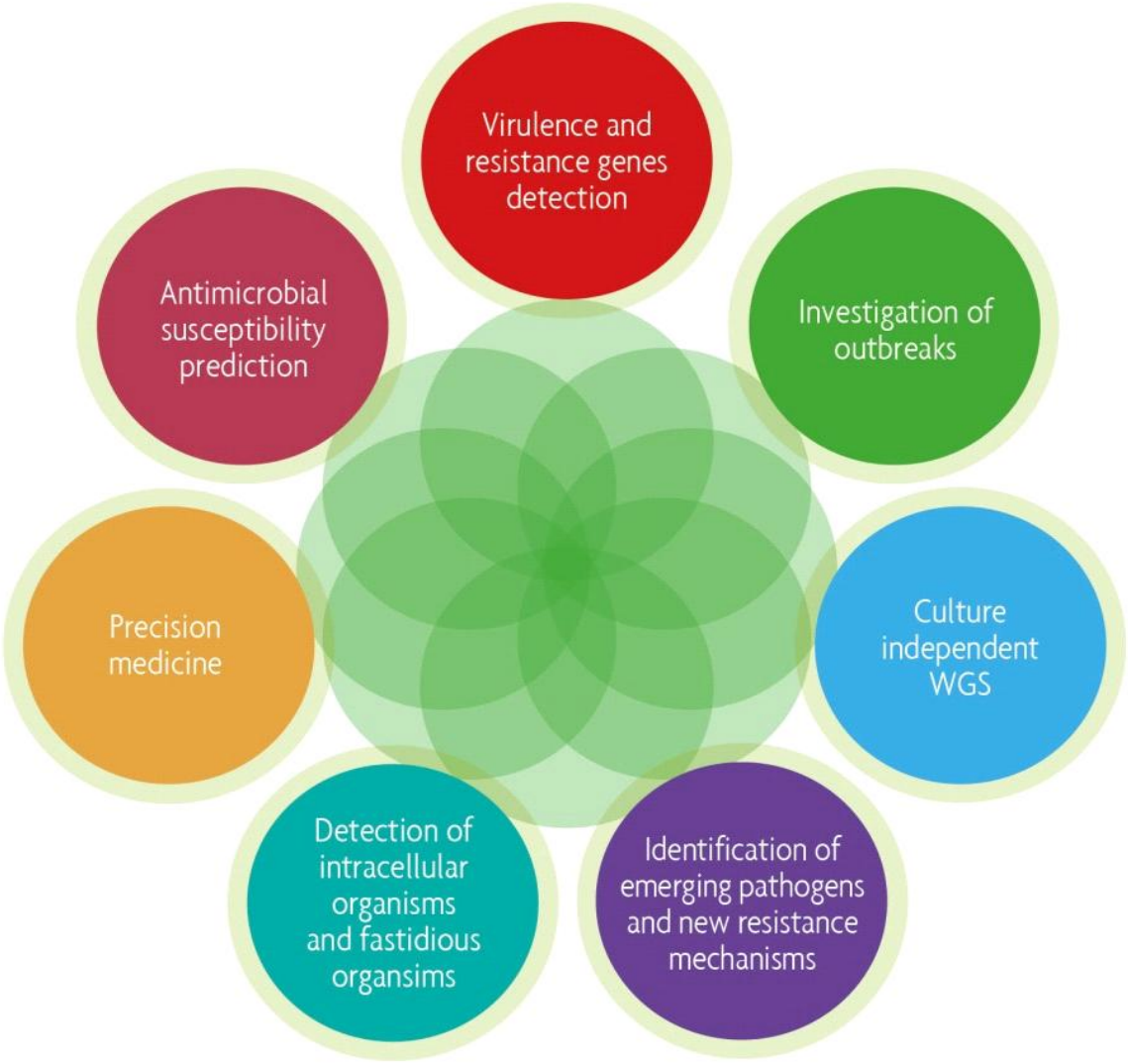


Figure 1.10. The applications of WGS. WGS has been used to detect virulence and resistance genes along with the detection of fastidious microbes. It can be used to detect microbes without the need for culture and can also be used to identify emerging pathogens. One application is that WGS can predict antibiotic sensitivity results. WGS has been used in precision medicine to allow accurate treatments.

In a proof of concept study WGS was successfully integrated into the routine workflow of a clinical laboratory for a working day with researchers sequencing the genomes of all isolates recovered from culture.(10) Additional applications of WGS include virulence-associated genes detection, antimicrobial susceptibility prediction, detection of intracellular organisms and fastidious bacteria and the identification of emerging pathogens and new resistance mechanisms.(5). Culture independent WGS is also increasingly being carried out and has the potential to improve clinical laboratory workflow and turn-around times (11) Other potential applications include genome-scale metabolic modelling, biosurveillance, and

bioforensics.(155) At present, the most common way of defining the taxonomy of organisms is to determine its 16S rRNA gene sequences. (116) WGS has the potential to provide greater resolution than 16S rRNA sequence analysis providing additional information on structural, metabolic and functional aspects of bacteria. (156) Researchers studying the microbiome commonly use amplicon analysis of the 16S ribosomal RNA (rRNA) gene. It has been shown that shotgun WGS can provide an increase in the detection of diversity, bacterial species and prediction of genes compared to 16S amplicon sequencing.(157) WGS now has the potential to revolutionise the delivery of clinical microbiology services including outbreak detection.

To investigate the number of published articles utilising WGS a search of Pubmed was carried out 19/09/2019 using the term “whole genome sequencing". This identified 24,276 articles. To further investigate these, a search of MeSH was carried out to identify the MeSH categories that researchers had published their WGS work under. The commonest MeSH category that researchers were publishing their work against was ‘organisms’. To assess the application of WGS for outbreak investigation a bibliographic search was carried out using keywords ‘whole genome sequencing AND (healthcare outbreak OR public health outbreak)’ using Pub Venn (https://pubvenn.appspot.com/). This revealed a total of 108,150 citations associated with a healthcare outbreak or public health outbreak. There were 24,215 citations associated with WGS and 1,012 citations in which WGS had been applied to outbreak investigations. This search is displayed as a Venn diagram in Figure 1.11.

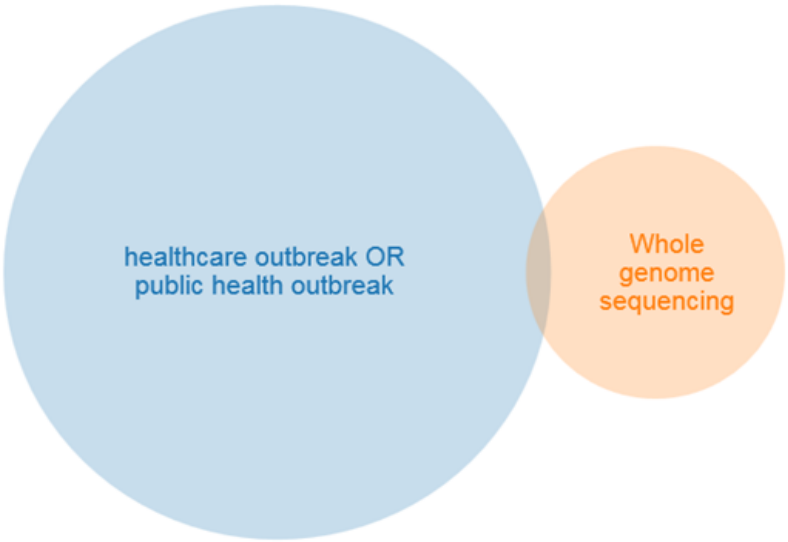


Figure 1.11. Venn diagram displaying citations associated with WGS and healthcare or public health outbreaks. Keywords 'whole genome sequencing AND (healthcare outbreak OR public health outbreak)' were searched for using Pub Venn (https://pubvenn.appspot.com/).

The terms 'WGS AND outbreak' were used to search Mapping MEDLINE (https://esperr.github.io/mapping-medline/) in order to gain on overview of where WGS is being carried out for outbreak investigation. MeSH headings can include geographic regions and topic areas with geographic regions such as name of a continent, country or city being tagged in approximately 15% of all indexed records. Some papers also state a country name in a title or abstract without a corresponding index term and this is also included in the search. In total there were 220 geographically tagged results for "WGS AND outbreak". Table 1.9 shows the geographically tagged results for "WGS AND outbreak".

Table 1.9. Geographically tagged results for "WGS AND outbreak"

Geographical area	Tagged Results
Africa	17
Caribbean Region	2
Europe	128
Oceania	11
Asia	35
Central America	1
North America	41
South America	8

Figure 1.12 shows the proportion of MEDLINE citations for “WGS AND outbreak” in Europe.

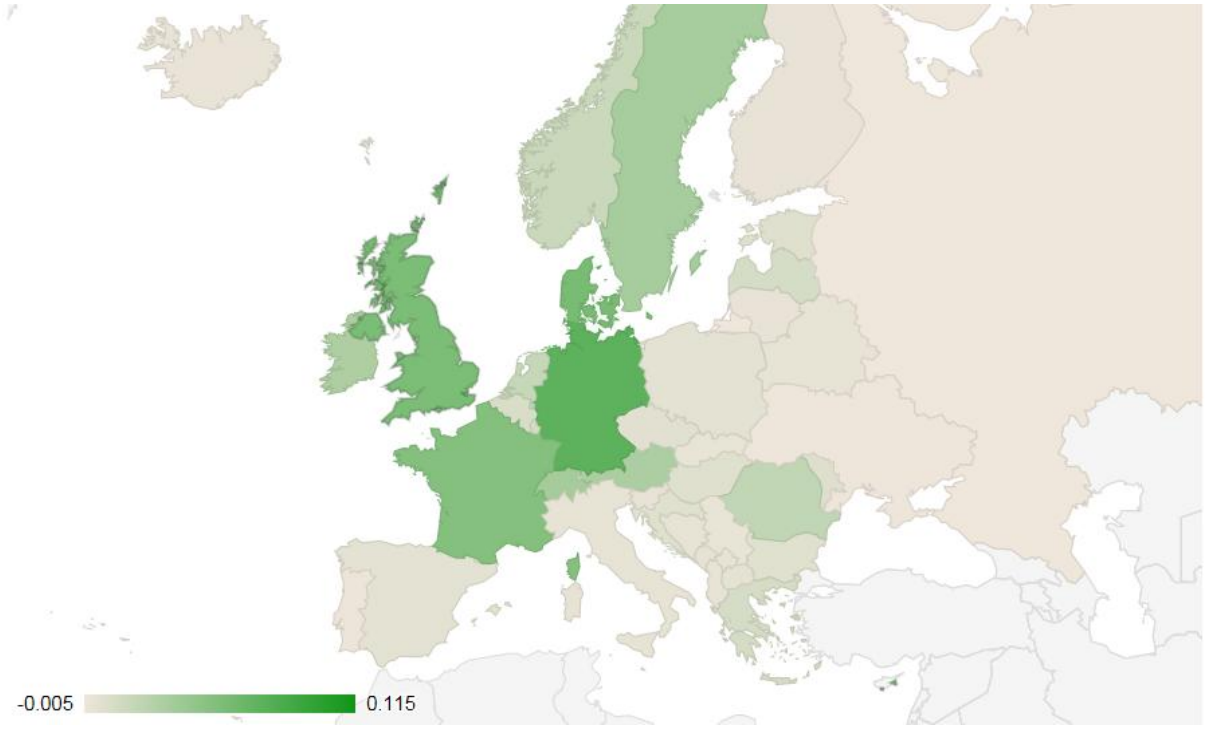


Figure 1.12. Map showing the proportion of MEDLINE citations for “WGS AND outbreak” in Europe. The terms 'WGS AND outbreak' were used to search Mapping MEDLINE (https://esperr.github.io/mapping-medline/) in order to gain on overview of where WGS is being carried out for outbreak investigation.

Google trends can also be used to measure global interest and online searches. The term ‘outbreak sequencing’ was entered into Google trends (https://trends.google.com/trends/) to investigate the number of global online searches using this term from 2004 -29. (158) It can be seen in the Figure 1.13 below that there was a spike of activity with online searches for this term in 2006.

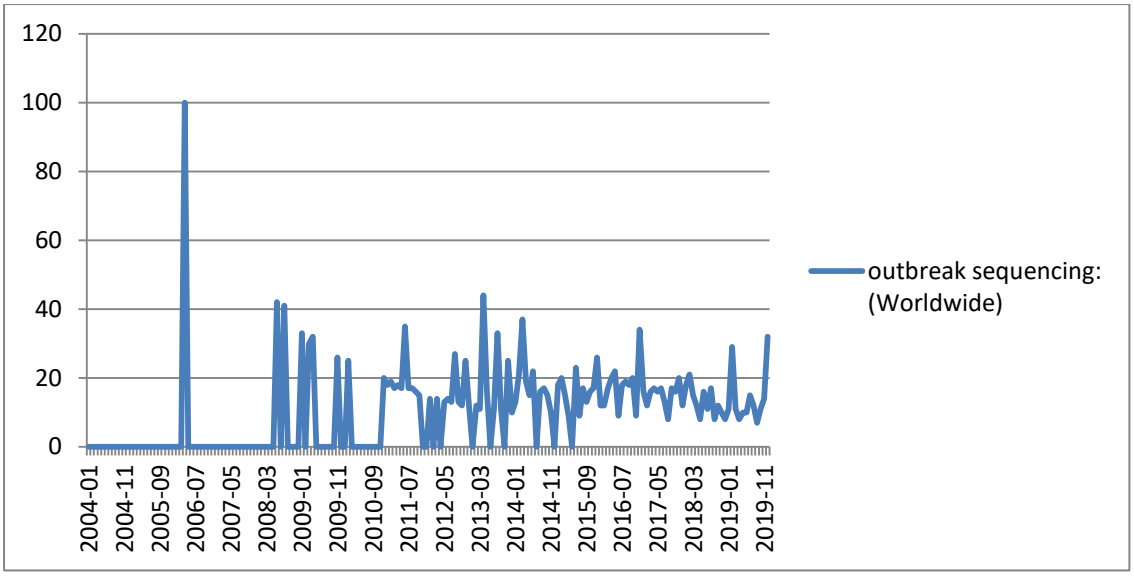


Figure 1.13. The number of global online searches of search engine Google using the term 'outbreak sequencing' from 2004 to 2019. The term 'outbreak sequencing' was entered into Google trends (https://trends.google.com/trends/).

A selection of notable publications describing the application of WGS for outbreak investigation will be discussed below to give an overview of the development of WGS for IPC. A landmark paper that proposed the use of WGS for outbreak investigation was published in 2010. In this study 63 MRSA isolates were sequenced using an Illumina Genome Analyzer GAII.(159) The MRSA isolates were ST239 and had been part of global collection.(159) When results were compared to PFGE and *spa* typing WGS was confirmed to have given greater discrimination of the isolates. This study identified that there had been intercontinental transmission and 5 isolates taken at a Thailand hospital differed by 14 SNPs suggesting they were related.(159)

A year later Ion Torrent technology was applied in a real-time investigation of a large *E. coli* 0104:H4 outbreak in Germany in 2011.(77) Using WGS in this way was a paradigm shift for outbreak investigation. This outbreak was unusual as there was a high incidence of haemolytic uremic syndrome, involving a high proportion of adults, the majority being female. The use of WGS was innovative as it involved crowd-sourcing. The first draft of the genome was completed after just 3 days and uploaded onto the internet changing the precedent for outbreak investigation with WGS. This prompted great interested amongst researchers across the world with each research group racing against each other to upload their work onto open-source wiki. They identified that the *E. coli* contained a Shiga toxin-

encoding prophage and a plasmid encoding CTX-M-15 extended-spectrum β-lactamase (ESBL).(160)

A proof of principle study was published in 2012 (a year later) using benchtop HiSeq to investigate a NICU MRSA outbreak. Retrospective samples of MRSA ST22 from NICU patients underwent sequencing. As part of this a phylogenetic tree was constructed revealing a distinct cluster of outbreak isolates. It identified there had been a missed transmission event as two isolates were found to be related from two patients with bacteraemia. The researchers also found concordance between the antibiotic genes identified by WGS and phenotypic susceptibility testing results.(161)

A pilot study of rapid WGS for the investigation of a *Legionella* outbreak was published a year later in 2013.(14) This described the genomic analysis of seven *L. pneumophila* isolates (three clinical and four environmental) from a 2003 outbreak. The researchers reported that they were able to distinguish outbreak from non-outbreak isolates confirming the probable environmental source which supported the findings from the initial outbreak investigation.(14) In the same year a paper was published describing the use of WGS in the investigation of *C. difficile* outbreaks over a 3-year period. In this study WGS revealed that 45% of *C. difficile* patients were genetically distinct from all previous patients which was an unexpected finding.(162)

In a study by Harris *et al.* WGS was applied to the investigation of an MRSA outbreak in NICU and a special care baby unit (SCBU).(78) In this unit, MRSA screening was carried out routinely each week and 3 patients were found to be positive for MRSA. These patients had been on the unit at the same time and an outbreak investigation was carried out which included a 6-month review of previous patients. In this 6-month period, 13 patients were identified to be MRSA positive however these patients were not all thought to be part of the outbreak as there had been numerous times when no MRSA had been isolated on the unit. When sequencing was applied it was found that 14 MRSA positive patient's isolates were highly related suggesting they were part of the outbreak. Following this, the IPCT screened staff on the unit. There was de-convolution of this outbreak after a staff member who was identified to be a carrier was decolonised. This study also identified that MRSA transmission had occurred beyond the SCBU and had spread into the community after further sequencing of 19 MRSA isolates in the laboratory. These had been selected for sequencing based on

antibiogram information. The source of sample submission included family doctors, emergency department and breast surgery outpatient clinic.(78) After establishing genomic surveillance in the university hospital diagnostic microbiology laboratory a second study revealed that the same outbreak strain, ST2371 was still being identified from GP and hospital ward patients. All strains were found to be highly related after phylogenetic comparison of these isolates with the 45 isolates from the original outbreak.(163)

The following year in 2014, a paper was published detailing the investigation and reconstructed evolutionary history of the pandemic *Vibrio cholera* in Haiti. WGS was applied to 116 serotype O1 strains from different global sources, including 44 Haitian genomes. Researchers identified that the source of *V. cholerae* was not indigenous to Haiti and that it had been introduced from Nepal.(164)

A large scale prospective longitudinal genomic surveillance study focusing on proactive sequencing was published in 2017, as part of this, 2282 MRSA screening isolates were sequenced from 1465 patients. These isolates had been collected over a 12-month period from a routine diagnostic microbiology laboratory in England and included isolates from samples submitted by hospital and GP users. Using epidemiology and WGS results 173 transmission clusters were identified each involving 2 to 44 patients. These had not been detected by the IPCT or by the laboratory.(165) It was later described that phylogenetic analysis had been used to identify an E-MRSA15 cluster of isolates from 15 patients from the same GP surgery. Two of these patients had died with MRSA bacteraemia and 11 had been receiving leg ulcer GP clinic care. This demonstrated that WGS could be used to detect outbreaks that could have been missed by routine IPC intervention.(166)

1.9.2 Development of sequencing techniques

1.9.2.1 First-generation DNA sequencing

The three-dimensional structure of DNA was first identified by Watson and Crick in 1953 after they reviewed Rosalind Franking and Maurice Wilkins crystallographic data.(167) Researchers first focused on sequencing RNA from microbial ribosomes or transfer RNA and single stranded RNA bacteriophages as these could be grown in culture, did not have a complementary strand and RNAse enzymes were available to cut RNA chains.(168) In 1965

Holley *et al.* produced the first whole nucleic acid sequence (alanine tRNA) from *Saccharomyces cerevisiae*.(169) The next advancement occurred in the 1970s with Sanger *et al.* publishing the first complete DNA genome belonging to a bacteriophage ϕX174 ('PhiX') based on using the "plus and minus" method.(170)(171) The Maxam and Gilbert technique also used polyacrylamide gels but differed as radiolabelled DNA broke chains at certain bases.(172) A further advancement was made in 1977 with Sangers 'chain-termination' dideoxy techniques which used chemical analogues of deoxyribonuclotides (dNTPs) as monomers of DNA stands.(173) This method involved mixing radiolabelled dideoxynucleotides (ddNTPS) during DNA extension reactions. As part of this the ddNTPS lacked 3'hydroxyl groups and could not form a bond with the 5'phosphate of the next dNTP.(168) DNA strands of various lengths were produced as dideoxynucleotides became incorporated. Four parallel reactions were carried out each with a ddNTP on four lanes of polyacrylamide gel. Following this autoradiography was used to show the nucleotide sequence in the initial template by looking at the radioactive band in the corresponding lane in the gel.(168)

1.9.2.2 Second-generation sequencing

Developments in sequencing continued when Pål Nyrén *et al.* developed an luminescent sequencing by synthesis (SBS) method called pyrosequencing in which pyrophosphate synthesis is measured.(168) In this method ATP sulfurylase converts pyrophosphate into ATP. ATP is then used as a substrate for firefly luciferase and light was produced in portion to pyrophosphate. Sequences were then inferred by measuring pyrophosphate production making this technique more sensitive.(174)(168) The technique was licensed to 454 Life Sciences (later owned by Roche) and was the first major commercial "next generation sequencing" (NGS) technology.(168)(174)(9) The GS20 was the first machine available to consumers.(168) Other methods developed with DNA libraries being attached to beads using adapter sequences. Each bead was treated with a water-in-oil emulsion PCR amplifying one DNA molecule on one bead.(168)The next step involved the DNA –coated beads being washed while pyrosequencing occurred as smaller bead-link enzymes and dNTPs were washed over the beads on a plate. A charged couple device (CCD) sensor was used to measure pyrophosphate release. Reads of 400-500 base pair (bp) long could be created.(168). The Genome Sequencer FLX instrument (454, Inc.) was used to sequence James D. Watsons

human genome in 2008 and was faster and approximately one hundredth the cost of traditional capillary electrophoresis methods.(175)

Another SBS technique took the name of Ion Torrent (Life Technologies).(9) It is the first NGS technology that worked without optical sensing and used novel 'post-light sequencing technology' in which clonal populations of DNA fragments were created using emPCR and attached to beads. The beads and nucleotides were washed over a picowell plate.(9)(168) Complementary metal-oxide-semiconductor (CMOS) technology allowed nucleotide incorporation and this was measured using the difference in pH (due to the release of protons (H^+ ions).(168)

Further sequencing methods were developed such as Solexa sequencing in 1998. Illumina Inc. acquired Solexa in 2007 and the technology is still used today.(176) Illumina technology focuses on using flow cells and does not use bead-based emPCR.(168) The technique involves adapter-bracketed DNA molecules which are moved over a lawn of complementary olignonucleotides on flow cells. Following this a solid phase PCR creates clusters of clonal populations from the flow cell binding DNA strands.(168) Bridge amplification occurs as the replicating DNA strands move over to prime the following polymerisation from neighbouring olignonucleotides. Fluorescent 'reversible terminator' dNTPs are also used.(177) These dNTPS are not able to bind to other nucleotides as there is a fluorophore which sits in the 3'hydroxyl position. For polymerisation to occur this must be cleaved away.(168) Modified dNTPs and DNA polymers are washed over flow cells in cycles and the nucleotides that are incorporated are detected using CCD by exciting the fluorophores with lasers.(168) To begin with Genome Analyzer (GA) machines could only produce short reads (up to 35 bp long) but could produce paired-end (PE) data (each sequence ends of DNA clusters can be detected). This produces more information improving accuracy especially when mapping reads to reference sequences. (168) Illumina Genome Analyzers have been periodically updated. The HiSeq 2000 was developed for high throughput parallel sequencing and in 2012 the MiSeq was released as a faster, lower-throughput, lower cost machine capable of longer read lengths.(176)(178)

Developments in second-generation sequencing continued with a number of new techniques being created by different companies. Sequencing by ligation (SBL) was utilised by Applied Biosystems with the oligonucleotide ligation and detection (SOLiD) system. This system uses

a two-base-encoded probes with each fluorometric signal representing a dinucleotide.(9) SOLiD platforms do not produce the same read length or degree of depth as Illumina machines but it has been reported that its the cost of sequencing has remained lower than other techniques.(168) Complete Genomics DNA nanoballs technique is also based on SBL. DNA sequencing is carried out with combinatorial probe- anchor synthesis (cPAS) or combinatorial probe-anchor ligation (cPAL).(9)

1.9.2.3 Third-generation DNA sequencing

It has been proposed that third generation technologies should be defined by their ability to sequence single molecules.(168) Helicos BioSciences developed the first single molecule sequencing (SMS) technology which was similar to the technology used by Illumia but without amplification. This technology was innovative as it was the first to sequence non-amplified DNA (avoiding errors and associated biases).(168) As part of this DNA templates were used and attached to planar surfaces, following this fluorescent reversible terminator dNTPs (virtual terminators) were washed over each base and an image taken. This produced short reads and was slow and expensive and the Helicos Genetic Analysis System became one of the casualties of the NGS arms race.(168)(9)

The most popular used long-read technology is the single molecule real-time (SMRT) sequencers by Pacific Biosciences (PacBio).(9) Specialized flow cell with thousands of picolitre wells are used called zero mode waveguides (ZMW). These have transparent bottoms and light moves through apertures which have a diameter smaller than the wavelength.(9)(168) This triggers decay to occur in an exponential fashion and this is illuminated at the bottom of the walls. When this occurs single flurophore molecules can be visualised at the bottom of ZMWs. Single DNA polymerase molecules are placed in the ZMWs within the illuminated region. Extension of DNA chains by single nucleotides are monitored as the DNA library and fluorescent dNTPs are washed. This can occur in real-time and can produce very long reads.(168)

Further developments in sequencing include nanopore sequencing. By using electrophoresis researchers identified that single stranded RNA or DNA could be moved across a lipid bilayer through alpha-haemolysin ion channels and as this happens ions flow is blocked. This

reduces the current for a time which is proportional to nucleic acid length.(168) The nanopores developed by the Oxford Nanopore Technologies are considered by some to be fourth generation sequencing.(179) In 2012 the larger GridION was released which can be used to produce several gigabytes of data in a day. The technology was miniaturized with the MinION becoming available to consumers in 2014.(179)(9) Recently in 2015, the MinION was utilised for the real-time genome sequencing of the Ebola virus disease (EVD) epidemic in West Africa. This technology was innovative due to its portability (weight was under 100 grams) and results were read from a laptop powered by a Universal Serial Bus (USB) port. (180) Oxford Nanopore technologies also developed the PromethION benchtop platform that has a higher output with 48 flow cells and parallel sequencing involving 144,000 nanopores simultaneously.(129) Table 1.10 and Table 1.11 shows an overview of third generation sequencing platforms and their performance properties (adapted from Quainoo *et al.* 2017 and Deurenberg *et al.*, 2017).(129)(181)

Table 1.10. Overview of third generation sequencing platforms read length, output, coverage, run time, number of reads and costs

Company	Platform	Read length (bp)	Output (GB)	Coverage	Run time (hr)	Number of reads	Cost per Gb max ($)
Illumina	MiniSeq Mid output	2x150	2.1-2.4	8.6	4-24	14-16million	2,953
	MiSeq v3	2x75	3.3-3.8	13.6	5-55	44-50 million	1,362-1,568
	NextSeq 500 High output	1x75	25-30	107.7	12-30	<400 million	312-374
PacBio	Pacific biosiences RS II P6-C4	>20,000	8-16	57.4	0.5-4	55,000	250-500
Oxford Nanopore	MinION MK1 (2D)	>882,000	10-20	71.8	1min-48hr	138,000	49.95-99.9
	PromethION single flow cells	<300,000	233	836.2	1.67->72	26million	NA
	PromethION 48 flow cells	<300,000	11,000	3,9475.8	1.67->72	1.25 billion	NA

GB, gigabytes.

Table 1.11. Overview of third generation sequencing platforms consumable and instrument costs, error rate and instrument dimensions

Company	Platform	Consumables cost	Instrument cost	Error rate	Dimensions (width x depth x height) (cm)
Illumina	MiniSeq Mid output	6,201	55,411	0.1% >80% base calls	45.6x48x51.8
	MiSeq v3	5,,174	108,244	0.1% >85% base calls	68.6x56.5x52.3
	NextSeq 500 High output	9,347	266,835	0.1% >80% base calls	53.3x63.5x58.4
PacBio	Pacific biosiences RS II P6-C4	4,000	695,000	14% errors per base	203.0x90.0x160.0
Oxford Nanopore	MinION MK1 (2D)	999	1,000	15% errors per base	10.5x3.3x2.3
	PromethION single flow cells	NA	135,000	NA	44.0x24.0x40.0
	PromethION 48 flow cells	NA	135,000	NA	44.0x24.0x40.0

Comparative microbial genomics is essential in order to answer molecular epidemiological questions, identify risk factors and understand the transmission of pathogens.(182) Data analysis is considered to be the biggest challenge in relation to introducing WGS in a clinical microbiology environment.(181) The process of pathogen sequencing usually involves a clinical sample being taken and cultured. After this DNA is extracted from bacterial colonies and a DNA library is created for sequencing. Fragments (reads) are produced after sequencing in FAST Q or BAM formats (for PacBio sequencers), these are typically several hundred nucleotides long and can be assembled into contiguous DNA sequences (consigns). (181)(183)(184)(185)(186). Bioinformatics brings together the database storage of sequences and the application of computational algorithms to analyse DNA.(187) Bioinformatics originated in the 1960s when computational methods were applied to protein sequence analysis and Margaret Dayhoff is credited with being "the mother and father of bioinformatics" as she developed COMPROTEIN a computer program to identify protein primary structures.(188) As major improvements were made in sequencing technology in the 1990s there was an exponential rise in data and increase requirement for expertise in computer science and since then bioinformatics has become a discipline in many universities.(188)(188) There are three way to analyse microbial genomes: reference mapping; *de novo* assembly and assembly free methods.(182)

1.9.2.4 Reference mapping

Reference mapping is also known as guided assembly. Reference mapping is ideal in situations in which organisms have low population diversity and when mobile genetic elements do not need to be investigated.(182). Pipelines such as bwa and bowtie2 include alignment and can be used for this.(182) As part of alignment a reference genome is used for mapping reads from the isolate being investigated. Researchers look for differences between the two and this forms the basis for SNP typing.(182)(176)(185) Occasionally certain reads will not map to the reference genomes due to them containing novel regions and if this occurs filtering is carried out.(185) Pipelines in which reference mapping is carried out can contain variant calling options such as GATK, VarScan and freebayes.(182) Annotation can also be completed with breseq or Prokka. Phylogenetics and prediction of resistance can also be included.(182)

1.9.2.5 *De novo* assembly

De novo assembly is an alternative to mapping.(185) In *de novo* assembly reads are aligned against each other and a reference organism is not required.(181) This method may be preferred for organisms that are difficult to characterise due to high genomic plasticity and extrachromosomal sequences. It is useful for the analysis of chromosomal and extrachromosomal sequences such as plasmids.(182) Reads are assembled based on overlap and scaffolding of the data. When short read data is used problems can occur as it can be more difficult to assemble compound repeat regions.(183). Genome assembly by *de novo* methods is generally more accurate and easier with larger fragments from long-read sequencing such as sequences produced by Pacific Biosciences and Nanopore instruments.(183)(182) *SPAdes* is one of the most common *de novo* assemblers.(182) Other software packages include CLC Genomic Workbench (Qiagen), and Velvet.(181)

1.9.2.6 Assembly free methods

Assembly free methods are a new development in comparative microbial genomics. Usually k-mer or hash-based methods are used to examine query sequences.(182) They can be used for rapid molecular analysis such as SNP-based methods with kSNP and MLST for example SRST2 and can also be used to predict resistance e.g. k-mer resistance.(182)

1.9.2.7 Quality control

It is important that sequence quality is assessed for all of these approaches in terms of read level quality control, trimming poor quality and adaptor sequences and demultiplexing the sample into individual isolate read sets.(182) Quality control processes detect miscalled bases as sequencing machines carry out base calling in which 'raw' signals are converted to reads for each nucleotide.(186) When they are not able to do this with confidence the call will be recorded with an N. Software tools such as Fast QC can assess quality of data. Phred-like quality scores (Q scores) can be used to measure the probability that bases have been correctly called. In general, a score of Q25 and above is an accepted score to continue processing the data. Following this quality can be assessed by filtering and pre-processing sequence reads before the genome is assembled.(186)

1.9.2.8 Phylogenetic analysis

Molecular typing is often used to establish the relatedness of isolates and Spratt defines the term clone as *"isolates of bacterial species that are indistinguishable in genotype"*.(138) Genomes of bacteria are subject to genome re-arrangement due to movement of insertion sequences , deletions and localized recombinational replacements.(138) The rate of clonal diversification and development of clonal complexes depends upon on the rate of recombination which can vary amongst bacteria.(138) Phylogenetics is the basis for comparative genomics in which comparisons can be made between genomes assisting us in understanding the evolutionary relationships amongst genomes.(189) Phylogenetic trees are branching diagrams that contain nodes (junctions between lines) which reflect ancestors for the sequences. (187)(189). The ends of lines represent sequences divergence.(187) Phylogeny is derived from the Greek word "*phylo*" meaning family and "*geny*" meaning ancestry.(187) Charles Darwin sketched one of the first evolutionary trees in 1837, his work entitled 'On the Origin of Species' was published in 1859 offering one of the first descriptions of biological evolution. (190)(191) Since then Willi Hennig has promoted the use of phylogenetic trees and cladistics for describing evolutionary relationships and classifications.(192) Carl Woese published a paper on bacterial evolution in 1987 which pointed out that microbiology had been "devoid of evolutionary concepts" for many years and with the rapid developments of sequencing nucleic acids he made the prediction that " no discipline within biology will be more changed by this revolution than microbiology".(193)WGS is now a fully recognised

tool that can be used to identify the degree of relatedness of bacterial isolates and assist us in identifying the chain of transmission when used alongside epidemiological findings.(185) It has been shown that it is possible to infer phylogenetic relationships between bacterial isolates and a number of programmes are now available online which can be used for inferring phylogenetics directly from next generation sequencing (NGS) data. Programmes such as CSI Phylogeny (Center for Genomic Epidemiology), snpTree, NDtree, and Reference sequence Alignment based Phylogeny builder (REALPHY) can be used.(194)

1.10 Summary

Traditionally outbreaks are investigated using surveillance and epidemiology by identifying patients with alert/indicator organisms. Early in suspected outbreak situations we routinely rely upon the use of phenotypic methods to identify alert organisms and their resistance patterns from patient/environmental samples. An outbreak investigation is usually initiated if the same alert organism is identified from two separate patients' taken at the same time and in a similar place. To manage this, IPC measures are employed and screening of patient contacts may be carried out using phenotypic methods to identify the same organism. Following this, patients who are positive are likely to be isolated or cohorted and in some circumstances will be offered a course of antibiotics for decolonisation. The basis of this approach is to break the ongoing chain of transmission from patients but this relies on phenotypic techniques and routine typing to accurately detect alert organisms and their resistance patterns. It does not assume that dissemination of genetic elements such as transposons and plasmids carrying antibiotic resistance genes is a dynamic process or that traditional typing may lack the discriminatory power to identify the genetic lineage of isolates. Potentially this strategy could result in a failure of IPC measures if an alert organism is misidentified or not detected. Currently confirming transmission relies on traditional typing methods such as PFGE, VNTR and *spa* typing although WGS has recently been introduced for certain typing at reference laboratories e.g. *Salmonella* typing. It is recognised that routine typing results can obstruct outbreak investigations. This can occur when results are not rapid and also when they are unable to unequivocally show whether isolates are linked or not. For instance, PFGE is the considered by many to be the gold standard for microbiological typing and although it is standardized internationally it cannot differentiate isolate to the same degree as WGS.(195)

1.11 Objectives

The first and main objective of this work is to establish a WGS service for the investigation of suspected HAI outbreaks. As part of this I also intend to identify the challenges of implementing this new technology including identifying the instances in which increased WGS discriminatory power and phylogenetic analysis is required (objective 2). Additionally, I will identify clinical benefits that can occur from using WGS for suspected outbreak management (objective 3). Meeting these objectives will inform the development of a clinical decision aid on how best to utilise WGS in HAI outbreak investigations in real-time.

1.12 Hypotheses

This work tested the following hypotheses in relation to using WGS for the management of suspected HAI outbreaks in healthcare settings:

Hypothesis 1. - Establishing a WGS service in the NHS in the UK is achievable.

This will be tested by implementing a WGS service for healthcare outbreak detection in NHS Tayside and NHS Grampian and identifying practical barriers for this. Sequencing is still relatively expensive compared to certain other typing techniques and in practice it is not feasible to sequence all suspected outbreaks. To determine the basis for phylogenetic analysis and increased WGS discriminatory power I will therefore review WGS findings from all the suspected outbreak situations and compare these to conventional typing results to determine the situations in which increased resolution from WGS may be required to unequivocally define strains.

Hypothesis 2. - Utilising WGS for the management of outbreaks will result in clinical benefits.

To test this, I will review the impact of using WGS results compared to conventional typing results in suspected healthcare outbreaks from a microbiologist, IPCT and clinical care team perspective.

2 General Methods

2.1 Ethics and Caldicott approvals

Caldicott Guardian approval for NHS Grampian (Appendix 1) and NHS Tayside (Appendix 2) was gained in order to carry out this study. This enabled protection of patient confidentiality and appropriate information sharing. NHS Biorepository approval (Appendix 3) was also gained so that surplus patient bacterial isolates could be used. Ethics approval from the University of St Andrews ethics was also granted for this study (Appendix 4).

2.2 Bacterial identification

Clinical specimens were first collected as part of routine care and initially processed at the Departments of Medical Microbiology in NHS Tayside and NHS Grampian (both accredited to ISO 15189:2012 by UK Accreditation Service (UKAS)). Specimens were inoculated by Biomedical Scientists on Columbia blood agar plate and cultured aerobically, anaerobically, in carbon dioxide atmosphere and on MacConkey agar aerobically for 48 hours at 37 °C. Bacterial identification was carried out by Biomedical Scientists using matrix-assisted laser desorption/ionization machines (bioMérieux, Marcy L'Etoile, France and Bruker Diagnostics, Germany). If MALDI-TOF MS was not available or failed to identify the organism Vitek 2 (bioMérieux, Marcy L'Etoile, France) was used. Antibiotic susceptibility testing was undertaken by Biomedical Scientists using Vitek 2 (bioMérieux, Marcy L'Etoile, France) or disc testing with both techniques using MICs in accordance with the European Committee on Antimicrobial Susceptibility Testing (EUCAST). Environmental water sampling in ICU was carried out by an external contractor who sampled all water outlets for *P. aeruginosa* using pre- and post-flush samples.

2.3 Vitek 2

The following process was used by Biomedical Scientists for identification and antibiotic testing using Vitek 2 (bioMérieux, Marcy L'Etoile, France) in the NHS laboratories as part of routine care. A bacterial suspension was created using inoculums of 3ml saline at room

temperature and 3 colonies of test organism from the culture plate. This was vortexed and tested on a densitometer to ensure a defined McFarland as per Table 2.1.

Table 2.1. Vitek identification and antibiotic sensitivity testing cards and McFarlane standards

Vitek Cards		McFarland
Identification	Gram negative identification (GN)	0.5 – 0.63
	Gram positive identification	0.5 – 0.63
	Anaerobe and *Corynebacteria* identification (ANC)	2.7 – 3.3
	Fastidious organism identification (NH)	1.8 – 2.2.
Antibiotic testing	AST-N297/N381 Non- urine Coliform sensitivity	0.5 – 0.63
	AST-N253 Pseudomonas and *Acinetobacter* sensitivity	
	AST-N254/N382 Urinary Gram Negative Rod sensitivity	
	AST-ST03 *Streptococcus* including *S. pneumoniae* sensitivity	
	AST-P634 *Staphylococci* sensitivity	
	AST-P607 *Enterococci* and urinary Group B *Streptococcus* sensitivity	

AST, antibiotic sensitivity testing.

A purity plate was inoculated using the inoculums and incubated 16-24 hours at 36^{o}C +/- 1. The inoculums were inserted into a Vitek 2 cassette at the Smart Carrier Station™. Once the cassette was loaded it was incubated by the Vitek 2 machine with card reading undertaken by the system.(196)

2.4 Matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS)

Isolates were identified in NHS laboratories by Biomedical Scientists as part of routine care using matrix-assisted laser desorption/ionization machines (bioMérieux, Marcy L'Etoile, France), (Bruker Diagnostics, Germany). The direct transfer method was used for routine identification in which a wooden cocktail stick was used to collect a very small amount of a colony. This was then smeared to create a thin film directly onto a spot on a MALDI-TOF MS target plate. An even distribution was made across the spot. Following this 1μL of HCCA α-Cyano-4-hydroxycinnamic acid (HCCA) solution was overlaid on top of the spot and allowed to dry at room temperature. After this the target plate was inoculated and matrix had dried it was transferred to the biotyper and a run created using the software.

2.5 Disc diffusion testing and minimum inhibitor concentration (MIC) evaluation

As part of routine care EUCAST disc sensitivity testing was carried out by Biomedical Scientists by preparing an inoculum preparation of several morphologically similar colonies from a culture following overnight growth on non-selective medium. The colonies were suspended in saline with a sterile loop or a cotton swab. A suspension to the density of a McFarland 0.5 using the Vitek densitometer was prepared. Plates were inoculated within 15 minutes of this being carried out by dipping a sterile cotton swab into the suspension and using a rotary plater to spread the inoculums evenly over the entire surface of the plate. Control isolates were also tested on separate plates. Discs were applied within 15 minutes of this by rigidly attaching them to the inoculated surface and dried agar plate. No more than six discs were applied per plate. The required disc strengths for antibiotic panels are listed in Table 2.2.

Table 2.2. EUCAST antimicrobial susceptibility testing methods and media

Organism	Antimicrobial susceptibility testing	
	Sensitivity method	**EUCAST disc testing media**
Enterobacterales species	Vitek and Disc	If vitek fails MH
Enterococcus species	Vitek	If vitek fails MH
L. monocytogenes	Disc	MHF
Pseudomonas species	Vitek	If vitek fails MH
Pseudomonas species. CF	Disc	MH
Staphylococcus species	Vitek	If vitek fails MH
Streptococcus species	Vitek and Disc	MHF

MH, Mueller Hinton; MHF, Mueller Hinton Fastidious (MH +5% defibrinated horse blood and 20mg/L NAD); CF, cystic fibrosis.

Plates were incubated within 15 min of disc application. Plates that contained *Streptococcus* and *L. monocytogenes* were incubated in 4-6% CO_2, and others in air. Plates were incubated at 35+/-1^{o}C for 16-20 hours.(196)

2.6 Referral of isolates

I was involved in the IPC team assessment of outbreaks and selected isolates that should be sent for reference laboratory typing during bench rounds and by reviewing ICNet infection surveillance software Version 5.9 findings. My decision to send isolates for typing was also based on SPC information, patient epidemiology and antibiotic sensitivity results. Figure 2.1 shows integration of WGS into outbreak investigation.

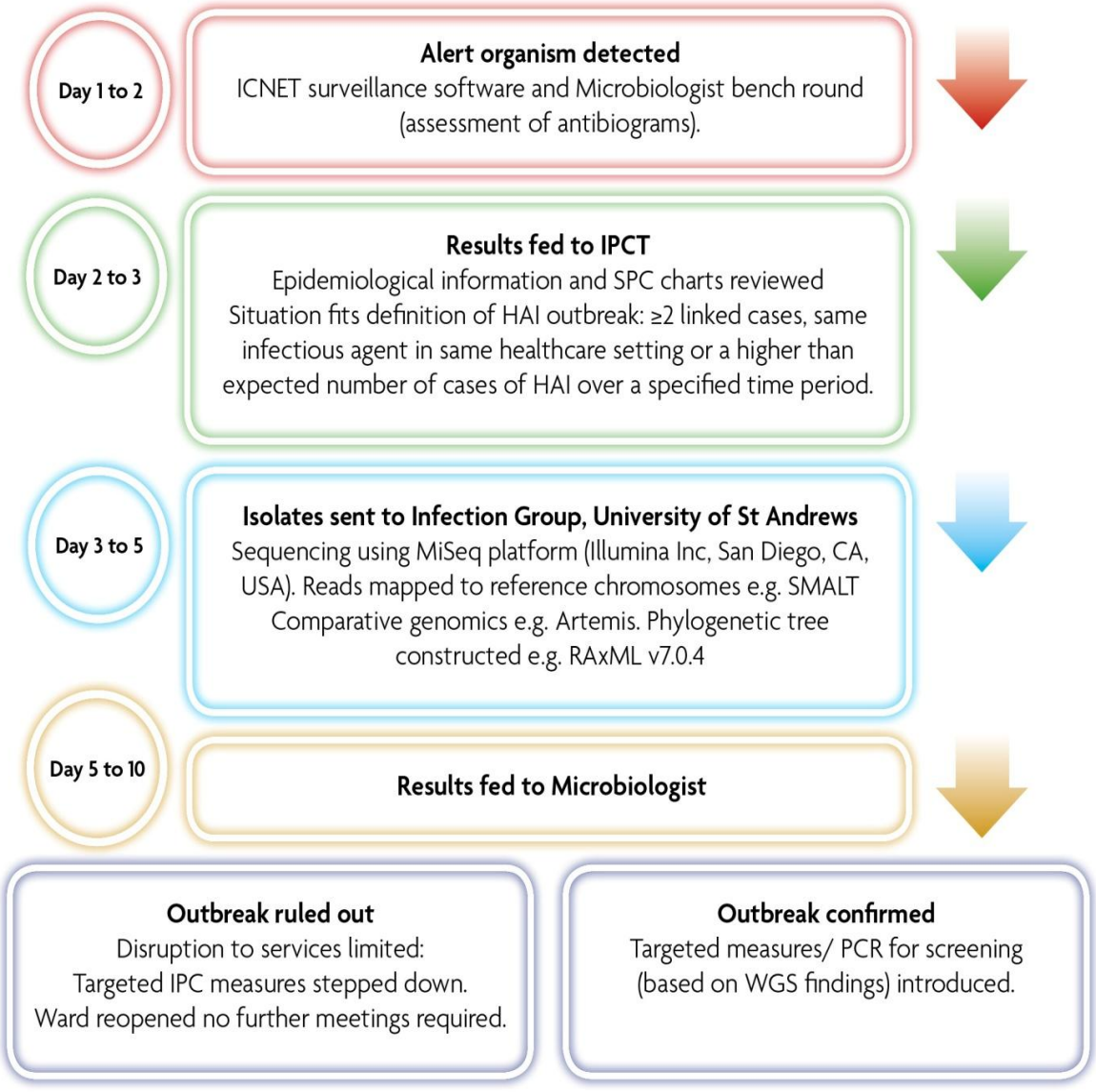


Figure 2.1. Figure demonstrating the various stages in which WGS was integrated into outbreak investigations. Bacterial isolates from NHS Tayside and NHS Grampian were sent to the Infection and Global Health Research Division, School of Medicine, University of St Andrews.

Figure 2.2 shows the referral form that was developed. Isolates were sent to the following reference laboratories; Antimicrobial Resistance and Healthcare Associated Infections Reference Unit (AMRHAI), Public Health England, Colindale (for VNTR typing and PFGE), Scottish MRSA Reference Laboratory (for *spa* typing or PFGE) and the Scottish AMR Satellite Reference Laboratory, Glasgow. In parallel with this, bacterial isolates were also referred to the Infection and Global Health Research Division, School of Medicine, and University of St Andrews for WGS.

Scottish Healthcare Associated Infection Prevention Institute

(*Clinical Pathogen Genomics Laboratory*)

Infection Group
University of St Andrews
School of Medicine
North Haugh
St Andrews, KY16 9TF

DX319801 ST ANDREWS 90 KY
Phone: +441334 463539
Email: Clinical queries:fife-uhb.ustan-shaipi@nhs.net
General enquiries: shaipi@st-andrews.ac.uk
Information www.infection.st-andrews.ac.uk/shaipi

Sender's information

Senders name and address:

Post code

Report to be sent FAO:
b.parcell@nhs.net
Direct Phone:07753379093
Direct Phone (out of hours):

☐SHAIPI ☐Charge (*specify*):..........................

Source information

☐Human ☐Animal ☐Food ☐Environment (*specify*).............. ☐Other (*specify*)......................

Hospital name (*if different from sender's*):...

Sample information

Presumptive identification
☐*S. aureus* MRSA ☐*E. faecalis* VRE ☐Klebsiella spp. – ESBL ☐Other (specify)..........klebsiella.........
☐*S. aureus* MSSA ☐*E. faecium* VRE ☐*E. coli* – ESBL ☐Pseudomonas spp (*specify*)...............

Identification method ☐x Vitek ☐MALDI-TOF ☐API....................... ☐Other (*specify*)......................

Date submitted:.. Priority status ☐High ☐Medium ☐Low
*a The laboratory must be contacted before sending high risk and urgent samples

Reason for referral of isolates

☐Strain relatedness - outbreak ☐SNAPSHOT[a]
☐Biomarker detection –Toxin-Gene......................... ☐Genomovar determination[a]
☐Unusual phenotype[a] (*specify*)................................ ☐Other (*specify*)
[a] These investigations must be agreed in advance before dispatch

Comments/Epidemiological Information

Clinical Pathogens[c] ***Characterisation and molecular epidemiology*** (multiple isolates)

Continuation page. Not to be used without Page 1. Page_____of_____(total number of pages sent)

SHAIPI – Clinical Pathogen Genomics & Molecular Epidemiology – Infection Group, University of St Andrews 1

Source Information	Sample information	SHAIPI Use Only
☐Patient ☐Staff☐ Environment Ward/Speciality: __________	**Laboratory ID:**______________ Body site:_________Sample date :_________	
☐Patient ☐Staff☐ Environment Ward/Speciality: __________	**Laboratory ID:**______________ Body site:_________Sample date :_________	
☐Patient ☐Staff☐ Environment Ward/Speciality: __________	**Laboratory ID:**______________ Body site:_________Sample date :_________	
☐Patient ☐Staff☐ Environment Ward/Speciality: __________	**Laboratory ID:**______________ Body site:_________Sample date :_________	
☐Patient ☐Staff☐ Environment Ward/Speciality: __________	**Laboratory ID:**______________ Body site:_________Sample date :_________	
☐Patient ☐Staff☐ Environment Ward/Speciality: __________	**Laboratory ID:**______________ Body site:_________Sample date :_________	
☐Patient ☐Staff☐ Environment Ward/Speciality: __________	**Laboratory ID:**______________ Body site:_________Sample date :_________	
☐Patient ☐Staff☐ Environment Ward/Speciality: __________	**Laboratory ID:**______________ Body site:_________Sample date :_________	
☐Patient ☐Staff☐ Environment Ward/Speciality: __________	**Laboratory ID:**______________ Body site:_________Sample date :_________	
☐Patient ☐Staff☐ Environment Ward/Speciality: __________	**Laboratory ID:**______________ Body site:_________Sample date :_________	

[c] If more than 10 isolates are to be sent, agreement with the laboratory must be obtained.

SHAIPI – Clinical Pathogen Genomics & Molecular Epidemiology – Infection Group, University of St Andrews 2

Figure 2.2. Example of an Infection and Global Health Research Division referral form for WGS. This form was used to record bacterial isolates being sent from NHS Tayside and NHS Grampian microbiology laboratories to the Infection and Global Health Research Division for WGS.

2.7 *Staphylococcus aureus* toxin genes

The detection of resistance, Panton-Valentine leukocidin (PVL) toxin, toxic shock syndrome toxin (TSST), and exfoliative toxins A (ETA) and B (ETB) genes were performed using two multiplex PCR assays and standard gel electrophoresis at the Scottish MRSA Reference Laboratory by laboratory staff.

2.8 *Spa* typing

Spa typing and analysis of this was carried out by laboratory staff using Ridom StaphType (Ridom GmbH, Germany) at the Scottish MRSA Reference Laboratory.

2.9 Variable-number tandem repeat (VNTR)

Environmental and patient isolates were sent to the Antimicrobial Resistance and Healthcare Associated Infections Reference Unit (AMRHAI), Public Health England, Colindale for typing by Biomedical Scientists (VNTR typing at nine loci and PFGE) as described previously.(197)

2.10 Pulsed field gel electrophoresis (PFGE)

PFGE was performed by laboratory staff using *SmaI* as per the manufacturer's protocol and CHEF DRII system (Bio-Rad Laboratories, Hemel Hempstead, UK). Profiles were analysed using BioNumerics 7.6 (Applied Maths, Sint-Martens-Latem, Belgium). PFGE of *Xba*I-digested genomic DNA was carried out as described previously at the AMRHAI laboratory.(198)

2.11 *emm* typing

GAS isolates were subcultured and DNA was extracted by laboratory staff using an achromopeptidase extraction method as described previously at the reference laboratory.(199). After sequencing was carried out on an AB 3500 xL sequencer sequences were manually edited using CLC Main Workbench software and FASTA files were

uploaded to basic local alignment search tool (BLAST) 2.0 server on the CDC website by the Scottish *Haemophilus*, *Legionella*, *Meningococcus*, *Pneumococcus* Reference Laboratory.(140)

2.12 Storage of isolates

Once isolates arrived at the Infection and Global Health Research Division, School of Medicine, University of St Andrews, they were stored on prolab beads at -80ºC until processed.

2.13 DNA extraction

The isolates were recovered from frozen and cultured by laboratory staff at the Infection and Global Health Research Division, School of Medicine, and University of St Andrews . The DNA was then extracted using either the QIAamp DNA Mini Kit (Qiagen Hilden, Germany) protocol for cultured cells or the MasterPure^{TM} Gram Positive DNA Purification Kit (Epicenter, Madison, WI, USA).(200)(201) In the first year of my MD I assisted in recovering frozen isolates, plating, culture, DNA extraction and sequencing.

Following the QIAamp DNA Mini Kit (Qiagen Hilden, Germany) protocol for cultured cells. the first step involved heating the heating block to 56°C. Distilled water was then brought to room temperature (15-25°C) for elution.(200) Cells grown on culture were removed using a cell scraper, 5 x 10^6 cells were transferred to a 1.5ml microcentrifuge tube to be centrifuged for 5 minutes at 300 *x* g. After this, supernatant was removed completely and discarded. The cell pellet was resuspended in PBS to make a volume of 200ul. Following this, 20 ul QIAGEN Protease was added. The next step involved adding 200 ul of Buffer AL to the sample mix, pulse-vortexing for 15 seconds and incubation was carried out at 56 ° C for 10 minutes. After this, the 1.5 ml microcentrifuge tube was centrifuged to remove drops from inside of the lids. Following this, 200 µl 96-100% ethanol mix was added and the tube was pulse vortexed for 15 seconds. After mixing the 1.5ml microcentrifuge tube was centrifuged to remove drops from inside lid. After this the mixture was applied to the QIaAMP Mini spin column (using a 2ml collection tube), the cap was closed, and mixture centrifuged at 6000 x g (8000 rpm) for 1 minute. The QIA amp Mini spin column was then placed in a clean 2 ml

collection tube and the tube containing the filtrate was discarded. The QIAamp Mini spin column was opened and 500 ul Buffer AW2 added. The cap was closed after this and column centrifuged full speed at 20,000 x g (14,000 rpm) for 3 minutes. The QIamp Mini spin column was placed in a new 2ml collection tube and the old tube with the filtrate was discarded after this and the column was centrifuged at full speed for 1 minute. The QIAamp Mini spin column was then placed in a clean 1.5ml microcentrifuge tube and the collection tube containing filtrate was discarded. The QiaAMP Mini spin column was opened and 200 ul distilled water added. The column was then incubated at room temperature (15-25°C) for 1 minute and centrifuged at 6000 x g (8000 rpm) for 1 min. Control checks were also carried out using a NanoDrop (1.5ul sample) – 260:280 ratio >1.8. Qubit BR (2ul sample) – concentration >10ng/ul.(200)

The MasterPure™ Gram Positive DNA Purification Kit (Epicenter, Madison, WI, USA) protocol involved Gram Positive DNA purification and separate DNA precipitation steps.(201) The first step of the Gram Positive DNA purification part involved centrifugation of 1.0 mL of an overnight Gram-positive bacterial culture to create a pellet. After this, the supernatant was discarded. Following this, 150 µL of TE Buffer was added and vortexed to resuspend the cell pellet. The next step involved adding1 µL of Ready-Lyse Lysozyme to each resuspended pellet (from 1.0 mL culture) of bacteria. Incubation at 37°C for 30 minutes to overnight was carried out after this. The next step involved diluting 1 µL of Proteinase K (50 µg/µL) into 150 µL of Gram Positive Lysis Solution for each 1.0 mL of culture pellet. 150 µL of the Proteinase K/Gram Positive Lysis Solution was then added to the sample and the sample was mixed thoroughly. Incubation was then carried out at 65-70°C for 15 minutes, and vortexing briefly carried out every 5 minutes. The samples were cooled to 37°C and placed on ice for 3-5 minutes.(201) After this, the DNA precipitation steps were carried out which involved adding 175 µL of MPC Protein Precipitation Reagent to 300 µL of lysed sample. This was vortexed for 10 seconds to mix vigorously. Centrifugation at 4°C for 10 minutes at >10,000 × g in a microcentrifuge was carried out in order to pellet the debris. The supernatant was transferred to a clean microcentrifuge tube and the pellet was discarded. Following this, 1 µL of RNase A (5 µg/µL) was added to each sample and mixed thoroughly. The samples were incubated at 37°C for 30 minutes and after this 500 µL of isopropanol was added to the recovered supernatant. The tube was inverted 30-40 times and then centrifugation at 4°C for 10 minutes was carried out at >10,000 x g in a microcentrifuge in order to pellet the DNA. A pipette tip was used to remove the isopropanol without dislodging

the DNA pellet. Next the pellet was rinsed with 70% ethanol. The centrifuge was used briefly if the pellet became dislodged. After this, the DNA was resuspended in 35 µL of TE Buffer.(201)

2.14 Nextera XT sample preparation protocol

The isolates underwent sequencing by laboratory staff at the Infection and Global Health Research Division, School of Medicine, and University of St Andrews . In the first year of my MD I assisted in this.

2.14.1 Sample preparation

The quality of the DNA was measured as per the University of St Andrews Nextera XT sample preparation protocol determining a A280 nm/A260 nm ratio on NanoVue (GE Healthcare) as previously described.(202)(203)(204) Using NanoVue a reference sample was pippetted (2 ul of sample) to the black spot between the four alignment spots.(203) Following this, the top and bottom was cleaned and the next sample was pipetted. Concentrations were based on the absorbance at wavelength 1. These steps were repeated for each sample.(203) The concentration of double stranded DNA was assessed as described previously using dsDNA BR Kit (2µl DNA, 198µl Qubit working solution) on a Qubit 2.0 fluorometer (Thermo Fisher Scientific) Abs 260/280 ratio (optimal absorbance ratios were 1.8-2.0).(202)(69)

2.15 Tagmentation

Tagmentation was carried out following the University of St Andrews Nextera XT sample preparation protocol.(202) TD and ATM were removed from reagent box 1 and thawed. NT was removed from reagent box 2 vortexed, spun down and refrigerated in a pre-PCR area. Thawed TD and ATM tubes were gently inverted 3-5 times and spun briefly in a microcentrifuge. NTA was assembled by labelling a new 96-well plate and adding 10µl TD Buffer to each well with a multichannel pipette. Following this, using a multichannel pipette 5µl of DNA (0.4ng/µl = 2ng total) was added to each sample well of the plate. After this, 5µl ATM was added to each well using a multichannel pipette.(202) The sample then underwent

mixing using a multichannel pipette five times. Pipette tips were changed between samples. The plate was then covered with a PCR seal and centrifuged at 280 x g at 20°C for 1 minute. Following this, the plate was placed in a thermocycler and run (with heated lid) at 55°C for 4 minutes and held at 10°C. After the sample reached 10°C neutralisation steps were taken. As part of this, the plate was centrifuged, PCR seal removed and 5µl chilled NT Buffer was added to each well of the plate using a multichannel pipette.(202) Again this underwent mixing using a multichannel pipette 5 times and tips were changed between samples. The plate was then covered with a PCR seal, centrifuged at 280 x g at 20°C for 1 minute and plate left at room temperature for 5 minutes.(202)

2.16 PCR amplification

PCR amplification was carried out using the University of St Andrews Nextera XT sample preparation protocol.(202) The next stage involved thawing NPM and index primers on the bench at room temperature. These were mixed gently by inverting the thawed tubes 3-5 times and centrifuging them using 1.7ml Eppendorf tubes as adaptors. For the 24 libraries, the index primers were arranged in TruSeq Index Plate Fixture. As part of this, index 1 (i7) primers (orange caps) were arranged in order horizontally (N701 in column 1, N706 in column 6). Index 2 (i5) primers (white caps) were arranged in order vertically (S517 in row A, S502 in row B, S503 in row C, S504 in row D).(202) The plate was then placed in the TruSeq Index Plate Fixture. Following this, 15µl NPM was added to each well containing index primers using a multichannel pipette. Another 5µl of index 2 primers (white caps) were added to each column of the plate using a multichannel pipette. After this, 5µl of index 1 primers (orange caps) was added to each column of the plate, using a multichannel pipette. To avoid index cross-contamination tips were changed between samples. The original white and orange caps were then discarded and new white and orange caps applied. After this was carried out the plate was pipetted gently 3-5 times to mix using a multichannel pipette with the tips being changed between samples.(202) The plate was then covered with PCR seal and centrifuged at 280 x g at 20° for 1 minute. PCR (with heated lid) was then carried out at 72°C for 3 minutes and 95°C for 30 seconds. This underwent 12 cycles of 95°C for 10 seconds, 55°C for 30 seconds, 72°C for 30 seconds, 72°C for 5 minutes and held at 10°C.(202)

2.17 PCR clean-up

As part of the PCR clean-up AMPure beads were brought to room temperature (30 minutes). The University of St Andrews Nextera XT sample preparation protocol was followed and the plate was centrifuged at 280 x g at 20°C for 1 minute to collect condensation. (202) A new deep well plate was labelled. After this, 50µl of PCR product was transferred from the PCR plate to the deep well plate with tips being changed between samples. The AMPure XP beads were then vortexed for 30 seconds until homogenous. After this, 25ul beads were added to each well using a repeater pipette. The next step involved shaking the deep well plate on a microplate shaker at 1,800 rpm for 2 minutes. Following this, the plate was then incubated at room temperature without shaking for 5 minutes.(202) The deep well plate was then placed on a magnetic stand for 2 minutes or until the supernatant had cleared. The discard supernatant was carefully removed and discarded using a multichannel pipette (changing tips between each sample). After this, the beads were aspirated, dispensed back into plate on the magnet. The next step involved waiting for 2 minutes until the supernatant had cleared. The beads were washed with freshly prepared 80% ethanol with the deep well plate on a magnetic stand. After this, 200µl freshly prepared 80% ethanol was added to each well using a multichannel pipette. Beads were not resuspended. The plate was then incubated on the magnetic stand for 30 seconds. The supernatant was carefully removed and discarded. A second 80% ethanol wash was performed. A P20 multichannel pipette with fine tips was used to remove residual ethanol. With the deep well plate on a magnetic stand the beads were allowed to air-dry for 15 minutes. The deep well plate was removed from the magnetic stand. After this step, 45µl RSB was added to each well of the deep well plate using a repeater pipette. The S deep well plate was then placed on a microplate shaker at 1,800 rpm for 2 minutes.(202) Following this it was incubated at room temperature for 2 minutes. The deep well plate was then placed on the magnetic stand for 2 minutes or until the supernatant cleared. The next step involved labelling l new 96-well PCR plate. After this, 50µl of supernatant from the deep well plate was transferred to the PCR plate using a multichannel pipette (tips were changed between samples). The plate was then sealed with PCR seal and stored at -15 to -25°C for up to one week.(202)

2.18 Manual library normalisation and pooling

For this stage the University of St Andrews Nextera XT sample preparation protocol was followed and the concentration of the PCR product was quantified using Qubit High Sensitivity (2ul product added to 198ul Qubit mastermix).(202) The product was run on the Bioanalyser High Sensitivity DNA chip taking note of the average fragment size between markers of each sample.

Using excel, the molarity of each sample was calculated using the following:

Molarity of each sample (in nM) =B2/1*(500/C2*3) in which: B2=concentration in ng/ul (from Qubit HS); C2=average fragment size in bp (from Bioanalyser).

The average library concentration (nM) was calculated along with the volume of each individual library added to pool in order to achieve pooled library of average library concentration using the following:

Volume of individual library to add to pool ul = average library concentration nM* 5ul)/concentration of individual library nM

Individual libraries were pooled to average library concentration. Following this, they were vortexed, mixed and centrifuged to pool the library. The pooled library was diluted to the desired loading concentration (to a final volume of 5ul) using HT1 buffer. Between 4nM and 6nM of library was loaded (4.5nM is optimal to give a cluster density of 1 million per mm^2). The pooled library was kept in the fridge until ready to be processed.(202)

2.19 Preparing Miseq for run

Following the University of St Andrews Nextera XT sample preparation protocol the reagent cartridge was thawed in a basin of warm water (up to maximum water level on cartridge).(202) The MiSeq Control Software was rebooted.(202) (204) The Maintenance Wash (3 washes of ~20 minutes each) was carried out. Following this, a fresh wash solution was prepared using 25ml Tween 20 10% and 475ml MilliQ or Elga water.(202) In total, 6ml of wash solution was added to each reservoir of the wash tray, and the remaining ~350ml

transferred to the wash bottle. ‘Perform Wash’ mode was selected on the MiSeq Control Software by selecting ‘Maintenance Wash’ and following instructions. After this the Illumina Experiment Manager option was opened and ‘Create Sample Plate’ was selected. Sample IDs were entered along with the indices.(202) Following on from this, the work was saved and ‘Create Sample Sheet’ option was selected. As part of this the select ‘Small Genome Resequencing’ option was selected and a reagent cartridge barcode entered along with the run ID. Cycle number was set to 241.(202) The Illumina Experiment Manager instructions were followed and after this fresh 0.2N NaOH (200ul 1N NaOH + 800ul autoclaved MilliQ water) was prepared. After this, a new flow cell was removed from the fridge/cold room. (205) The flow cell was removed from the plastic container and storage solution retained in the container. Storage solution was washed off by flushing with MilliQ water and the flow cell was dried carefully with lint-free tissue. The next step involved adding several drops of 70% ethanol to a lint-free tissue and wiping the glass part of the flow cell to remove any residue or smears.(202)

To denature the library 5ul of the pooled library was added to 5ul fresh 0.2N NaOH. This was vortexed and centrifuged then incubated at room temperature for 5 minutes.(202) After this, 990ul of chilled HT1 buffer was added and inverted several times to mix using a pulse centrifuge. Directly after this, 600ul of library was loaded onto a reagent cartridge. The MiSeq Control Software was then opened and ‘Sequence’ option was selected. Instructions to load the flow cell, cartridge and incorporation buffer were followed. The ‘Start run’ option was selected after this and a green tick given. (202)

2.20 PhiX control

The PhiX Control (v3) was prepared as per instructions from the University of St Andrews Nextera XT sample preparation protocol using the Full MiSeq System Denature and Dilute Libraries Guide. (202)(206) The following instructions were used to prepare the 10 nM PhiX library to 20 pM. The following volumes were combined to dilute the PhiX library to 4 nM: 2 µl 10 nM PhiX library; 3 µl 10 mM diluent (Tris-Cl, pH 8.5 with 0.1% Tween 20). A fresh dilution of 0.2 N NaOH was also prepared. In a microcentrifuge tube 5 µl 4 nM PhiX library was combined with a 5 µl 0.2 N NaOH. After this, it was vortexed briefly to mix the 2 nM PhiX library solution. Following this, the template solution was centrifuged to 280 × g for 1 minute. It was then incubated for 5 minutes at room temperature to denature the PhiX library

into single strands.(202)(206) The following volume of pre-chilled HT1 was added to the tube containing denatured PhiX library to result in a 20 pM PhiX library: denatured PhiX library (10 µl); pre-chilled HT1 (990 µl). The MiSeq Reagent Kit v3 was used and the PhiX control was used at 20 pM concentration. For the majority of libraries, a low-concentration PhiX control spike-in at 1% was used. For low diversity libraries, PhiX control spike-in was increased to at least 5%. The following volumes of denatured PhiX control library were added to the denatured sample library. Most Libraries (1%): denatured and diluted PhiX control 6 µl; denatured and diluted sample library 594 µl. Low Diversity Libraries (≥ 5%): denatured and diluted PhiX control 30 µl*;* denatured and diluted sample library 570 µl. The combined sample library and PhiX control were set aside on ice until it was ready to load it onto the MiSeq reagent cartridge.(202)(206)

2.21 Read mapping and phylogenetic tree construction

After normalized libraries were sequenced using a 2×250 pair-end read of a 500-cycle v2 kit on a MiSeq platform (Illumina Inc, San Diego, CA, USA) using a resequencing workflow the Illumina sequences generated were deposited in the European Nucleotide Archive (ENA). Using SMALT (Wellcome Trust Sanger Institute; www.sanger.ac.uk/resources/software/smalt/), reads were initially mapped to reference chromosomes (Table 2.3).

Table 2.3. Reference genomes

Suspected outbreak	Reference chromosome
MRSA	EMRSA-15 reference (GenBank accession number HE681097)
L. monocytogenes	Strain: 4b F2365 (Genbank accession number AE017262)
E. coli O25b:H4-ST131	Strain: EC958 (GenBank assembly accession: GCA_000285655.3)
P. aeruginosa	PAO1 reference genome (Genbank accession number AE004091)

Bioinformatics was carried out by Professor Matthew Holden and Dr Miguel Pinheiro. Artificial 250bp pair-end reads fastq files were generated using a python script. Recombination was detected in the genomes using Gubbins (http://sanger-pathogens.github.io/gubbins/). Core genome regions of the reference genomes were defined

by human curation using pairwise BLAST comparisons with each other, and other outbreak strains. The Artemis Comparison Tool (ACT) was used to visualize the comparisons. SNPs falling inside Mobile Genetic Elements (MGEs) regions were excluded from the core genome, as well as those falling in regions predicted by Gubbins to have occurred by recombination. Phylogenetic trees were constructed separately using RAxML v7.0.4 for all sites in the core genomes containing SNPs, using a General Time Reversible (GTR) model with a gamma correction for among site rate variation.

2.22 MLST and detection of AMR-associated genes

For selected outbreaks SRST2, a read mapping-based tool was used for rapid detection of multi-locus sequence types (MLST) from WGS data by staff at the Infection and Global Health Research Division, School of Medicine, University of St Andrews. (207) ARIBA software was used for the detection of AMR-associated genes and SNPs by staff at the Infection and Global Health Research Division, School of Medicine, University of St Andrews. (208) In the last year of my MD I received training in this and used it to investigate several outbreaks.

2.23 Feedback of results

Throughout my MD I led and chaired the outbreak investigation meetings e.g. problem assessment groups (PAGS) and incident management team (IMT) meetings. As part of this I fed back sequencing results to the clinical teams and gave interpretation of the results.

3 Results

3.1 WGS for Gram positive HAI outbreaks

3.1.1 Meticillin-resistant *Staphylococcus aureus* (MRSA)

3.1.1.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of MRSA on a ward of an elderly rehabilitation hospital. An overview of the epidemiology and IPC investigation carried out for this suspected outbreak is shown in Table 3.1. From October 2014 to January 2015 there was a cluster of six cases of MRSA on this unit breaching the upper control limit.

Table 3.1. Overview of the epidemiology and IPC investigation carried out a suspected MRSA outbreak

Healthcare facility	41 bedded unit for elderly rehabilitation following acute care input
Microorganism	MRSA
Description of suspected outbreak	• Two patients identified to have leg wound infection due to MRSA. • Screening of further patients on the ward revealed 4 patients to be colonised with MRSA.
IPCT hypothesis	Likely outbreak based on epidemiology (patients with history of prior healthcare contact and antibiotic use) and cluster of cases occurring in a short time frame breaching the upper control limit of the SPC chart. Transmission due to breakdown in hand hygiene and wound care dressing.
Mode of spread	Most likely contact e.g. hands. It was identified that dressing changes were not appropriately carried out.
Case definition	Any patient identified to have a clinical sample positive for MRSA which is found to be PF15-314/315 on PFGE.
Total number of suspected patients	Overall 6 patients (2 infected, 4 colonised).
Outbreak specific control measures	Contact precautions, isolation, and education in relation to hand hygiene and wound care.
HIIAT Score	Green
Outbreak confirmed/not confirmed	Routine *spa* and PFGE typing supported the hypothesis that this was an outbreak.
WGS influence on IPC management	WGS results did not impact on IPC management. Routine typing turn-around time (TAT) was 27 days, and WGS TAT was 49 days. WGS results were not discussed at outbreak meetings but were shared with the IPC team.
Impact of WGS	**Greater discrimination** - WGS supported routine typing

	results and provided greater granularity demonstrating that patients 3, 4, 5 all clustered very closely together differing by only 9 SNPs suggesting recent transmission.

The SPC chart used in the investigation of this outbreak can be seen below in Figure 3.1

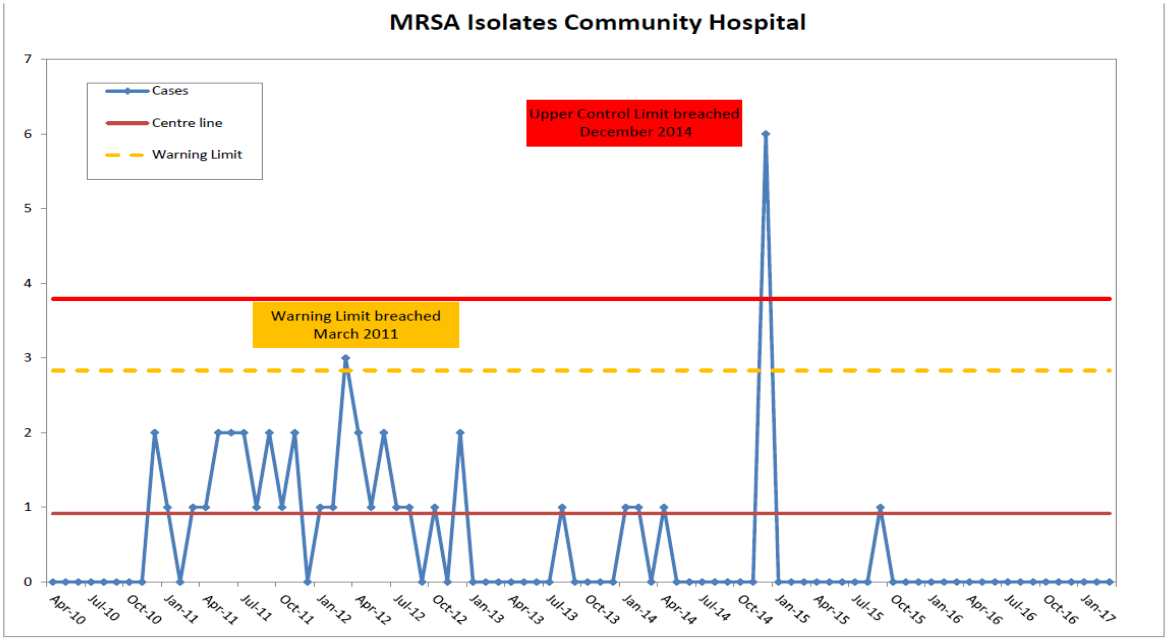


Figure 3.1. SPC chart of suspected MRSA outbreak demonstrating a breach of the upper control limit in NHS Grampian. Produced by NHS Grampian IPC department using data from ICNet outbreak surveillance software.

3.1.1.2 Routine microbiology results

Heat map of antibiograms of the MRSA isolates are shown in Table 3.2.

Table 3.2. Antibiograms of MRSA isolates

P	Sample	AMC	CIP	SXT	CXM	FLC	PEN	CHL	CLI	DOX	ERY	FA	GEN	LZD	RIF	TEC	VAN
1	Leg wound swab	R	R	R	R	R	R	S	S	S	S	S	S	S	S	S	S
2	Nasal swab	R	R	R	R	R	R	S	R	S	R	S	S	S	S	S	S
3	Nasal swab	R	R	R	R	R	R	S	R	S	R	S	S	S	S	S	S
4	Leg wound swab	R	R	R	R	R	R	S	S	S	S	S	S	S	S	S	S
5	Nasal swab	R	R	R	R	R	R	S	S	S	S	S	S	S	S	S	S
6	Nasal swab	R	R	S	R	R	R	S	S	S	S	S	S	S	S	S	S

P, patient; AMC, amoxicillin-clavulanic acid; CIP, ciprofloxacin; SXT, trimethoprim-sulfamethoxazole; CXM, cefuroxime; FLC, flucloxacillin; PEN, penicillin; CHL, chloramphenicol; CLI, clindamycin; DOX, doxycycline; ERY, erythromycin; FA, fuscidic

acid; GEN, gentamicin; LZD, linezolid; RIF, rifampicin; TEC, teicoplanin; VAN, vancomycin; R, resistant; I, intermediate; S, sensitive.

3.1.1.3 Reference laboratory typing

A summary table of Reference laboratory typing results is shown in Table 3.3

Table 3.3. MRSA *spa*-type and PFGE results

P	Date of sample	*Spa*-type	Inferred clonally complex	PFGE
1	10/12/14	NT		
2	12/12/14	t032	CC22	PF15-166 closely related but not identical
3	11/12/14	t032	CC22	PF15-314/315 indistinguishable
4	19/12/14	t032	CC22	PF15-314/315ndistinguishable
5	23/12/14	t032	CC22	PF15-314/315 indistinguishable
6	24/12/14	T2231	NK	PF15-46/75 closely related

P, patient; NT, isolate not tested or sent to reference laboratory.

3.1.1.4 Whole genome sequencing results

Five MRSA samples were sequenced one from each of patient 2,3,4,5,6. All of these isolates were EMRSA15. According to MLST samples all belonged to CC22, four of the samples to ST22 and one was a single locus variant with a novel allele. The sequence data was mapped to an EMRSA-15 reference. Three of the samples (patients 3, 4, 5) all clustered very closely together and differed by 9 SNPs, with patient 5 and patient 3 being closest differing by 3 SNPs (Figure 3.2). The next nearest genetically related isolate was patient 2 which was 87 SNPs away from this cluster. Genome data alone suggested that patient 5, patient 3, patient 4 were close enough genetically to be part of an outbreak.

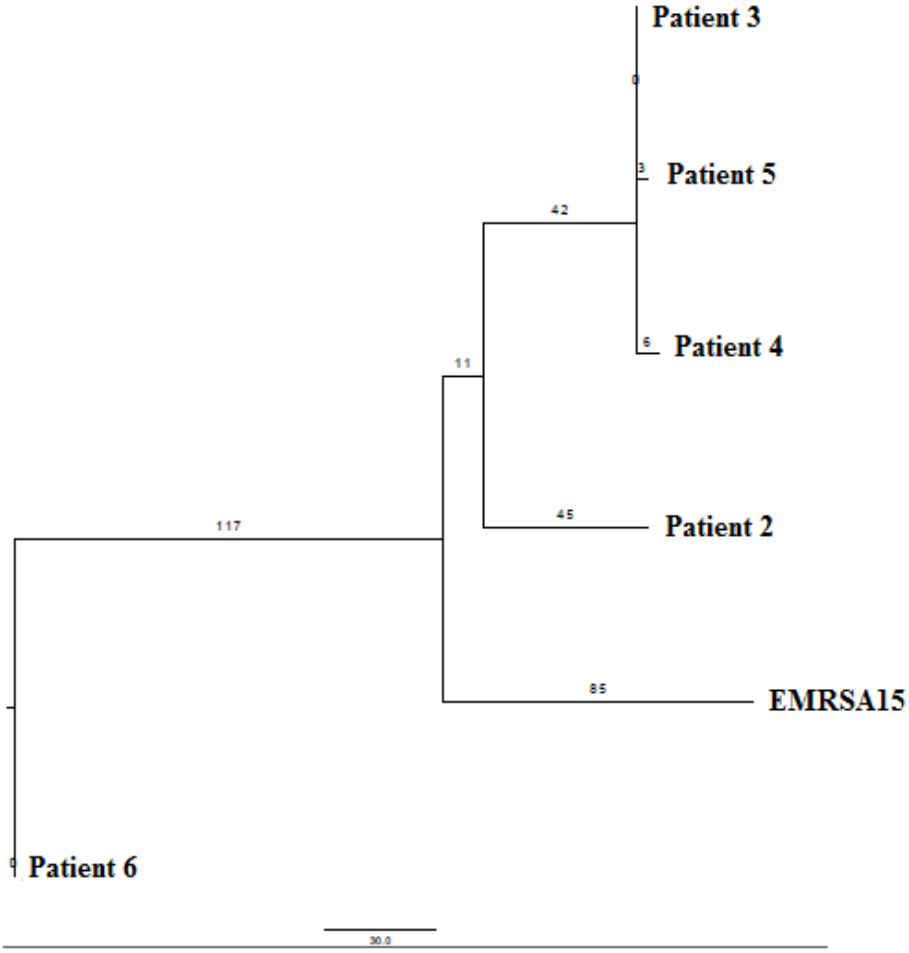


Figure 3.2. Phylogenetic tree of suspected MRSA outbreak isolates from a community hospital. Figure shows that isolates from patients 3, 4, 5 were close enough genetically to be part of an outbreak. These bacterial isolates all clustered very closely together. They differed by 9 SNPs, with patient 5 and patient 3 being closest differing by 3 SNPs. The next nearest genetically related isolate was patient 2 which was 87 SNPs away from this cluster. Data source - Infection and Global Health Research Division, School of Medicine, University of St Andrews.

3.1.2 Vancomycin-resistant enterococci (VRE)

3.1.2.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of VRE in a tertiary hospital. An overview of the epidemiological and IPC investigation carried out for this suspected outbreak is shown in Table 3.4. From December 2015 to February 2016, it can be seen that the number of VRE positive cases beached the upper control limit.

Table 3.4. Overview of the epidemiological and IPC investigation carried out for suspected VRE outbreak

Healthcare facility	Tertiary hospital
Microorganism	VRE
Description of suspected outbreak	• The urine from two different patient's urine samples was found to be positive for VRE on an orthopaedic rehabilitation ward. Further samples were taken and a rectal swab was found to be positive for VRE 6 days later from patient 1. Of note, it was identified that one of these patients was colonised with two different strains of vancomycin-resistant *E. faecium* (VREfm) one identified from rectal swab and another from urine, each strain related to an entirely different outbreak cluster in the main hospital. • Hand hygiene score was 70% compliant. Both patients were nursed in single rooms. There were no invasive devices or similar therapies. • Over a two-year period further positive patients were identified on a SHDU and Renal ward of the main hospital. PFGE and WGS identified that two vancomycin-susceptible *E. faecium* (VSEfm) isolates from two separate patients previously identified during a separate VSEfm outbreak in the ICU the year before were related to the ST64 cluster.
IPCT hypothesis	• Interhospital transmission between local hospitals and also a regional hospital carrying out renal transplants. Transmission within the ward due to break down in hand hygiene.
Mode of spread	Hand contact
Case definition	Any patient identified to have a clinical sample positive for VRE
Total number of suspected patients	14
Outbreak specific control measures	Contact precautions, isolation, education in relation to hand hygiene
HIIAT Score	Yellow

Outbreak confirmed/not confirmed	Confirmed
WGS influence on IPC management	WGS results were not produced in real-time due to batching. Results were shared with the IPC nurses but did not result in direct clinical impact.
Impact of WGS	**Greater discrimination**- PFGE identified 5 clusters in total (three ST80 clusters, one ST64 culture and one ST203 cluster). WGS provided greater granularity revealing there was only one ST80 cluster in total. **Enhanced 'alert organism' detection** -Both PFGE and WGS identified that two VSEfm isolates from two separate patients previously identified during a separate VSEfm outbreak in the ICU the year before were related to the ST64 cluster. The four isolates were differentiated by only 21 SNP sites suggesting a common source.

The SPC chart used in the investigation of this outbreak can be seen below in Figure 3.3.

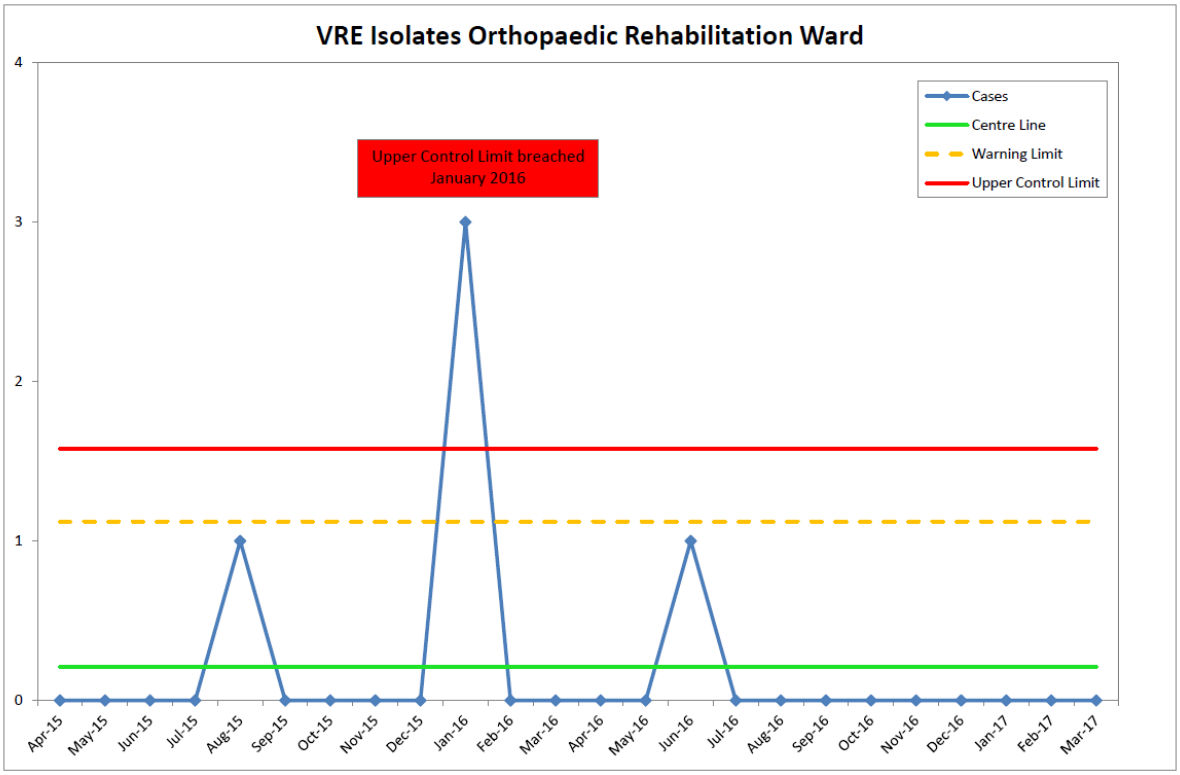


Figure 3.3. SPC chart of suspected VRE outbreak isolates from an orthopaedic rehabilitation ward. Demonstrating a breach of the upper control limit in NHS Grampian. Produced by NHS Grampian IPC department using data from ICNet outbreak surveillance software.

3.1.2.2 Routine microbiology results

Heat map of antibiograms of the isolates are shown in Table 3.5.

Table 3.5. Enterococci antibiograms

P	Date	Sample	Q-D	AMX	SXT	GEN	PEN	VAN	CHL	LZD	NIT	TGC
1	14/01/16	Rectal	I	R	R	R	R	R	S	S	S	S
	08/01/16	Urine	I	R	R	R	R	R	S	S	S	S
2	11/01/16	Urine	R	R	R	R	R	R	S	S	S	S
3	08/12/15	Urine	R	R	R	R	R	R	S	S	S	S
4	22/05/15	T tube fluid	I	R	S	R	R	R	S	S	S	S
5	23/04/15	Blood culture	R	R	S	R	R	R	S	S	S	S
6	3/01/15	BAL	S	R	R	R	R	S	S	S	S	S
7	2/1/15	BAL	S	R	R	R	R	S	S	S	S	S
8	04/09/15	Urine	I	R	R	R	R	R	S	S	R	S
9	27/04/16	Urine	R	R	R	R	R	R	S	S	S	S
10	21/09/15	Faeces	S	R	R	R	R	R	S	S	R	S
11	06/02/16	Urine	S	R	R	R	R	R	S	S	R	S
12	16/06/16	Urine	S	R	R	R	R	R	S	S	I	S
13	12/05/16	Blood culture	R	R	R	R	R	R	S	S	S	S
	25/05/16	Blood culture	S	R	R	R	R	R	I	R	I	S
	25/05/16	Blood culture	R	R	S	R	R	R	S	S	S	S
14	15/02/16	Urine	R	R	S	R	R	R	S	S	S	S

P, patient; Q-D, quinupristin-dalfopristin; AMX, amoxicillin; SXT, trimethoprim-sulfamethoxazole; GEN, gentamicin; PEN, penicillin; VAN, vancomycin; CHL, chloramphenicol; LZD, linezolid; NIT, nitrofurantoin; TGC, tigecycline; R, resistant; I, intermediate; S, sensitive.

3.1.2.3 Reference laboratory typing

Routine typing by PFGE of all the VREfm isolates from all patients identified that there were 5 clusters in total (three ST80 clusters, one ST64 culture and one ST203 cluster). Results of this are shown in Table 3.6

Table 3.6. PFGE of enterococci isolates

P	Sample	PFGE
1	Rectal	ABER13EC-4
	Urine	ABER13EC-3
2	Urine	ABER13EC- 4
3	Urine	ABER13EC-4
4	T tube fluid	ABER13EC-4
5	Blood culture	ABER13EC-4
6	BAL	ABER13EC-3
7	BAL	ABER13EC-3
8	Urine	ABER13EC-4
9	Urine	ABER13EC-4
10	Faeces	ABER13EC-1
11	Urine	ABER13EC-1
12	Urine	ABER13EC-3
13	Blood culture	ABER13EC- 5
	Blood culture	ABER13EC- 5
	Blood culture	ABER13EC- 5
14	Urine	ABER13EC-5' a highly similar, but not quite identical, profile to those from another patient 13

P, patient.

3.1.2.4 Whole genome sequencing results

When WGS was applied it revealed that in fact there was only one ST80 cluster in total (Table 3.7).

Table 3.7. Enterococci MLST result.

P	Sample	ST
1	Rectal	ST80
	Urine	ST64
2	Urine	ST80
3	Urine	ST80
4	T tube fluid	ST80
5	Blood culture	ST80
6	BAL	ST64
7	BAL	ST64
8	Urine	ST80
9	Urine	ST80
10	Faeces	ST203
11	Urine	ST203
12	Urine	ST64
13	Blood culture	ST80
	Blood culture	ST80
	Blood culture	ST80
14	Urine	ST80

P, patient.

WGS identified that two VSEfm isolates from two separate patients previously identified during a separate VSEfm outbreak in the ICU the year before were related to the ST64 cluster supporting PFGE findings. The four isolates were differentiated by only 21 SNP sites suggesting a common source.

3.1.3 *optrA* gene positive *Enterococcus faecalis*

3.1.3.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of *optrA* gene positive *E. faecalis*. Patients had been identified to be positive with *optrA* gene positive *E. faecalis* in the community and also in a tertiary hospital. An overview of the epidemiological and IPC investigation carried out for this suspected outbreak is shown in Table 3.8.

Table 3.8. Overview of the epidemiological and IPC investigation carried out for suspected *optrA* gene positive *E. faecalis* outbreak

Healthcare facility	Tertiary hospital and community
Microorganism	*optrA* gene positive *E. faecalis*
Description of suspected outbreak	• A patient was hospitalised and found to have linezolid resistant *E. faecalis* in a urine culture. The isolate was submitted by the routine microbiology laboratory to AMRHAI and the *optrA* gene was identified. • A retrospective search of stored linezolid resistant *E. faecalis* isolates by the routine microbiology laboratory resulted in the identification of 2 further *optrA* gene positive *E. faecalis* isolates (both were urine samples submitted by GPs). • The isolates submitted in 2014 and 2015 were reported as *cfr* gene negative – which was the plasmid-borne mechanism of resistance known at that time
IPCT hypothesis	• Due to this being a rare resistance mechanism the IPCT were concerned that cases could be related involving community transmission. Epidemiology data was limited and route of transmission was unknown.
Mode of spread	Unknown of note, one patient was a beef cattle farmer.
Case definition	Any patient identified to have a clinical sample which has *optrA* gene positive *E. faecalis*
Total number of suspected patients	3
Outbreak specific control measures	Side room, contact precautions
HIIAT Score	Yellow
Outbreak confirmed/not confirmed	Outbreak ruled out
WGS influence on IPC management	WGS results were not produced in real-time due to batching. They did not directly influence IPC management.
Impact of WGS	**Used to investigate new/unusual resistance mechanism. Greater discrimination**-WGS was utilised to identify a novel resistance gene *optrA* gene which could not be detected in the clinical NHS microbiology laboratory, MLST results supported PFGE finding that all the isolates were unrelated.

3.1.3.2 Routine microbiology results

For the agent linezolid to be sensitive, the EUCAST guidelines state that linezolid MIC should be ≤4 mg/L. The suspected outbreak isolates MICs were tested by Vitek 2 (bioMérieux Marcy L'toile, France) and were between 6-8 mg/L. Sensitivities can be seen below in heat map Table 3.9.

Table 3.9. Antibiograms of *optrA* gene positive *E. faecalis* patients

P	Year	AMX	SXT	NIT	TEC	TGC	VAN	CHL	DOX	ERY	PEN	LZD	TET	TMP
1	2016	S	S	S	S	S	S	R	R	R	R	R	R	I
2	2015	S	S	S	S	S	S	R	R	R	R	R	R	S
3	2014	S	S	S	S	S	S	R	R	R	R	R	R	R

P, patient; AMX, amoxicillin; SXT, trimethoprim-sulfamethoxazole; NIT, nitrofurantoin; TEC, teicoplanin; TGC, tigecycline; VAN, vancomycin; CHL, chloramphenicol; DOX doxycycline; ERY, erythromycin; PEN, penicillin; LZD, linezolid; TET, tetracycline; TMP, trimethoprim; ; R, resistant; I, intermediate; S, sensitive

3.1.3.3 Reference laboratory typing

Routine Reference laboratory typing by PFGE demonstrated that all the isolates were unrelated. Results of this are shown in Figure 3.4.

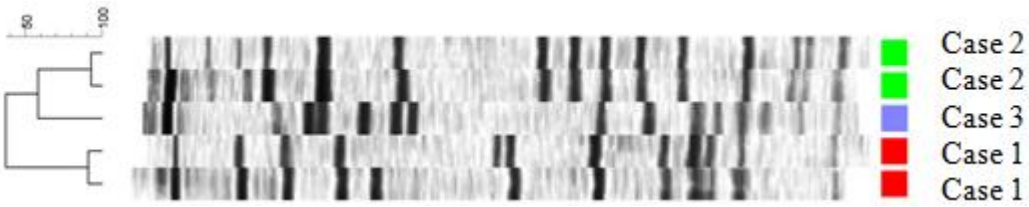


Figure 3.4. PFGE of *optrA* gene positive *E. faecalis* from tertiary hospital and community patients. PFGE results show isolates from different patients were unrelated. Data source- National Reference Laboratory (Antimicrobial Resistance and Healthcare Associated Infections Reference Unit (AMRHAI), Public Health England, Colindale

3.1.3.4 Whole genome sequencing results

MLST results supported PFGE findings. Isolates had various sequence type (ST) results: patient 1 (ST330); patient 2 (ST19); patient 3 (ST480).

3.1.4 *Listeria monocytogenes* outbreak

3.1.4.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of *L. monocytogenes* in a tertiary hospital. An overview of the epidemiological and IPC investigation carried out for this suspected outbreak is shown in Table 3.10. It can be seen that the upper control warning limit was not breached with these cases.

Table 3.10. Overview of the epidemiological and IPC investigation carried out for suspected *L. monocytogenes* outbreak

Healthcare facility	Tertiary hospital
Microorganism	*L. monocytogenes*
Description of suspected outbreak	• Two patients had a bacteraemia due to *L. monocytogenes.* Both patients were immunocompromised and had been admitted to different hospital wards and clinics at different times. • Routine typing identified the *Listeria* as serotype 4 (clonal complex 1). • Hospital kitchens had been inspected and no areas of concern were reported. No obvious links between the cases were identified in the community
IPCT hypothesis	The IPCT and HPT were unsure whether there had been transmission since this was one of the more common serotypes found in clinical isolates (1 in 6 invasive *Listeria* isolates are clonal complex 1).
Mode of spread	Contact –food
Case definition	Any patient identified to have a clinical sample which has *L. monocytogenes*
Total number of suspected patients	3
Outbreak specific control measures	Hospital kitchen inspections and closure until remedial action was undertaken
HIIAT Score	Red
Outbreak confirmed/not confirmed	Outbreak confirmed by WGS
WGS influence on IPC management	WGS results were produced in 12 days. Routine typing results took 8 days. WGS agreed with serotyping but gave much greater clarity that transmission had occurred in 2 patients who had a rare organism and common serotype justifying an intensive investigation carried out by the IPC team and Public Health
Impact of WGS	**Greater discrimination and enhanced 'alert organism' detection**-WGS revealed the two isolates were indistinguishable and therefore highly likely to be epidemiologically linked. This prompted further action to look

	again for a common link. The outbreak team repeated hospital kitchen inspections and as a result of this identified that the handling of salads and meat did not meet national recommendations and subsequently hospital catering facilities were temporary closed until remedial action was undertaken.

The SPC chart used in the investigation of this *Listeria* outbreak can be seen below in Figure 3.5. IPCT initially had suspicions initially that this may be an outbreak as the warning limit was reached on the SPC chart.

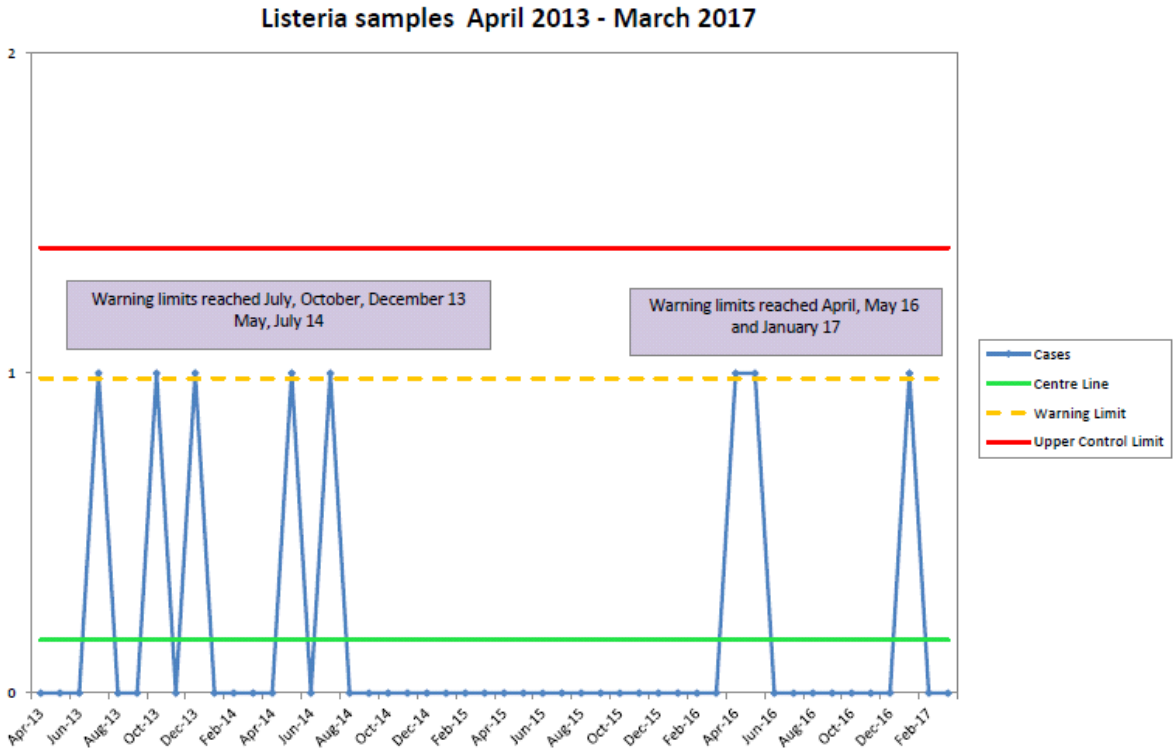


Figure 3.5. SPC chart of suspected *Listeria* outbreak in a tertiary hospital from April 2013-March 2017. The chart demonstrates a warning limit was reached. Produced by NHS Grampian IPC department using data from ICNet outbreak surveillance software.

3.1.4.2 Routine microbiology results

Antibiograms of the isolates are shown below in Table 3.11.

Table 3.11. Antibiograms of *Listeria* isolates

P	Date	LVX	TET	CLI	AMX	ERY	MXF	PEN
1	29/04/16	A	A	-	S	S	S	S
2	08/05/16	A	A	-	-	S	-	S
3	26/10/16	A	A	A	-	S	-	S

P, patient; LVX, levofloxacin; TET, tetracycline; CLI, clindamycin; AMX, amoxicillin; ERY, erythromycin; MXF, moxifloxacin; PEN, penicillin; A, active; ; R, resistant; I, intermediate; S, sensitive

3.1.4.3 Reference laboratory typing

All *Listeria* were identified as serotype 4. A summary table of Reference laboratory typing results is seen below in Table 3.12.

Table 3.12. *Listeria* serotyping results

P	Serotype	Clonal Complex
1	4	CC1
2	4	CC1
3	4	-

P, patient.

3.1.4.4 Whole genome sequencing results

Isolates of *L. monocytogenes* from the three patients underwent WGS and were mapped to the reference chromosome of strain F2365 which is a serotype 4b (genomic division II) cheese isolate from a Jalisco cheese outbreak of 1985 in California.(209) SNPs were called against this and patient 1 and patient 2 were found to be identical and likely to be epidemiological linked. Both patient 1 and 2 were found to be 164 SNPs different from the reference strain F2365. Patient 3 was found to be approximately 10,000 SNP different from patient 1 and 2.

3.1.5 Group A *Streptococcus* (GAS)

3.1.5.1 General Ward

3.1.5.1.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of Group A *Streptococcus* (GAS) on a single ward in a tertiary hospital. An overview of the epidemiological and IPC investigation carried out for this suspected outbreak is shown in Table 3.13.

Table 3.13. Overview of the epidemiological and IPC investigation carried out for suspected District Hospital ward GAS outbreak

Healthcare facility	District Hospital ward
Microorganism	GAS
Description of suspected outbreak	• Two patients developed bacteraemia due to GAS on the same ward. • The first patient developed hospital acquired pneumonia (HAP) and GAS was detected from blood cultures. They were treated with amoxicillin and improved. • The second patient was identified to have a blood culture positive with GAS 12 days later. These blood cultures were taken whilst patient one was still admitted to the ward. They had been diagnosed with an infected leg ulcer, treated with vancomycin and clindamycin and their clinical condition improved on this regime.
IPCT hypothesis	Transmission had occurred between two patients due to a breach in the HCW hand hygiene protocol.
Mode of spread	Hand contact
Case definition	Any patient identified to have a clinical sample with GAS.
Total number suspected patients	2
Outbreak specific control measures	Isolation, contact precautions
HIIAT Score	Green
Outbreak confirmed/not confirmed	Confirmed
WGS influence on IPC management	WGS results were not produced in real-time due to batching. They did not directly influence IPC management but gave greater insight into how the bacterial isolates were related.
Impact of WGS	**Greater discrimination** -WGS provided additional granularity to *emm* typing and confirmed that the isolates were ST28 and indistinguishable.

3.1.5.1.2 Routine microbiology results

Antibiograms of the GAS isolates are displayed in Table 3.14.

Table 3.14. Antibiograms of the GAS isolates

P	Date	Sample	PEN	CLI	SXT	DOX	ERY	LZD	LVX	VAN
1	13/03/2018	Blood culture	S	S	S	S	S	S	S	S
2	25/03/2018	Blood culture	S	S	S	S	S	S	S	S

P, patient; PEN, penicillin; CLI, clindamycin; SXT, trimethoprim-sulfamethoxazole; DOX, doxycycline; ERY, erythromycin; LZD, linezolid; LVX, levofloxacin; VAN, vancomycin; R, resistant; I, intermediate; S, sensitive

3.1.5.1.3 Reference laboratory typing

Both patients' blood culture isolates of GAS were typed using *emm* typing and found to be *emm* type 1.0.

3.1.5.1.4 Whole genome sequencing results

WGS confirmed that the isolates were ST28 and indistinguishable.

3.1.5.2 Maternity Unit

3.1.5.2.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of GAS in a maternity unit of a tertiary hospital. An overview of the epidemiological and IPC investigation carried out for this suspected outbreak is shown in Table 3.15.

Table 3.15. Overview of the epidemiological and IPC investigation carried out for suspected Maternity unit GAS outbreak

Healthcare facility/Patient	Tertiary hospital - maternity unit
Microorganism	GAS
Description of suspected outbreak	• Two patients were admitted to the same maternity unit and found to have GAS positive vaginal swabs. Patient 1 had been admitted to the unit and had delivered at home (unplanned) at 32 weeks and subsequently admitted to hospital where the placenta was delivered. The patient was given amoxicillin-clavulanic acid and gentamicin and after becoming septic clindamycin was added. • The second patient had a C-section 10 day's later and developed sepsis 12 days after this. CT showed endometritis, but no collection. This patient improved on cefuroxime, gentamicin and clindamycin.
IPCT hypothesis	Transmission of GAS on a ward between two patients due to non-complaint hand hygiene.
Mode of spread	Hand contact
Case definition	Any patient identified to have a clinical sample which has GAS.
Total number of suspected patients	2
Outbreak specific control measures	Isolation, contact precautions
HIIAT Score	Green
Outbreak confirmed/not confirmed	Ruled out
WGS influence on IPC management	WGS results were not produced in real-time due to batching. They did not directly influence IPC management but gave the IPC team a greater insight into how the bacterial isolates were related.
Impact of WGS	**Greater discrimination**- MLST supported *emm* typing and both techniques assisted the IPCT in ruling out transmission.

3.1.5.2.2 Routine microbiology results

Antibiograms of the GAS isolates are shown below in Table 3.16.

Table 3.16. Antibiograms of GAS isolates from a suspected outbreak in a Maternity unit

P	Sample	Date	PEN	CLI	AMC	DOX	ERY	LZD	LVX	VAN
1	Vaginal swab	26/03/16	S	S	S	S	S	S	S	S
2	Vaginal swab	05/04/16	S	S	S	S	S	S	S	S

P, patient; PEN, penicillin; CLI, clindamycin; AMC, amoxicillin-clavulanic acid; DOX, doxycycline; ERY, erythromycin; LZD, linezolid; LVX, levofloxacin; V, vancomycin; R, resistant; I, intermediate; S, sensitive.

3.1.5.2.3 Reference laboratory typing

Emm typing was carried out which revealed that both isolates were different. Patient 1 *emm* typing was 1.0 and patient 2 had an *emm* typing result of 28.

3.1.5.2.4 Whole genome sequencing results

MLST was carried out and confirmed that both isolates were unrelated and had different STs. The isolate from patient 1 was ST28, patient 2 had an isolate that was found to be ST458.

3.1.5.3 Midwife unit

3.1.5.3.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of GAS in a tertiary hospital. An overview of the epidemiological and IPC investigation carried out for this suspected outbreak can be seen in Table 3.17

Table 3.17. Overview of the epidemiological and IPC investigation carried out for suspected Midwife unit GAS outbreak

Healthcare facility	Tertiary hospital midwife unit
Microorganism	GAS
Description of suspected outbreak	• Two babies developed infection of their umbilicus with GAS after discharge from a labour suite. Both had been born within 8 days of each other and spent time on the same ward. • Both patients were treated with amoxicillin. Both babies were not in hospital at the same time. • One baby was discharged from the bay before the mother of other the baby was admitted. They were therefore in Labour suite on different days. • Different staff had looked after the babies and no staff had been reported to be unwell.
IPCT hypothesis	Transmission of GAS between the babies was unlikely
Mode of spread	Not applicable
Case definition	Any patient identified to have a clinical sample which has GAS.
Total number suspected patient	2
Outbreak specific control measures	NA
HIIAT Score	Green
Outbreak confirmed/not confirmed	Not an outbreak
WGS influence on IPC management	WGS results were not produced in real-time due to batching. They did not directly influence IPC management.
Impact of WGS	**Greater discrimination**- MLST confirmed that transmission had not taken place.

3.1.5.3.2 Routine microbiology results

Antibiograms of the isolates are shown below in Table 3.18.

Table 3.18. Antibiograms GAS isolates from patients admitted to a Midwife unit

P	Sample	Date	PEN	SXT	DOX	ERY	LVX
1	Umbilical swab	02/05/18	S	S	S	S	S
2	Umbilical swab	030/5/18	S	S	S	S	S

P, patient; PEN, penicillin; SXT, trimethoprim-sulfamethoxazole; DOX, doxycycline; ERY, erythromycin; LVX, levofloxacin; R, resistant; I, intermediate; S, sensitive

3.1.5.3.3 Reference laboratory typing

The reference laboratory carried out *emm* typing and patient 1 *emm* typing was 6.4 and patient 2 umbilical swab was 89.0.

3.1.5.3.4 Whole genome sequencing results

MLST was carried out which showed that the isolates were unrelated. The GAS from patient 1 was ST101 and the GAS from patient 2 was a ST382.

3.1.5.4 Care Home

3.1.5.4.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of GAS in a care home. An overview of the epidemiological and IPC investigation carried out for this outbreak is shown in Table 3.19.

Table 3.19. Overview of the epidemiological and IPC investigation carried out for suspected Care Home GAS outbreak

Healthcare facility	Care home
Microorganism	GAS
Description of suspected outbreak	• Two patients with invasive GAS were diagnosed 6 days apart (both had positive blood cultures) in neighbouring rooms of 11 bedded wing of a care home. • The first patient was diagnosed with aspiration and the following patient had periorbital cellulitis • A further patient was diagnosed with epiglottitis four days later.
IPCT hypothesis	Likely transmission based on epidemiology, care home setting and timeframe of presentation of infection likely due to direct hand contact and respiratory droplets.
Mode of spread	Hand contact and droplet
Case definition	Any care home resident identified to have a clinical sample positive for GAS
Total number suspected patients	3
Outbreak specific control measures	• Prophylaxis to all other residents in 11 bedded wing and all staff who had close physical contact with the patients • Staff cohorting • Leaflets for all other staff/visitors and increased vigilance amongst all staff for all residents in care home (around 50 residents in total). • Closure of care home with no acceptance of new admissions till the outbreak ceased - this resulted in delayed admission for one new resident.
HIIAT Score	Yellow
Outbreak confirmed/not confirmed	Outbreak confirmed
WGS influence on IPC management	WGS results were not produced in real-time due to batching. They did not directly influence IPC management.
Impact of WGS	**Greater discrimination**- WGS provided additional granularity to *emm* typing and confirmed that the first two isolates were indistinguishable supporting that transmission had taken place. Patient 3 was confirmed to be ST101 and unrelated to the other isolates.

3.1.5.4.2 Routine microbiology results

Antibiograms of the isolates are shown below in Table 3.20.

Table 3.20. Antibiogram of Care home GAS isolates

P	Sample	Date	PEN	CLI	SXT	DOX	ERY	LZD	LVX	VAN
1	Blood culture	07/07/2017	S	S	S	S	S	S	S	S
2	Blood culture	13/07/2017	S	S	S	S	S	S	S	S
3	Blood culture	17/07/2017	S	S	S	S	S	S	S	S

P, patient; PEN, penicillin; CLI, clindamycin; SXT, trimethoprim-sulfamethoxazole; DOX, doxycycline; ERY, erythromycin; LZD, linezolid; LVX, levofloxacin; VAN, vancomycin; S, sensitive. R, resistant; I, intermediate; S, sensitive.

3.1.5.4.3 Reference laboratory typing

Both patient 1 and 2 were found to have an *emm* typing result of 1 and patient 3 had an *emm* typing result of 89.

3.1.5.4.4 Whole genome sequencing results

MLST was carried out which showed that the first two isolates were indistinguishable. Patient 1 and 2 were found to be ST28 on MLST and indistinguishable on WGS. Patient 3 was confirmed to be ST101 and unrelated to the other isolates.

3.1.6 Borderline oxacillin resistant *Staphylococcus aureus* (BORSA)

3.1.6.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of borderline oxacillin resistant *Staphylococcus aureus* (BORSA) associated with a dermatology ward. An overview of the epidemiological and IPC investigation carried out for this suspected outbreak is shown in Table 3.21.

Table 3.21. Overview of the epidemiological and IPC investigation carried out for suspected BORSA outbreak

Healthcare facility	Dermatology ward tertiary hospital
Microorganism	BORSA
Description of suspected outbreak	• In June 2016, an outbreak of a rare spa type, t10939 of *S. aureus* was identified by the Scottish National MRSA Reference Laboratory in NHS Tayside. This type of *S. aureus* had 0.01% prevalence worldwide. • Since September 2015 thirty-five *S. aureus* isolates were identified as t10939. • Outbreak investigations identified complex epidemiological links between dermatology inpatient (ward) and outpatient clinical areas such as phototherapy. • The majority of patients had chronic skin conditions with frequent admissions to hospital and/or outpatient attendance for treatment.
IPCT hypothesis	Suspected transmission based on epidemiology, ward and community setting, patient clinical features e.g. dermatology and timeframe of presentation of infections/colonisations. Transmission due to breach in hand hygiene and contamination of the environment with skin scale shedding.
Mode of spread	Hand contact and contamination of the environment
Case definition	Any patient identified to have a clinical sample which has BORSA
Total number suspected patients	35
Outbreak specific control measures	• Contact precautions, including isolation. • Audited compliance of hand hygiene, the environment, and observations of practice. • Enhanced environmental cleaning with chlorine based agent. • Environmental sampling - areas sampled included top side of dressing trolley, showerhead, mattress cover, dustpan, patient chair, tympanic probe, ward floor, toilet seat, bath tap, and plug hole. • Development of patient screening and decolonisation

	protocols. • Patients with BORSA were identified on re-admission to hospital by an alert in the patient administration system and are placed appropriately within the clinical setting. • Closure of ward and relocation to a new site.
HIIAT Score	Yellow
Outbreak confirmed/not confirmed	Confirmed
WGS influence on IPC management	WGS results were produced in real-time throughout the outbreak. Sequencing results were presented at outbreak meetings. As WGS identified that isolates were closely related increased efforts were made to contain the outbreak with a range of layered mitigations.
Impact of WGS	**WGS was used to investigate a new/unusual resistance mechanism** **Greater discrimination**-WGS identified 14 different STs (ST59/1207/8/30/45) and in total 50 isolates were found to be ST188. WGS provided additional granularity to PFGE and confirmed amongst the ST188 isolates that there were three clusters in total with a maximum genetic distance between any two isolates of 45 SNPS. **Streamlining testing-** WGS could have identified isolates that were BORSA and part of the cluster cutting down on unnecessary additional testing.

A timeline of actions taken by the IPCT can be seen in Figure 3.6. The total number of new t10939 isolates per month from August 2015- June 2018 can be seen in Figure 3.7 and bar chart of individuals and their skin conditions can be seen in Figure 3.8.

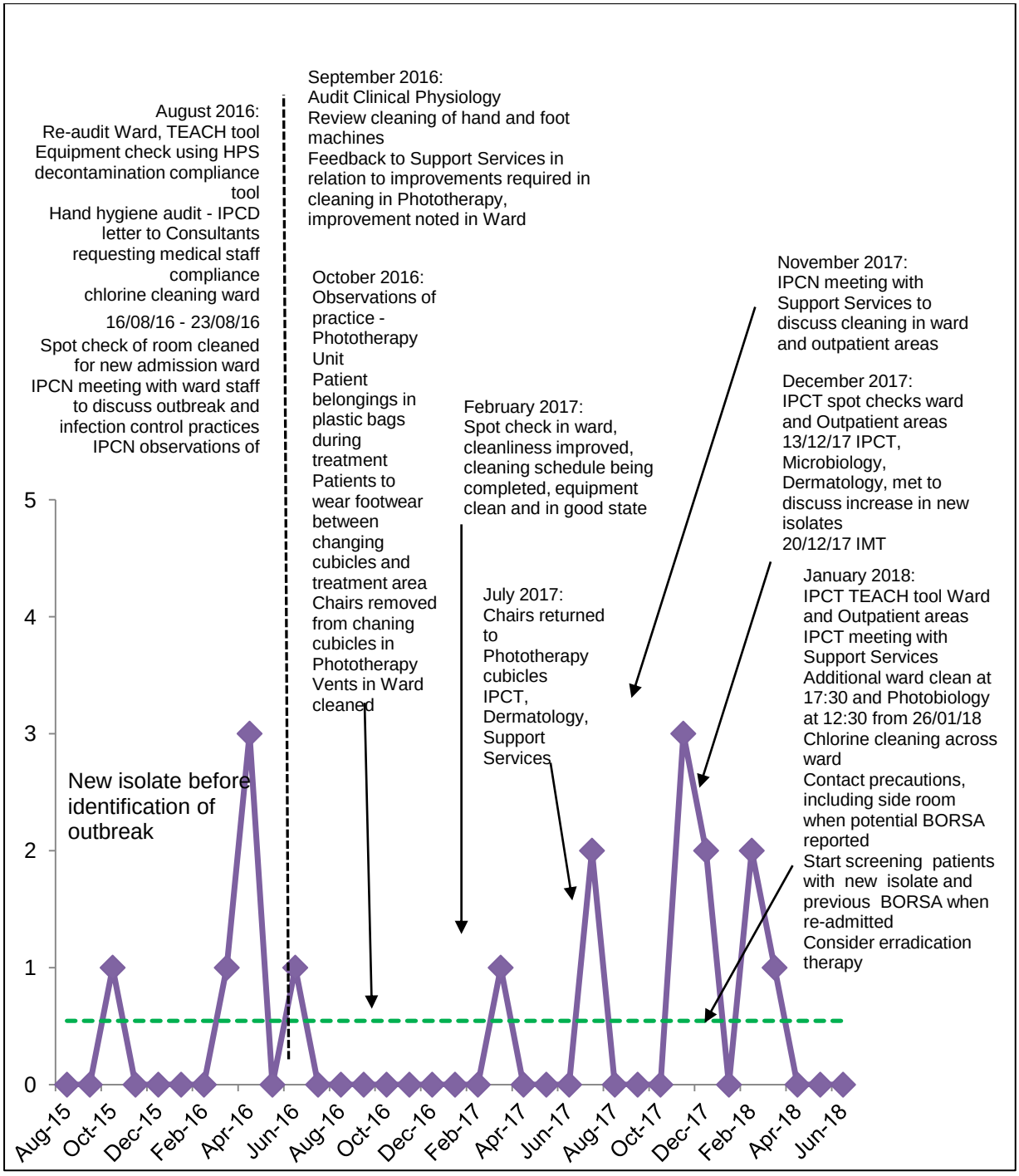


Figure 3.6. New isolates of hospital acquired BORSA associated with a dermatology unit August 2015 – June 2018. Timeline shows layered mitigation measures used to control the outbreak. Produced by NHS Tayside IPC department using data from ICNet outbreak surveillance software.

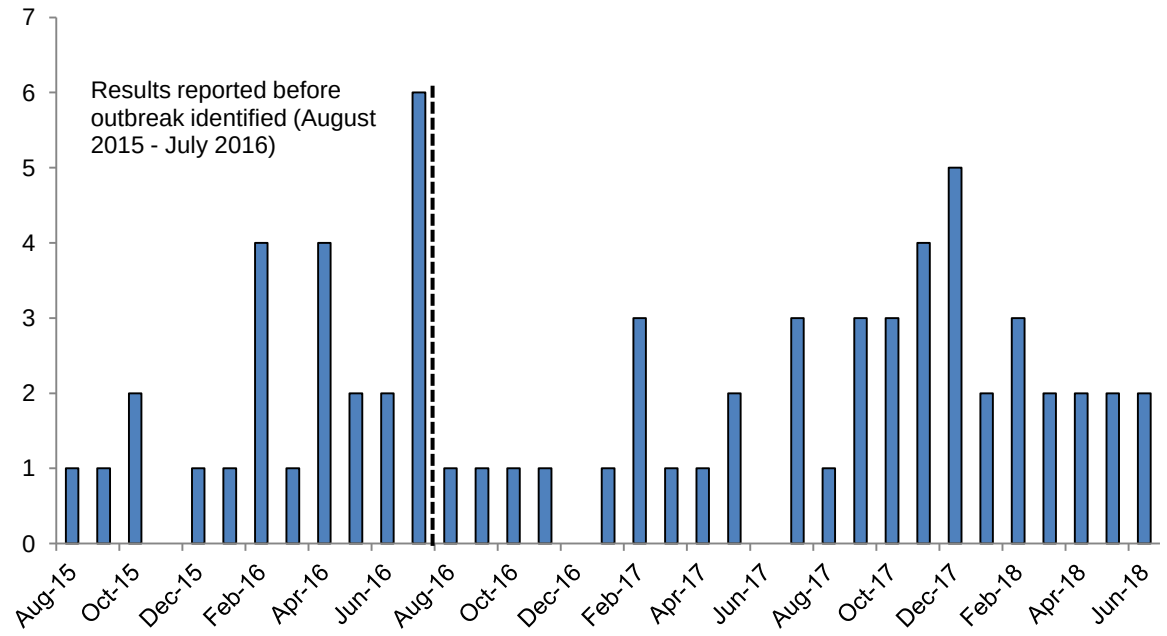


Figure 3.7. Number of new t10939 isolates per month (all patients) from August 2015-June 2018. Produced by NHS Tayside IPC department using data from ICNet outbreak surveillance software.

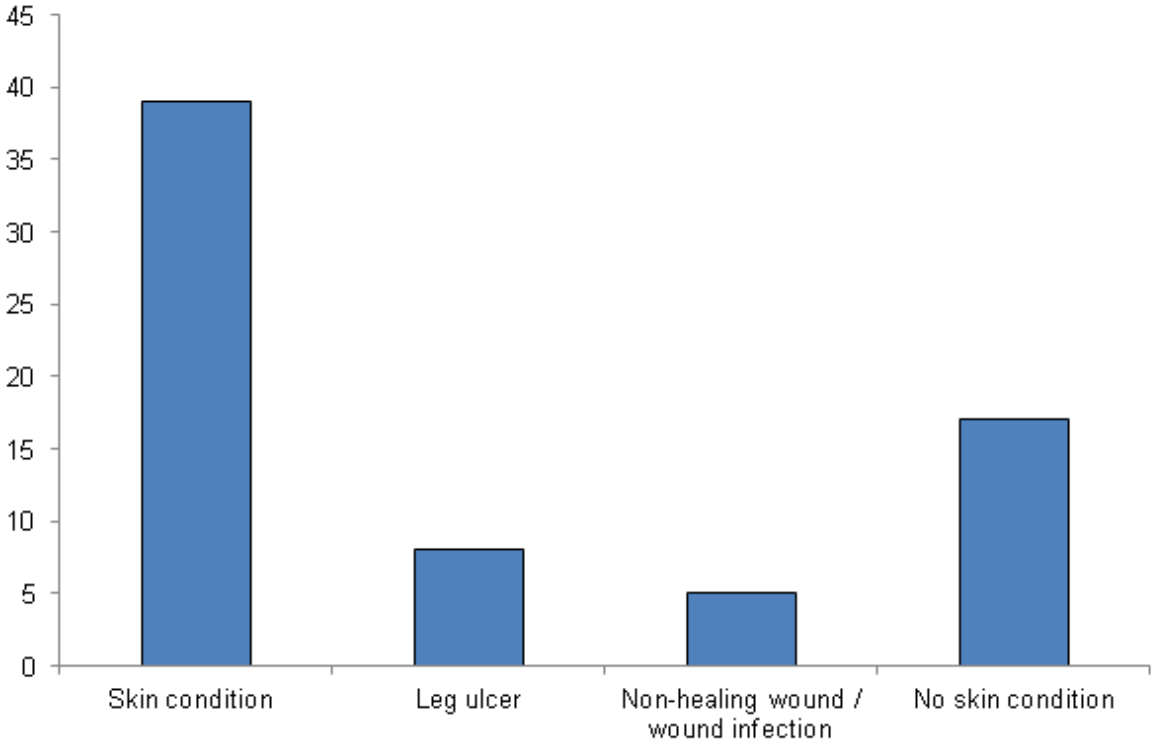


Figure 3.8. Count of dermatology patients with new isolate t10939 and patients clinical features confirmed from August 2015 – June 2018. Produced by NHS Tayside IPC department using data from ICNet outbreak surveillance software.

3.1.6.2 Routine microbiology results

In total 35 BORSA isolates were identified as t10939 since September 2015. On testing at the reference laboratory all were found to be *mecA* gene negative. In total 24 (69%) had an oxacillin MIC greater than 2μg/ml. They were also reported as. PVL, TSST, ETA and ETB toxin genes negative.

3.1.6.3 Reference laboratory typing

Initially PFGE analysis was carried out by the Scottish National MRSA Reference Laboratory (figure 1). This revealed that isolates were closely related with little variation between more recent isolates.(210) There were occasional band differences reflecting loss/gain of DNA as seen in Figure 3.9.

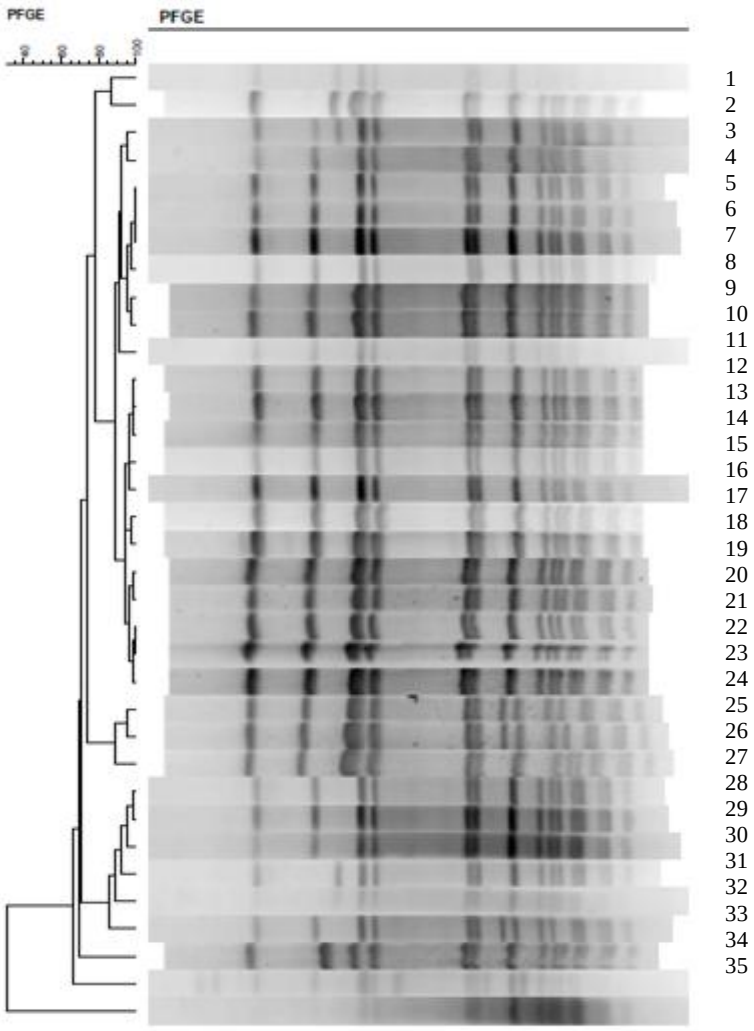


Figure 3.9. BORSA isolates PFGE profiles carried out at the Scottish National MRSA Reference Laboratory. PFGE shows bacterial isolates were closely related with little variation. Data source -Scottish MRSA Reference Laboratory.

3.1.6.4 Whole genome sequencing results

Following on from this WGS was carried out on a total of 64 samples (isolated dated from August 2015 to May 2017). Of these, 14 were of different STs (ST59/1207/8/30/45) and in total 50 isolates were found to be ST188 with repeat samples received from 8 patients. On comparison of the ST 188 isolates it was identified that there were three clusters in total with a maximum genetic distance between any two isolates of 45 SNPS.

3.2 Results from utilising WGS for Gram Negative outbreaks

3.2.1 Carbapenamase-producing *Enterobacterales* (CPE)

3.2.1.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of carbapenamase-producing *Enterobacterales* (CPE) in a tertiary hospital. An overview of the epidemiological and IPC investigation carried out for this suspected outbreak is shown in Table 3.22.

Table 3.22. Overview of the epidemiological and IPC investigation carried out for suspected CPE outbreak

Healthcare facility/hospital	Intensive Care Unit and Renal Ward
Microorganism	CPE
Description of suspected outbreak	• Three patients were identified to be colonised or infected by *Klebsiella pneumoniae* carbapenemase (KPC)-producing *K. pneumoniae* (KPC-KP) resistant to meropenem and ertapenem. • Patient A (index patient) had a positive rectal screen four days after admission (7/9/17) taken as part of a national screening protocol as the patient had been in hospital abroad. They had been admitted into a side room. The patient also attended the renal dialysis unit (RDU) and also used a side room there. They had polycystic kidney disease and did not require antibiotic treatment • Patient B was found to be positive for CPE on midstream urine two months later (07-Nov-2017). They were also present on the renal ward at various times. They were treated with doxycycline and a course of ceftazidime/avibactam • Patient C wound swab of an ankle was positive (11/11/17) for CPE four days after Patient B whilst on the renal ward. Patient C was treated with tigecycline and fosfomycin. They had been in the ICU where wound swabs were negative. Patient B and C had both been in renal side room 2 however patient A had never spent time there.
IPCT hypothesis	• Suspected transmission based on epidemiology, crossover of patients and timeframe of presentation of infections/colonisations. Transmission due to non-compliance with hand hygiene protocols and potential contamination of the environment.
Mode of spread	Hand contact and environmental contamination in the ward

Case definition	Any patient with a KPC-KP confirmed from any clinical or screening sample.
Total number suspected patients	Three
Outbreak specific control measures	• Hand hygiene audit score was 85% support was given to healthcare staff to improve this. • Enhanced cleaning and a deep clean of the ward • A mapping exercise determined all patients who had been in contact with patient A. Contact screening was carried out on 55 contracts (patient is first admission) and 114 (patient is second admission). • HCWs in the satellite units, managers, relatives and HPS were informed of the situation.
HIIAT Score	Yellow
Outbreak confirmed/not confirmed	Confirmed
WGS influence on IPC management	WGS result turnaround time was 17 days. Routine typing took 8 days. This data was fed back to the IPCT. WGS and VNTR results were concordant and phylogenetic analysis revealed that isolates were closely related giving the IPCT greater confidence transmission had occurred recently and screening of contacts was warranted.
Impact of WGS	**Streamline testing**- Use of WGS could have streamlined the microbiology identification of CPE in the clinical microbiology laboratory removing the need for repeat and unnecessary PCR tests. WGS also identified that patient C sample was mixed with a ST3 *E. coli*. This had not been detected by the clinical microbiology laboratory. **Greater discrimination-** WGS revealed that the isolates belonged to ST258. Pairwise SNP differences were calculated from the whole genome alignment showing that patients A and B were 9 SNPs apart, patients B and C 3 SNPs apart and patient A and C being 9 SNPs apart.

The timeline of patient hospital ward transfers can be seen in Figure 3.10. It can be seen that the patients spend most of their time on the renal ward (ward 22 in dark yellow).

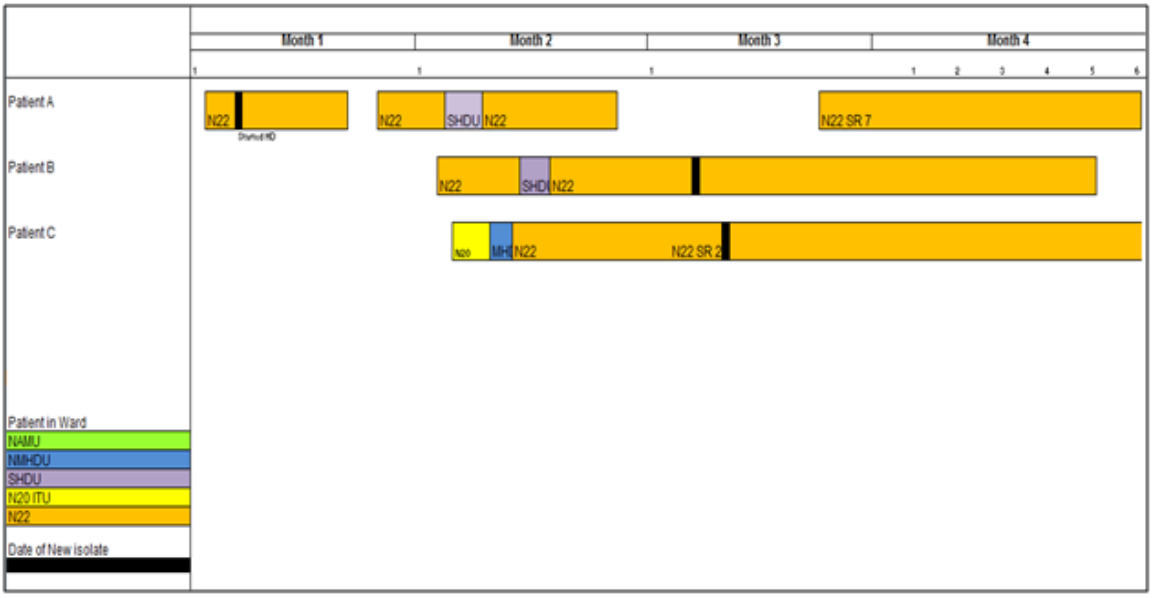


Figure 3.10. Timeline of CPE positive patient transfers during the CPE outbreak. Patients were admitted to the renal ward (ward 22 in dark yellow) for the majority of their admission. Produced by NHS Tayside IPC department using data from ICNet outbreak surveillance software.

The timeline of patient bed movements can be seen in Figure 3.11. It can be seen that two patients (B and C) spent time in bay one at the same time (dark yellow).

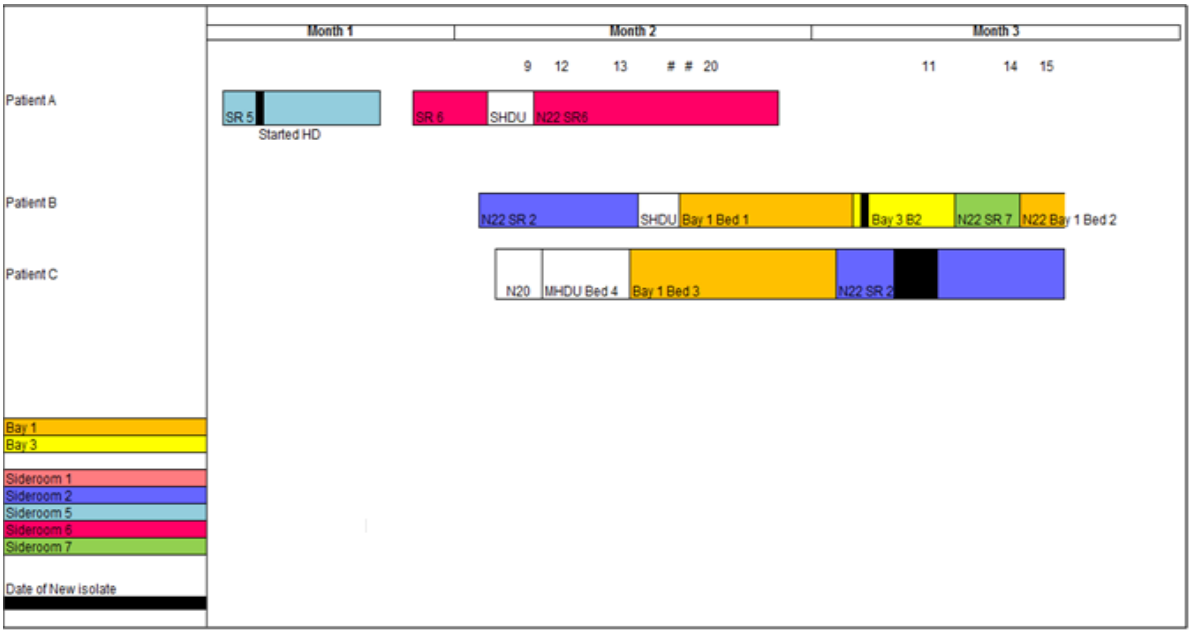


Figure 3.11. Timeline of CPE positive patient bed movements during the CPE outbreak. Patients (B and C) spent time in bay one at the same time (dark yellow). Produced by NHS Tayside IPC department using data from ICNet outbreak surveillance software.

3.2.1.2 Routine microbiology results

Isolates were found to have a remarkably similar susceptibility patterns (Patient A rectal screen, Patient B Midstream urine, Patient C wound swab) after testing on Vitek 2 (bioMérieux Marcy L'toile, France). Table 3.23 shows the antibiograms of the CPE isolates.

Table 3.23. Antibiograms of CPE isolates

P	AMC	CXM	GEN	TZP	CIP	TMP	AMX	TMO	CAZ	ATM	ETP	MEM	AMK	TOB	TGC	C/T	CZA	CST	DOX	FOF
A	R	R	S	R	R	R	R	R	R	R	R	R	R	R	S	R	S	S	-	-
B	R	R	S	R	R	R	R	R	R	R	R	R	R	R	R	R	S	S	S	-
C	R	R	S	R	R	R	R	R	R	R	R	R	R	R	S	-	-	R	R	S

P, patient; AMC, amoxicillin-clavulanic acid; CXM, cefuroxime; GEN, gentamicin; TZP, piperacillin-tazobactam; CIP, ciprofloxacin; TMP, trimethoprim; AMX, amoxicillin; TMO, temocillin; CAZ, ceftazidime; ATM, aztreonam; ETP, ertapenem; MEM, meropenem; AMK, amikacin; TOB, tobramycin; TGC, tigecycline; C/T, ceftolozane-tazobactam; CZA, ceftazidime-avibactam; CST, colistin; DOX, doxycycline; FOF, Fosfomycin; R, resistant; I, intermediate; S, sensitive

3.2.1.3 Reference laboratory typing

All isolates were confirmed to be *K. pneumoniae* blaKPC non-metallo-carbapenemase gene positive at the Scottish AMR Satellite Reference Laboratory and AMRHAI. Further testing by multiplex PCR for serotype specific targets for capsular types K1, K2, K5, K54 and K57 associated with pathogenicity or invasive disease at the AMRHAI were negative. Virulence factors associated with invasive disease such as regulators of mucoviscosity -rmpA and rmpA2 and also wcaG (a capsular fucose synthesis gene associated with capsular types K1, K16, K54 and K58) were also sought and were found to be negative.(211) (212) (213). VNTR analysis of the isolates revealed that all three were related corresponding to the ST258 lineage associated with KPC. A VNTR profile 3,2,2,13,2,1,3,3,3,3,3, was given. Repeat sampling from patient C revealed that tigecycline and fosfomycin resistance had developed on treatment.

3.2.1.4 Whole genome sequencing results

WGS and VNTR results were concordant. WGS also showed that the three isolates belonged to ST258. Phylogenetic analysis was performed in order to resolve the fine-scale relationship between outbreak isolates and explore epidemiological links between them. Pairwise SNP differences were calculated from the whole genome alignment showing that patients A and B were 9 SNPs apart, patients B and C 3 SNPs apart and patient A and C being 9 SNPs apart. WGS also identified that patient C sample was mixed with a ST3 *E. coli* which had not been detected by the clinical microbiology laboratory.

3.2.2 Pseudomonas aeruginosa

3.2.2.1 Adult ICU

3.2.2.1.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of *P. aeruginosa* in an Adult ICU of a tertiary hospital. An overview of the epidemiological and IPC investigation carried out for this suspected outbreak is shown in Table 3.24.

Table 3.24. Overview of the epidemiological and IPC investigation carried out for suspected *P. aeruginosa* outbreak

Healthcare facility	ICU
Microorganism	*P. aeruginosa*
Description of suspected outbreak	Six patient isolates and six environmental samples positive for *P. aeruginosa* over two-year period
IPCT hypothesis	Transmission likely based on epidemiology, clinical setting, and timeline of positive patients. Due to breakdown in hand hygiene and contamination of water supply and fixtures.
Mode of spread	Water, hand contact
Case definition	A patient with a sample positive with *P. aeruginosa* resistant to imipenem isolated from a patient admitted to ICU since 2012.
Total number suspected patients	Six
Outbreak specific control measures	• IPCT visited the ward and gave advice as per the national guidance for the disposal of wastewater. Fluids such as endotracheal aspirate (ETA) and bed bath water had been discarded via wash-hand basins (WHBs). Patient temperature probes ventilator flow sensors had been decontaminated in a sink with hot soapy water. The decontamination sink was also identified to have crusting on the taps which were later replaced. • To become compliant with NHS building regulations taps with flow straighteners were removed and sink basins were replaced to remove overflow drains. The ice machine flexible hose was replaced by a water regulations advisory scheme (WRAS)-approved hose to reduce biofilm formation and localized cleaning of all affected outlets was performed. An increased flushing regimen was introduced to remove biofilm. • An external contractor sampled water (pre- and post-flush samples) from 14 water outlets in the ICU for *P. aeruginosa.* Monitoring swabs were also taken from 11 water outlet drains on the same day domestic service room WHB, Bed 7 WHB, Bed 8 WHB, kitchen sink, kitchen drinking water tap, domestic service room sink, ventilator room sink, ICU entrance WHB, ward area WHB 1, ward

	area WHB 2, Bed 4 WHB. Patient and environmental isolates were typed. • Remediation works were successful with no growth of *Pseudomonas* species on repeat testing of outlets and water.
HIIAT Score	Red
Outbreak confirmed/not confirmed	Confirmed
WGS influence on IPC management	WGS did not directly influence IPC investigation of the suspected outbreak as isolates were tested in retrospect. WGS had no direct clinical impact.
Impact of WGS	**Greater discrimination and streamlining of testing -** Conventional methods, antibiotic susceptibility, VNTR and PFGE, grouped the isolates in different ways. VNTR and PFGE identified two patients who were part of the outbreak but identified several false positive environmental links. WGS provided the necessary resolution to be used in place of these two typing techniques streamlining the outbreak investigation.

A map of the unit is shown in Figure 3.12.

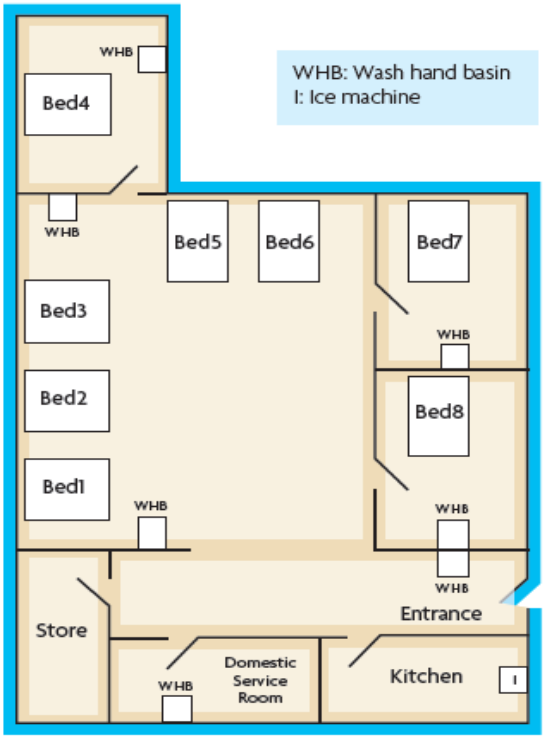


Figure 3.12. Floor plan of the ICU in NHS Tayside in which the *P. aeruginosa* outbreak occurred (drawn by Benjamin Parcell). Graphic map created by Wai-Lum Sung Graphic Designer at the University of Aberdeen.

3.2.2.1.2 Routine microbiology results

Table 3.25 shows results of microbiological detection for *P. aeruginosa* in water (pre- and post-flush samples) samples from water outlets in the ICU.

Table 3.25. *P. aeruginosa* count in water from water outlets in the ICU

Source	*Pseudomonas* count (cfu/ml) pre-flush samples	*Pseudomonas* count (cfu/ml) post-flush samples
Ice Machine	>100	>100
Domestic Service Room WHB	37	1
Bed 7 WHB	>100	28
Bed 8 WHB	41	0
Kitchen Sink	0	0
Kitchen Drinking Water Tap	0	0
Kitchen Hydroboil	0	0
Domestic Service Room Sink	0	0
Ventilator Room Sink	0	0
ICU Entrance WHB	0	0
Chilled Drinking Water Dispenser	0	0
Ward Area WHB 1	0	0
Ward Area WHB 2	0	0
Bed 4 WHB	0	0

All isolates were confirmed to be *P. aeruginosa* and five isolates were found to have an indistinguishable susceptibility pattern (Patient A Abdominal Drain Fluid, Patient B endotracheal aspirate (ETA), Patient D ETA, Patient E ETA, Bed 8 WHB water) after testing on Vitek 2 (bioMérieux, Marcy L'Etoile, France). The *P. aeruginosa* antibiograms can be seen in Table 3.26.

Table 3.26. Antibiotic resistance profile of patient and environmental *P. aeruginosa* isolates

Source	GEN	CIP	TZP	CAZ	MEM	TOB	ATM	IMP	AMK	CST
Bed 4 *	S	S	S	-	I	S	I	I	S	S
Bed 7	-	-	-	-	S	-	-	S	-	-
Bed 8	S	S	S	S	I	S	I	R	S	S
DSR	S	S	S	S	S	S	I	S	S	S
ICU *	S	S	S	S	S	S	I	R	S	S
Ice Machine	S	S	S	S	S	S	I	S	S	S
Patient A*	S	S	S	S	I	S	I	R	S	S
Patient B*	S	S	S	S	I	S	I	R	S	S
Patient C*	S	S	S	S	R	S	I	R	S	S
Patient D	S	S	S	S	I	S	I	R	S	S
Patient E*	S	S	S	S	I	S	I	R	S	S
Patient F	S	S	S	S	S	S	I	R	S	S

GEN, gentamicin; CIP, ciprofloxacin; TZP, piperacillin-tazobactam; CAZ, ceftazidime; MEM, Meropenem; TOB, tobramycin; ATM, aztreonam; IMP, imipenem; AMK, amikacin; CST, colistin; R, resistant; I, intermediate; S, sensitive; * refers to *OprD* porin loss.

3.2.2.1.3 Reference laboratory typing

VNTR analysis of the isolates from the ICU identified that 6 of the isolates belonged to a cluster of related profiles, which included Patients B and D and the four environmental isolates from Bed 8 WHB water, Bed 4 WHB water outlet drain, Bed 7 WHB water outlet drain, Kitchen Ice Machine water as displayed in Table 3.27.

Table 3.27. VNTR profiles of *P. aeruginosa* isolates from the ICU

Source	Date of Sampling	VNTR
Bed 4 WHB Water Outlet Drain	16/05/2013	12,2,1,5,5,2,4,5,11
Bed 7 WHB Water Outlet Drain	16/05/2013	12,2,1,5,5,2,4,5,11
Bed 8 WHB Water	16/05/2013	12,2,1,5,5,2,4,5,12
Domestic Service Room WHB Water	16/05/2013	12,3,6,3,1,4,14,5,10
ICU Entrance WHB Water Outlet Drain	16/05/2013	12,3,-,3,1,4,14,5,10
Ice Machine Water	16/05/2013	12,2,1,5,5,2,4,5,14
Patient A Abdominal Drain Fluid	11/03/2012	12,6,7,5,3,4,8,1,11
Patient B ETA	21/09/2012	12,2,1,5,5,2,4,5,12
Patient C ETA	04/01/2013	11,2,6, ,3,6,6,6,12
Patient D ETA	15/04/2013	12,2,1,5,5,2,4,5,12
Patient E ETA	11/05/2013	12,4,-,-,3,1,6,4,13
Patient F ETA	05/05/2013	12,2,-,3,2,2,-,5,6

WHB, wash hand basin; ETA, endotracheal aspirate.

All of these isolates had VNTR profiles that were similar to the PA14 clone. The close relationship of these isolates in the PA14 cluster suggested that these isolates may be part of an outbreak. In contrast the isolates from patients A, C, E and F had distinct VNTR profiles both from one another and the PA14 cluster, and also from the remaining environmental samples suggesting that these were unlinked and therefore could be ruled out of the outbreak.

PFGE was used to compare the PA14 cluster isolates. Analysis of the banding pattern divided the isolates into three distinct subtypes designated NINE04PA-1 (Bed 8 WHB water, Patient D ETA, Patient B ETA), NINE04PA-1' (Bed 4 and 7 WHB water outlet drain) and NINE04PA-1" (Kitchen Ice Machine). There were clear and definite band differences between the ice machine isolate and the patient isolates (Figure 3.13)

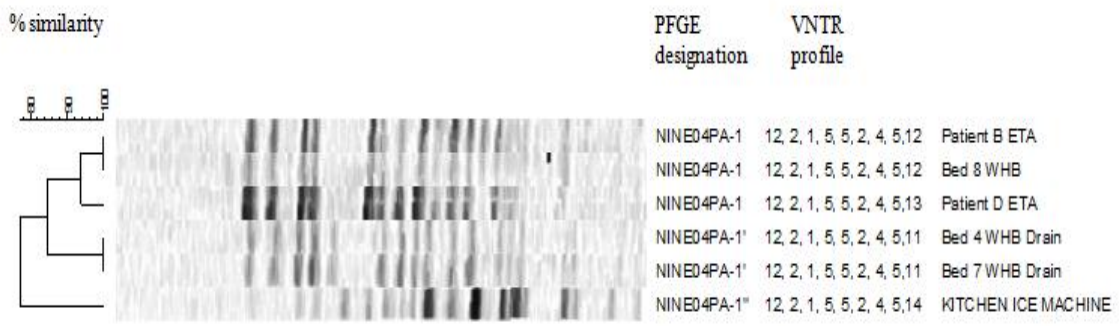


Figure 3.13. PFGE carried out at the reference laboratory comparing the PA14 cluster isolates. Analysis of the banding pattern divided the isolates into three distinct subtypes. These included NINE04PA-1 (Bed 8 WHB water, Patient D ETA, Patient B ETA), NINE04PA-1' (Bed 4 and 7 WHB water outlet drain) and NINE04PA-1" (Kitchen Ice Machine). Data source - National Reference Laboratory (Antimicrobial Resistance and Healthcare Associated Infections Reference Unit (AMRHAI), Public Health England, Colindale

3.2.2.1.4 Whole genome sequencing results

Figure 3.14 shows a phylogenetic tree was built with core SNPs identified by mapping to the PAO1 reference genome.

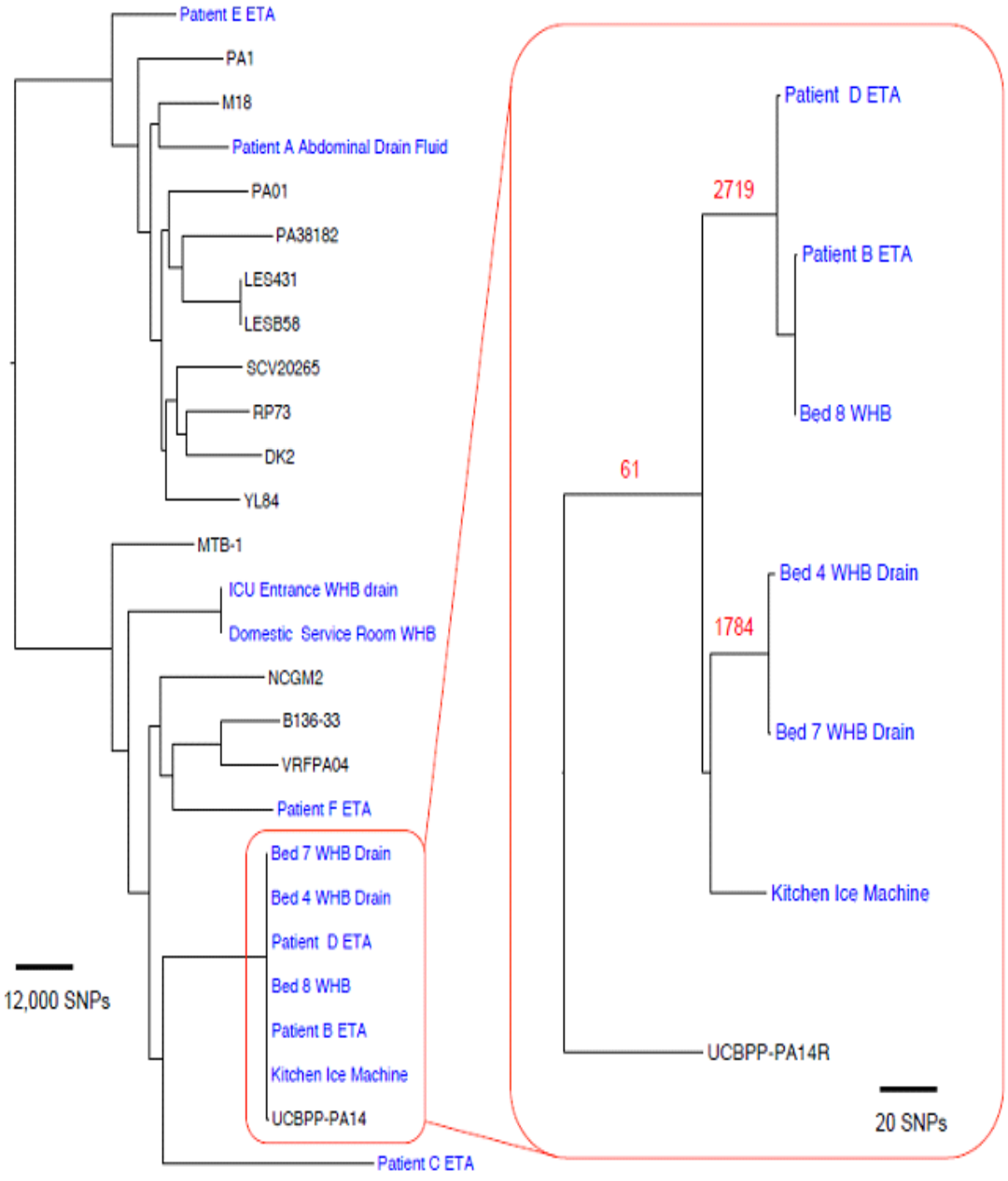


Figure 3.14. Phylogenetic analysis of *P. aeruginosa* ICU outbreak isolates. The isolates formed two separate clusters. In cluster one Patient B ETA and Bed 8 WHB isolates were indistinguishable. They were found to differ from the Patient D ETA sample by 4 SNPs strongly supporting transmission between the Bed 8 WHB and patients D and B. Data source - Infection and Global Health Research Division, School of Medicine, University of St Andrews

The box on the right of Figure 3.14 contains a maximum likelihood phylogeny of the PA14 clone ICU isolates. Reads were mapped to the PA14 reference genome of UCBPP-PA14R. The tree was built with core SNPs, SNPs identified in regions that had arisen by were excluded recombination (red text shows the SNPs associated with recombination). Fifteen *P. aeruginosa* reference isolates from the EMBL nucleotide database were included to show the diversity within the species. Overall 182,476 SNP sites were identified revealing a diverse population structure. The cluster of isolates identified by VNTR as belonging to PA14 clone formed a distinct clade in the tree. The WGS reads were remapped to reference chromosome of UCBPP-PA14 as this isolate is genetically closer to the outbreak isolates than PAO1 providing increased coverage and resolution. Phylogenetic analysis of this remapped data showed that the isolates formed two separate clusters and a further outlier that was each distinguish by over 1000 SNPs (cluster one containing Patient D ETA, Patient B ETA and Bed 8 WHB, cluster two Bed 4 WHB Drain and Bed 7 WHB Drain distinguished by 4515 SNPs and the Kitchen Ice Machine isolate distinguished by 1852 SNPs). In cluster one Patient B ETA and Bed 8 WHB isolates were indistinguishable. They were found to differ from the Patient D ETA sample by 4 SNPs strongly supporting transmission between the Bed 8 WHB and patients D and B.

3.2.2.2 Cystic fibrosis clinic tertiary hospital

3.2.2.2.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of *P.aeruginosa* in a CF clinic in a tertiary hospital. An overview of the epidemiological and IPC investigation carried out for this suspected outbreak is shown in Table 3.28.

Table 3.28. Overview of the epidemiological and IPC investigation carried out for suspected *P. aeruginosa* CF clinic Tertiary hospital

Healthcare facility/hospital	**CF clinic Tertiary hospital**
Microorganism	***P. aeruginosa***
Description of suspected outbreak	• Two CF patients were identified to be positive for *P.aeruginosa* from sputum on the same day (both isolates had similar antibiograms). • Patient 1 was a child who attended the paediatric outpatient clinic and had never attended the respiratory clinic. Patient 2 had been seen in the respiratory clinic (clinic 2). These patients have never crossed paths in any clinic or ward. A further 5 patients identified to have similar *Pseudomonas* antibiograms. Patient 1 and 2 on VNTR appeared to share the same strain. VNTR showed while patient 4, 5, 6 and 7 had isolates with similar profiles, there were differences at either the 7^{th} or 9^{th} locus (or both) for each of these patients.
IPCT hypothesis	Transmission had occurred between patients 1 and 2 due to break down in hand hygiene and staff working in the respiratory clinic and paediatric outpatient clinic. There could have been contamination of shared equipment in a room in which patients expectorated.
Mode of spread	Breach in hand hygiene protocols potential contamination of shared equipment
Case definition	Any patient with a meropenem resistant *P. aeruginosa* confirmed from any clinical or screening sample.
Total number of suspected patients	7
Outbreak specific control measures	• Audit and training for hand hygiene • Development of a cleaning protocol for equipment • Change in practice – patients given advice on how to expectorate and also given separate area to do this. Staff training and cleaning schedule implemented.
HIIAT Score	Green
Outbreak confirmed/not confirmed	Confirmed
WGS influence on IPC management	WGS results were not produced in real-time due to batching. They did not directly influence IPC management but gave the

	IPC team a greater insight into how the bacterial isolates were related.
Impact of WGS	**Greater discrimination**- WGS provided additional granularity to VNTR confirming that both patients isolates were ST1714 and sharing 53 SNP suggesting transmission or shared isolates, with possible common ancestor back in time. Patient to patient transmission between Patients 3 and 4, in clinic was ruled out as both patients' isolates were unrelated by MLST (ST111 and ST2140). The isolates were not related to any other CF patient's isolates. Patient isolates from three patients: patient 6 (2 isolates, 5/11/2015), patient 5 (1 isolate, 6/11/2015) and patient 7 (1 isolate, 15/11/2015) had similar VNTR. WGS found all patients' isolates to be ST500 sharing 2194 SNPs therefore it was possible there was a common ancestor or shared pool of isolates.

Figure 3.15 shows the SPC chart of this suspected *P. aeruginosa* outbreak.

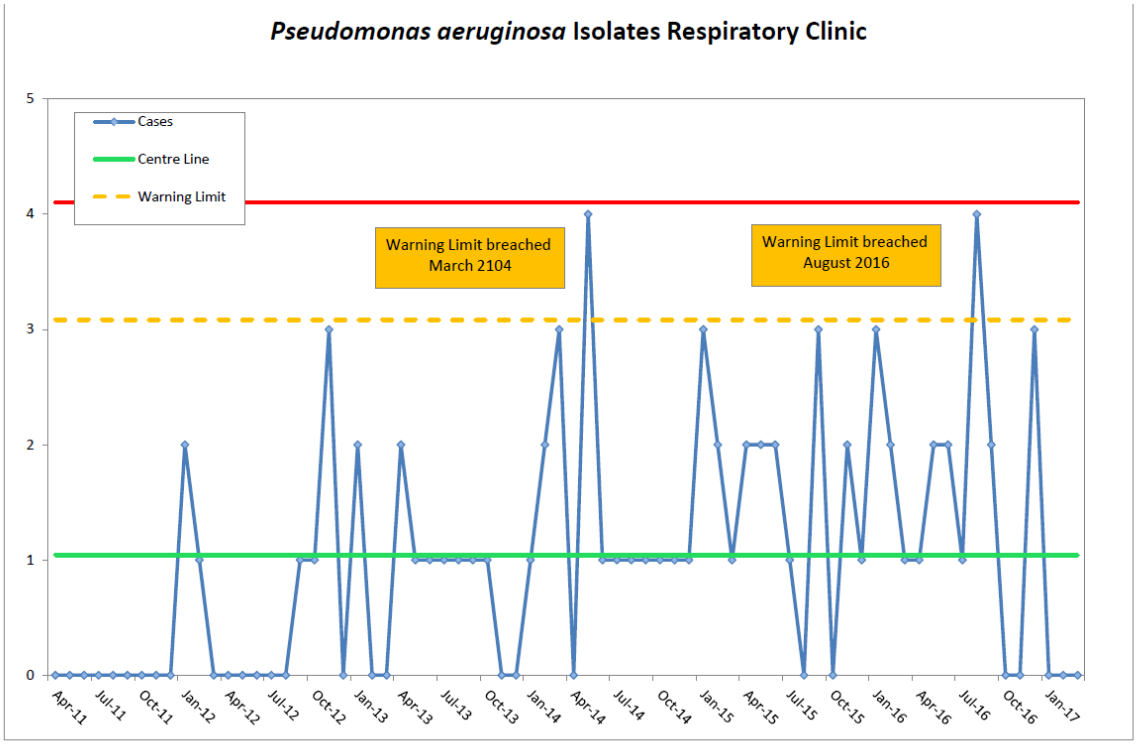


Figure 3.15. SPC chart of suspected *P. aeruginosa* isolates with similar antibiograms involving the respiratory clinic from April 2011-January 2017. Produced by NHS Grampian IPC department using data from ICNet outbreak surveillance software.

3.2.2.2.2 Routine microbiology results

Antibiograms of the CF clinic Tertiary hospital isolates are shown in Table 3.29.

Table 3.29. Antibiograms of CF clinic *P. aeruginosa* isolates

P	Date	Source	Strain	AMK	CIP	GEN	LVX	MEM	TOB	CST	CAZ	TZP
1	10/11/15	Paediatric OPD	1	R	R	R	R	R	R	S	S	S
2	10/11/15	Clinic C	1	R	R	R	R	R	R	S	S	S
	10/11/15	Clinic C	2	R	S	R	R	R	R	S	R	R
	10/11/15	Clinic C	3	R	S	R	S	R	R	S	S	S
	10/11/15	Clinic C	4	R	R	R	R	R	R	S	R	R
3	29/10/15	Clinic C	1	R	R	R	R	S	R	S	R	R
	29/10/15	Clinic C	2	R	R	R	R	R	R	S	R	R
4	02/11/15	Clinic C	1	S	R	S	R	R	S	S	R	R
	02/11/15	Clinic C	2	R	R	S	R	R	S	S	R	R
5	06/11/15	Clinic C	1	S	R	S	R	R	S	S	R	S
	06/11/15	Clinic C	2	R	R	S	R	R	S	S	R	S
6	05/11/15	Clinic C	1	R	R	R	R	R	S	S	R	S
	05/11/15	Clinic C	2	S	R	S	R	R	S	S	R	S
7	15/11/15	Ward	1	R	R	R	R	R	R	S	S	S

P, patient; AMK, amikacin; CIP, ciprofloxacin; GEN, gentamicin; LVX, levofloxacin; MEM, meropenem; TOB, tobramycin; CST, colistin; CAZ, ceftazidime; TZP, piperacillin-tazobactam; R, resistant; I, intermediate; S, sensitive; OPD, outpatient department.

3.2.2.2.3 Reference laboratory typing

A summary table of reference laboratory VNTR results can be seen in Figure 3.16 (colours indicate different patients).

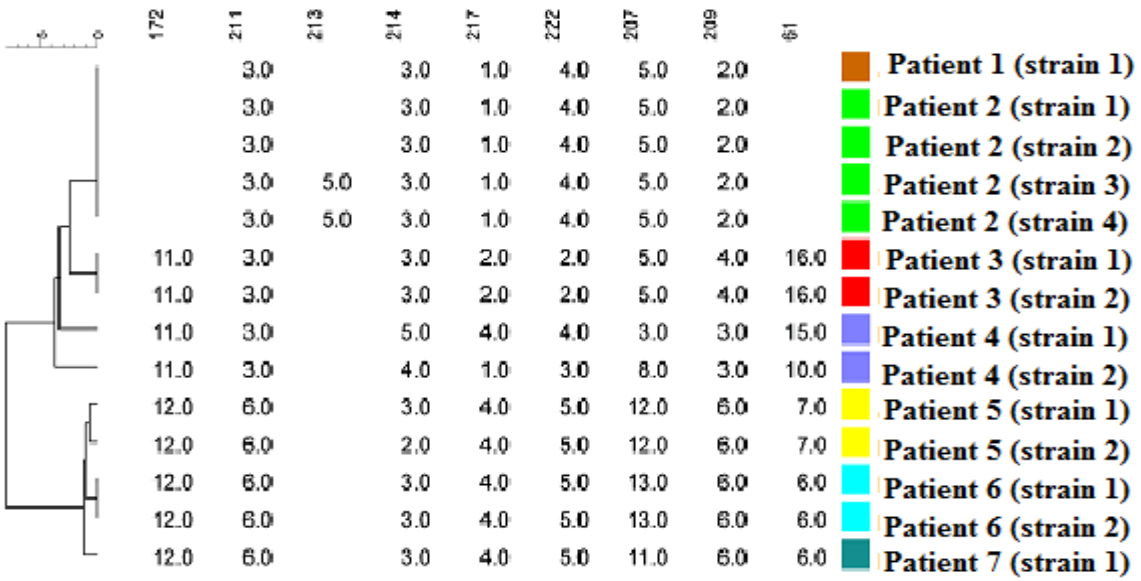


Figure 3.16. *P. aeruginosa* VNTR results produced at the reference laboratory. Patient 1 and 2 shared the same strain whilst patient 4, 5, 6 and 7 have isolates with similar profiles (with differences at either the 7th or 9th locus (or both) for each of these patients). Patient 4 has two, distinct strains compared to the other patients. Data source - National Reference Laboratory (Antimicrobial Resistance and Healthcare Associated Infections Reference Unit (AMRHAI), Public Health England, Colindale.

Patient 1 and 2 on VNTR appeared to share the same strain. VNTR showed patient 4, 5, 6 and 7 have isolates with similar profiles, however there were differences at either the 7th or 9th locus (or both) for each of these patients. Patient 4 has two, distinct strains

3.2.2.2.4 Whole genome sequencing results

Isolates from patient one and two were found to have the same VNTR result. On WGS both patient's isolates were identified to be ST1714 and share 53 SNP suggesting transmission or shared isolates, with possible common ancestor back in time, (CF patients are long-term colonized). Patient to patient transmission between Patients 3 and 4, in clinic was ruled out as both patients' isolates were unrelated by MLST (ST111 and ST2140). The isolates were not related to any other CF patient's isolates. Patient isolates from three patients: patient 6 (2 isolates, 5/11/2015), patient 5 (1 isolate, 6/11/2015) and patient 7 (1 isolate, 15/11/2015) had similar VNTR. WGS found all patients' isolates to be ST500 sharing 2194 SNPs therefore it was possible there was a common ancestor or shared pool of isolates (CF patients can be long-term colonised).

3.2.3 Extended-spectrum β-lactamase (ESBL) producing *Escherichia coli*

3.2.3.1 Community Hospital

3.2.3.1.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of ESBL producing *E. coli* in a community hospital for patients with dementia. An overview of the epidemiological and IPC investigation carried out for this suspected outbreak is shown in Table 3.30

Table 3.30. Overview of the epidemiological and IPC investigation carried out for suspected Community hospital ESBL producing *E. coli* outbreak

Healthcare facility/hospital	Community hospital
Microorganism	ESBL producing *E. coli*
Description of suspected outbreak	• Two patients identified to have urinary tract infection with *E. coli* ESBL. • The IPCT identified that some patients were incontinent and there had been spills on the floor. Screening was carried out which identified 9 further patients who were colonised with *E. coli* ESBL.
IPCT hypothesis	Transmission likely due to epidemiology, clinical setting, and patient group. Breakdown in hand hygiene measures and contact with a contaminated environment (some patients were incontinent and there had been spills on the floor).
Mode of spread	Hand contact and contaminated environment.
Case definition	Case definition – a patient with a positive *E. coli* ESBL result (ABER13ES-4).
Total number of suspected patients	Eleven
Outbreak specific control measures	• Contact precautions to prevent cross infection. • Hand Hygiene Audit – score 55% - Hand Hygiene PAG conducted All ward staff completed e-Learning Hand Hygiene and Glitter Box training. Hand hygiene audits will now be done weekly until score improved. Domestic Supervisor who will give domestic staff Hand Hygiene DVD training. • Patient placement – all patients are in side-rooms. • Environmental cleaning with chlorine based agent twice daily cleaning of hall way grab rails and corridors. • Equipment – allocated equipment present in patients rooms, cleaned daily

	• Environmental sampling carried out however results were negative.
HIIAT Score	Green
Outbreak confirmed/not confirmed	Confirmed outbreak
WGS influence on IPC management	Routine typing turnaround time was 12 days. WGS turnaround time was 11 days. Results were presented and discussed at outbreak meetings. Phylogenetic analysis was carried out to reveal the fine-scale relationship between outbreak isolates. Isolates were closely related and this information was used by the IPC team to inform their decision that a range of outbreak measures should be instigated to stop spread.
Impact of WGS	**Greater discrimination**- An outbreak was confirmed using PFGE and WGS. WGS provided additional granularity linking the isolates together in a cluster and showed that the 7 ST131 isolates were closely related differing by 31 SNPs.

3.2.3.1.2 Routine microbiology results

Antibiograms of the isolates are shown below in Table 3.31.

Table 3.31. Table of patients, symptoms, room number and *E. coli* ESBL positive sample type

P	Room	Urine	Date	Stool	Date	Symptomatic
1	5	*E. coli* ESBL	7/01/16	*E. coli* ESBL	14/01/16	Yes
2	2	*E .coli* ESBL	11/01/16	No growth	14/01/16	Yes
3	3	*E. coli* ESBL	12/02/16	*E. coli* ESBL	10/02/16	No
4	10	No growth	10/02/16	*E. coli* ESBL	10/02/16	No
5	4	*E. coli* ESBL	24/02/16	*E. coli* ESBL	11/02/16	No
6	10	*E. coli* ESBL	11/02/16	No growth	03/04/16	No
7	1	No growth	11/02/16	*E. coli* ESBL	14/02/16	No
8	6	No sample	-	*K. pneumoniae* ESBL	15/02/16	No
9	5	*E. coli* ESBL	31/08/16	*E. coli* ESBL	31/08/16	No
10	*1*	*E. coli ESBL*	07/09/16	No growth	24/08/16	No
11	9	No growth	16/06/16	*E. coli* ESBL	24/09/16	No

P, patient; ESBL, extended-spectrum β-lactamase.

The SPC chart used in the investigation of this outbreak is shown below in Figure 3.17.

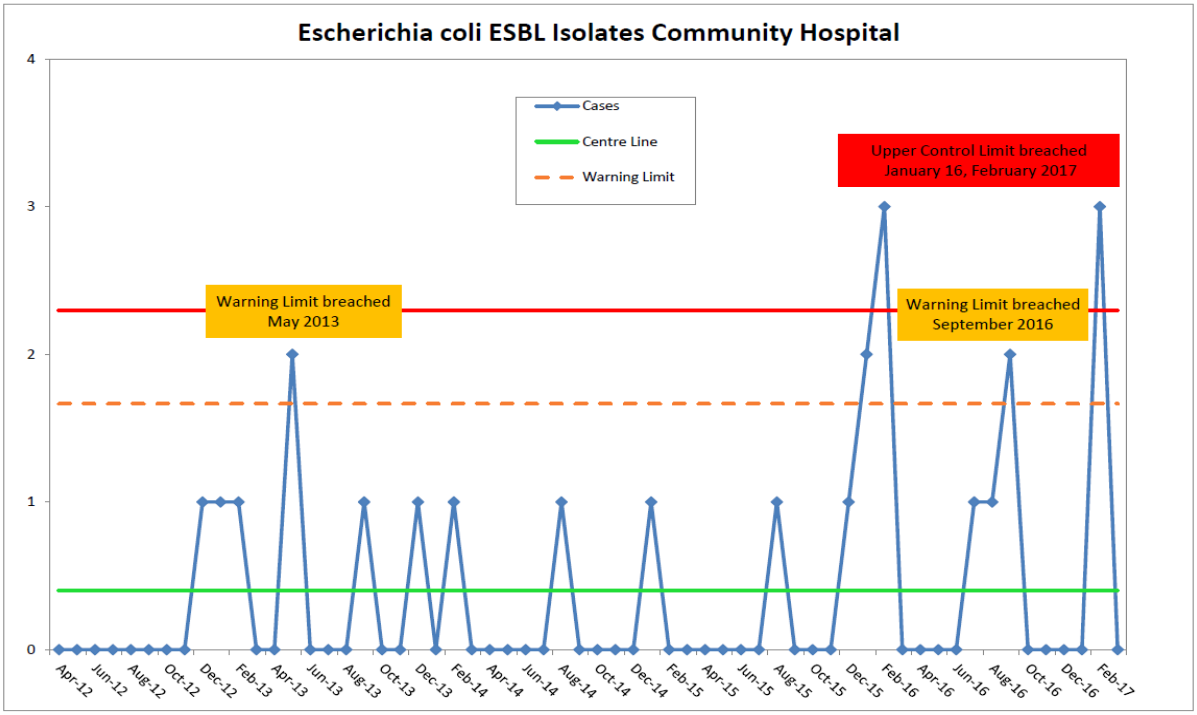


Figure 3.17. SPC chart of suspected ESBL producing *E. coli* from April 2012-February 2017. Produced by NHS Grampian IPC department using data from ICNet outbreak surveillance software.

Antibiograms of the isolates are shown below in Table 3.32.

Table 3.32. Antibiograms of ESBL producing *E. coli*

P	AMK	AMX	AMC	ATM	FEP	CIP	SXT	CAZ	ETP	GEN	MEM	TZP	TMO	TGC	TOB
1	S	R	S	-	R	S	R	R	S	S	S	S	S	S	S
2	-	R	S	-	S	R	S	R	S	S	S	S	S	S	S
3	S	R	S	R	R	S	R	R	S	S	S	S	S	S	S
4	S	R	S	I	R	S	R	R	S	S	S	S	S	S	S
5	S	R	S	R	R	S	R	R	S	S	S	S	S	S	S
6	S	R	S	R	R	S	R	R	S	S	S	S	S	S	S
7	S	R	S	R	R	S	R	R	S	S	S	S	S	S	S
8	S	R	S	R	S	S	S	R	S	R	S	R	S	R	S
9	S	R	R	R	R	S	R	R	S	S	S	S	S	S	S
10	S	R	S	R	R	S	R	R	S	S	S	S	S	S	S
11	S	R	S	R	R	S	R	R	S	S	S	S	S	S	S

P, patient; AMK, amikacin; AMX, amoxicillin; AMC, amoxicillin-clavulanic acid; ATM, aztreonam; FEP, cefepime; CIP, ciprofloxacin; SXT trimethoprim-sulfamethoxazole; CAZ, ceftazidime; ETP, ertapenem; GEN, gentamicin; MEM, meropenem; TZP, piperacillin-tazobactam; TMO, temocillin; TGC, tigecycline; TOB, tobramycin. R, resistant; I, intermediate; S, sensitive.

3.2.3.1.3 Reference laboratory typing

Table 3.33 displays the PFGE results produced by the reference laboratory.

Table 3.33. ESBL producing *E. coli* PFGE reference laboratory results

P	PFGE
1	ABER13ES-4
2	ABER13ES-4
3	ABER13ES-4
4	Unique
5	ABER13ES-4
6	ABER13ES-4
7	ABER13ES-4
8	Unique
9	ABER13ES
10	ABER13ES-4
11	Sample not kept

P, patient; PFGE, Pulse-field gel electrophoresis.

An example of the PFGE is shown in Figure 3.18.

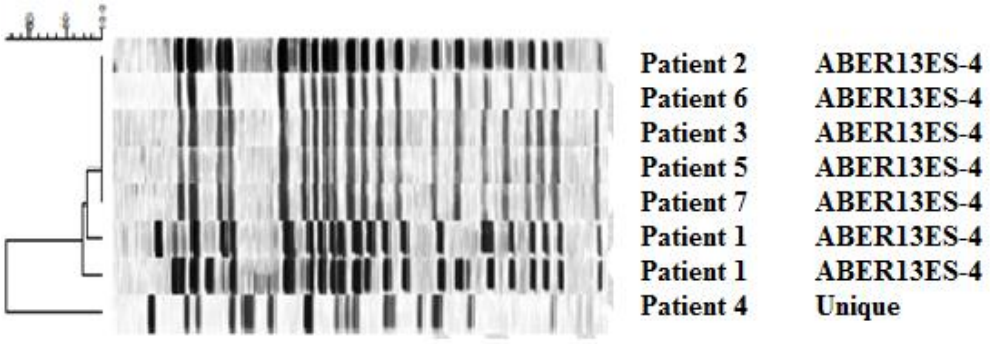


Figure 3.18. PFGE profiles of suspected *E. coli* ESBL outbreak isolates produced at the reference laboratory. Showing one patient's bacterial isolate was unique and the other patients isolates belonged to the same strain (ABER13ES-4). Data source - National Reference Laboratory (Antimicrobial Resistance and Healthcare Associated Infections Reference Unit (AMRHAI), Public Health England, Colindale

3.2.3.1.4 Whole genome sequencing results

For the *E. coli*, 7 of the isolates were from the same ST, ST 131: Patient 1, 2, 3, 5 and 6. Patient 4 were unrelated (belonging to ST131). The 7 ST131 isolates were mapped to the APOC_01 reference and all of the isolates were closely related differing by 31 SNPs linking the isolates together in a cluster.

3.2.3.2 Residential Care Home

3.2.3.2.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of ESBL producing *E. coli* in a residential care home as a GP had identified three patients who had urinary tract infection symptoms and had grown ESBL positive *E. coli* with similar antibiograms on urine culture. An overview of the epidemiological and IPC investigation is shown in Table 3.34.

Table 3.34. Overview of the epidemiological and IPC investigation carried out for suspected Residential care ESBL producing *E. coli* outbreak

Healthcare facility/hospital	Residential care home
Microorganism	ESBL producing *E. coli*
Description of suspected outbreak	Three patients identified to have *E. coli* ESBL from cultures with similar antibiotic sensitivity result from the same care home.
IPCT hypothesis	Transmission likely due to epidemiology, clinical setting, and patient group due to breach in hand hygiene.
Mode of spread	Hand contact and contaminated environment.
Case definition	Patients with positive with *E. coli* ESBL samples
Total number of suspected patients	3
Outbreak specific control measures	Not applicable
HIIAT Score	Green
Outbreak confirmed/not confirmed	Ruled out
WGS influence on IPC management	WGS results were available after 23 days and provided the IPC team with reassurance that an outbreak meeting was not required. Routine typing took 17 days.
Impact of WGS	**Greater discrimination** -WGS provided results with greater granularity than PFGE and suggested that isolates were not closely related enough to trace back to a transmission event.

3.2.3.2.2 Routine microbiology results

Antibiograms of the *E. coli* ESBL isolates are shown below (Table 3.35).

Table 3.35. Antibiograms of the ESBL producing *E. coli* isolates are shown below

P	Date	AMX	AMC	FEP	CIP	SXT	CAZ	ETP	GEN	MEM	TZP	TMO	FOF	DOX	NIT	FOX
1	05/08/15	R	R	R	R	R	R	S	R	S	S	S	S	R	S	I
2	05/08/15	R	S	R	R	R	R	S	S	S	S	S	S	R	S	R
3	05/08/15	R	R	R	R	R	R	S	R	S	S	S	S	R	R	R

P, patient; AMX, amoxicillin; AMC, amoxicillin-clavulanic acid; FEP, cefepime; CIP ciprofloxacin; SXT, trimethoprim- sulfamethoxazole; CAZ, ceftazidime; ETP, ertapenem; GEN, gentamicin; MEM meropenem; TZP, piperacillin-tazobactam; TMO, temocillin; FOF, fosfomycin; DOX, doxycycline; NIT, nitrofurantoin, FOX, cefoxitin. R, resistant; I, intermediate; S, sensitive.

3.2.3.2.3 Reference laboratory typing

Comparison by PFGE at the reference laboratory is shown in Figure 3.19. This revealed that these isolates represented distinct strains from one another. Although the isolates from patient 2 and patient 1 share some bands in common, the reference laboratory reported that there were in their opinion sufficient band differences for them to be considered distinct.

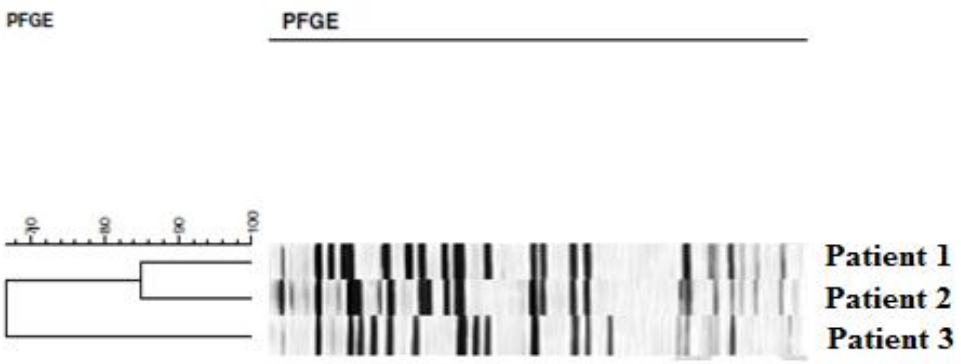


Figure 3.19. ESBL producing *E. coli* PFGE profiles produced at the reference laboratory. Showing isolates are distinct strains from one another. Data source - National Reference Laboratory (Antimicrobial Resistance and Healthcare Associated Infections Reference Unit (AMRHAI), Public Health England, Colindale

3.2.3.2.4 Whole genome sequencing results

Genomic analysis found that the three isolates belonged to ST131. Using a reference chromosome sequence from an isolate that originated in the UK patient 2 was found to be 99 SNPs from patient 1. Patient 2 was162 SNPs from patient 3, and patient 1 was 159 SNPs from patient 3. (214) In addition there was further evidence of recombination and variation in MGEs that distinguish the isolates. This suggested the isolates are related to a successful epidemic clone circulating widely in the Scotland and UK, and were not closely related enough to trace back to a transmission event.

3.2.3.3 Household transmission

3.2.3.3.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of ESBL producing *E. coli* between family members. An overview of the epidemiological and IPC investigation carried out for this suspected outbreak is shown in Table 3.36

Table 3.36. Overview of the epidemiological and IPC investigation carried out for suspected ESBL producing *E. coli* Household transmission outbreak

Healthcare facility/hospital	Household transmission
Microorganism	ESBL producing *E. coli*
Description of suspected outbreak	• A 74 year old man was admitted to hospital with urosepsis. He had a prior history of recurrent urinary tract infections (UTIs) with two admissions to hospital in the preceding 2 months. He was treated with intravenous 4.5g piperacillin/tazobactam via extended infusion improved clinically however had a myocardial infarction 3 weeks later and subsequently died. • Patient 2, a 61 year old lady who was the sister and carer of patient 1 was admitted to the same ward 35 days after patient 1 had been admitted with urosepsis. A midstream urine submitted by her GP and received by the laboratory 1 day prior to admission grew ESBL producing *E. coli* with the same antibiogram as patient 1. Blood cultures were negative. The patient was treated with intravenous meropenem 1g t.d.s which was continued for 8 days with clinical improvement. • No breaches in IPC measures had previously been identified on the ward. • The patient's sister informed staff that the patient had been incontinent at home which had contaminated furnishings.
IPCT hypothesis	Transmission had occurred in the community through a contaminated environment
Mode of spread	Hand contact and contaminated environment.
Case definition	Patient with ESBL producing *E. coli*
Total number of suspected patients	2
Outbreak specific control measures	Contact precautions, hand hygiene and isolation. Disposal of carpet and furniture
HIIAT Score	Green
Outbreak confirmed/not	Confirmed outbreak

confirmed	
WGS influence on IPC management	WGS results were not available in real-time and did not influence IPC management.
Impact of WGS	**Greater discrimination**- WGS provided additional layer of discrimination and showed both isolates were 3 SNPs apart supporting transmission between the two family members. This supported transmission in the community.

3.2.3.3.2 Routine microbiology results

Antibiograms of the isolates are shown below in Table 3.37

Table 3.37. Antibiograms of ESBL producing *E. coli* isolates

P	Date	AMX	CIP	SXT	CXM	CRO	LEX	DOX	GEN	AMC	ETP	FOF	MEM	NIT	TZP	TMO
1	03/01/16	R	R	R	R	R	R	R	R	S	S	S	S	S	S	S
2	21/01/16	R	R	R	R	R	R	R	R	S	S	S	S	S	S	S

P, patient; AMX, amoxicillin; CIP, ciprofloxacin; SXT, trimethoprim-sulfamethoxazole; CXM, cefuroxime; CRO, ceftriaxone; LEX, cephalexin; DOX, doxycycline; GEN, gentamicin; AMC, amoxicillin-clavulanic acid; ETP, ertapenem; FOF, fosfomycin; MEM, meropenem; NIT, nitrofurantoin; TZP, piperacillin-tazobactam; TMO, temocillin; R, resistant; I, intermediate; S, sensitive.

3.2.3.3.3 Reference laboratory typing

Isolates were sent for typing at the Antimicrobial Resistance and Healthcare Associated Infections Reference Unit (AMRHAI), Public Health England, Colindale. Comparison by pulsed –field gel electrophoresis suggested the isolate from patient 2 matched isolates from patient 1 both were named ABER13ES-3.

3.2.3.3.4 Whole genome sequencing results

Phylogenetic analysis using core genome SNPs was used to investigate the genetic relationships of strains. WGS showed isolates were ST131, both were highly similar being 3 SNPs apart supporting transmission between the two family members.

3.2.3.4 Maternity

3.2.3.4.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of ESBL producing *E. coli* in a tertiary hospital maternity unit. An overview of the epidemiological and IPC investigation carried out for this suspected outbreak is shown in Table 3.38.

Table 3.38. Overview of the epidemiological and IPC investigation carried out for suspected Maternity unit *E. coli* ESBL positive outbreak

Healthcare facility/hospital	Maternity unit
Microorganism	ESBL producing *E. coli*
Description of suspected outbreak	• A baby (patient 1) and their mother (patient 2) were found to be *E. coli* ESBL positive after admission to a hospital maternity unit. The baby was treated with antibiotics and died. • Another mother (patient 3) who had been attending the ward was identified to be positive for *E. coli* ESBL with a remarkably similar antibiogram.
IPCT hypothesis	Transmission of ESBL producing *E. coli* to patients from hand contact or contaminated equipment
Mode of spread	Hand contact contaminated equipment
Case definition	A patient positive for *E. coli* ESBL
Total number of suspected patients	3
Outbreak specific control measures	Not applicable
HIIAT Score	Green
Outbreak confirmed/not confirmed	Ruled out
WGS influence on IPC management	WGS results were available with an 8 day turnaround time. They directly influenced IPC management as an outbreak was ruled out and an outbreak meeting was not necessary. Routine typing took 10 days.
Impact of WGS	**Greater discrimination**- WSG revealed that patient 1 (Baby) and patient 2 (baby's mother) had the same ST type distinguished by a minimal genetic distance (1SNP in each patient), and patient 3 was a separate strain.

3.2.3.4.2 Routine microbiology results

Antibiograms of the ESBL producing *E. coli* isolates are shown below in Table 3.39.

Table 3.39. Antibiograms of *E. coli* ESBL positive isolates from the Maternity unit

P	Date	Sample	AMX	AMC	FEP	CIP	SXT	CAZ	ETP	GEN	MEM	TZP	TMO	ATM	CXM	AMK
1 (Baby)	20/04/15	Blood culture	R	S	R	R	R	I	S	S	S	S	S	I	R	S
2 (Mother of patient 1)	20/04/15	HVS	R	S	R	R	R	I	S	S	S	S	S	I	R	S
3 (another mother)	31/5/15	Blood culture and urine	R	S	R	R	R	R	S	S	S	S	S	R	R	S

P, patient; AMX, amoxicillin; AMC, amoxicillin-clavulanic acid; FEP, CIP, cefepime; CIP, ciprofloxacin; SXT, trimethoprim-sulfamethoxazole; CAZ, ceftazidime; ETP, ertapenem; GEN, gentamicin; MEM, meropenem; TZP, piperacillin-tazobactam; TMO, temocillin; ATM, aztreonam, CXM, cefuroxime, AMK, amikacin; R, resistant; I, intermediate; S, sensitive.

3.2.3.4.3 Reference laboratory typing

Comparison by PFGE (seen in Figure 3.20) showed that the isolates from patients 3 were distinct to other strains. The two isolates from patient 3 did appear to be different from one another. Isolates from patient 1 and 2 (family members mum and baby) were remarkably similar to one another, with many matching bands and they were given designations of ABER13ES-2 and ABER13ES-2' to reflect this.

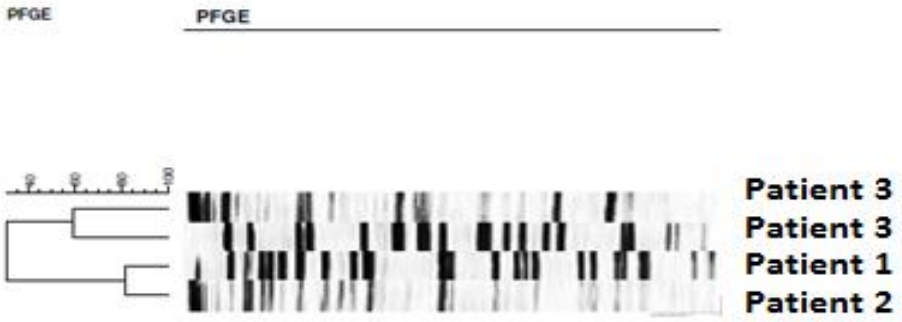


Figure 3.20. PFGE profiles of *E. coli* ESBL positive outbreak isolates produced at the reference laboratory. Isolates from patient 1 and 2 (family members mum and baby) were similar to one another. The isolates from patient 3 were distinct to other strains. Data source - National Reference Laboratory (Antimicrobial Resistance and Healthcare Associated Infections Reference Unit (AMRHAI), Public Health England, Colindale

3.2.3.4.4 Whole genome sequencing results

The isolates belonged to two sequence types: ST43 (patient 1 baby and patient 2 baby's mum) and ST1 (patient 3). The sequence data from each was mapped to a clonal complex reference (EPOA-01). The two pairs of ST43 isolates were distinguished by a minimal genetic distance (1SNP in each patient), and the two ST groups were distinguished by ~57,000 SNPs suggesting there were two separate strains.

3.2.4 Gram negative bacteria in Neonatal Intensive Care Units (NICUs)

3.2.4.1 *Pseudomonas aeurginosa* NICU

3.2.4.1.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of *P. aeruginosa* in an NICU of a tertiary hospital. An overview of the epidemiological and IPC investigation carried out for the suspected *P. aeurginosa* NICU outbreak is shown in Table 3.40.

Table 3.40. Overview of the epidemiological and IPC investigation carried out for suspected *P. aeruginosa* NICU outbreak

Healthcare facility/hospital	NICU
Microorganism	*P. aeruginosa*
Description of suspected outbreak	• *P. aeruginosa* was isolated from a conjunctival swab and stool sample of a neonate. A look back exercise was carried out which identified another patient to have had a umbilicus swab positive for *P. aeruginosa* 45 days prior to this (isolate not typed). An outbreak was suspected as both patients occurred in close conjunction and all isolates had the same antibiogram • Both babies had been treated with antibiotics and their clinical condition remained stable. • Several IMTs were held. • The Estates Department assessed the water supply including the fittings/ fixtures and sampled water from all sinks/water sources and taps. *P. aeruginosa* was identified in a cold pre-flush water sample and swab from drain of a stainless steel sink within the milk storeroom. This sink was used for cleaning breast pump equipment. In addition, three water samples from a visitor toilet tap were positive for *P. aeruginosa.* • Water sampling was repeated and no *P. aeruginosa* was identified from these areas. • A third patient was identified 120 days later. They were found to have a blood culture and throat swab positive for *P. aeurginosa*, had been discharged from the NICU and admitted to the Paediatric ward.
IPCT hypothesis	Transmission of *P. aeurginosa* between babies on the ward due contamination of water and following this contamination of hands and equipment used for patients.
Mode of spread	Hand contact, contamination of water/ fixtures.
Case definition	Any patient found to have a *P. aeurginosa* in a clinical sample.
Total number suspected patients	3
Outbreak specific control	• Resulting actions included screening and isolation of

measures	positive patients. The screening exercise was carried out on 23 babies and no further patients were identified. • The IPCT visited the ward and identified that frozen breast milk had been defrosted and warmed in warm tap water and surplus breast milk had been discarded down sinks. Ward staff were advised to discontinue both practices. • The milk storeroom sinks affected were put out-of-use, unused sinks were removed, an increased flushing schedule was enforced and documentation was encouraged. Tap water was replaced with sterilised bottled water for the care of neonates. When healthcare staff hands were visibly soiled they were advised to wash with soap and water, dry then use an alcohol based hand rub (ABHR).
HIIAT Score	Red
Outbreak confirmed/not confirmed	Ruled out
WGS influence on IPC management	WGS results were available before conventional reference laboratory results (10 days versus 12 days) and were fed back to the IPC team and clinicians. WGS results informed the IPC team that an outbreak could be ruled out and allowed the IPC team to open the ward earlier.
Impact of WGS	**Greater discrimination** - WGS and VNTR typing supported the hypothesis that there had not been transmission to or from patients 2 and 3 and that the *P. aeruginosa* contamination in the sink was specific for each sink and did not originate from one common water source.

An SPC chart used to investigate the outbreak can be seen in Figure 3.21. This demonstrates that the upper control limit was breached in September and October 2015.

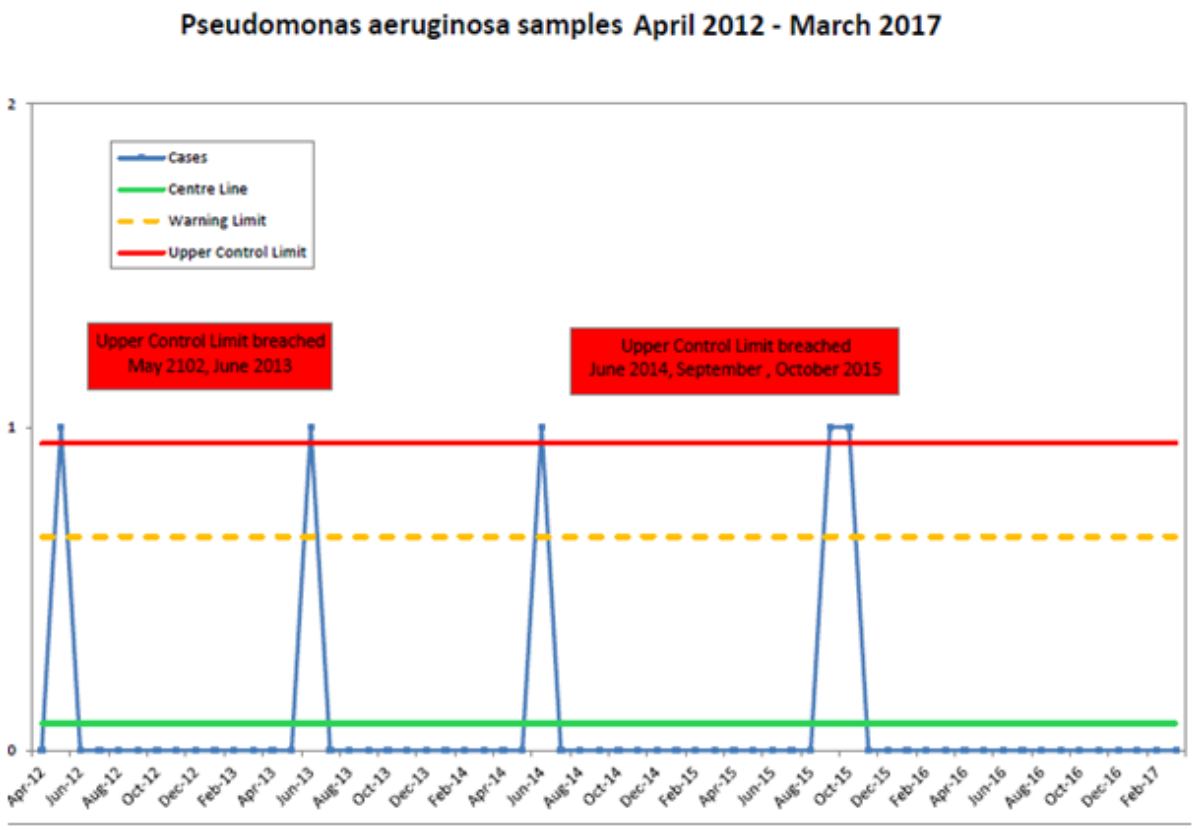


Figure 3.21. SPC chart of suspected NICU *P. aeruginosa* isolates from April 2012-February 2017. Produced by NHS Grampian IPC department using data from ICNet outbreak surveillance software.

3.2.4.1.2 Routine microbiology results

Antibiograms are shown in Table 3.41.

Table 3.41. Antibiograms of *P. aeruginosa* isolates from NICU

P	Date	Sample	ATM	AMK	CIP	CST	CAZ	GEN	IPM	MEM	TZP	TOB
1	12/09/15	Umbilical swab	I	S	S	S	S	S	S	S	S	S
2	26/10/15	Eye swab	I	S	S	S	S	S	S	S	S	S
2	02/11/15	Stool	-	-	-	-	-	-	-	-	-	
3	12/02/16	BC	I	S	S	S	S	S	S	S	S	S

P, patient; ATM, aztreonam; AMK, amikacin; CIP, ciprofloxacin; CST, colistin; CAZ, ceftazidime; GEN, gentamicin; IPM, imipenem; MEM, meropenem; TZP, piperacillin-tazobactam; TOB, tobramycin; R, resistant; I, intermediate; S, sensitive.

3.2.4.1.3 Reference laboratory typing

VNTR results produced by the reference laboratory are shown in Table 3.42.

Table 3.42. VNTR results for NICU *P. aeruginosa* isolates

Sample	VNTR
Patient 2 eye swab	11, 7, -, -, 2, 2, -, 3,13
Patient 2 stool sample	11, 7, -, -, 2, 2, -, 3,13
Patient 3 blood culture	10, 3, -, 5, 4, 1, 3, 7, 9
Milk store decontamination sink drain	8, 3, -, 5, 2, 3, 5, 2,12
NNU Milk store swab sink tap	11, 4, -, 2, 3, 2, -, 4, 9
NNU Visitor's toilet left basin cold pre flush	12, 2, -, 3, 2, 2, 8, 6, 9
NNU Visitor's toilet left basin mixed pre flush	12, 2, -, 3, 2, 2, 8, 6, 9
NNU Visitor's toilet left basin mixed post flush	12, 2, -, 3, 2, 2, 8, 6, 9

NNU, neonatal unit.

3.2.4.1.4 Whole genome sequencing results

The information in Table 3.43 demonstrates that WGS results were concordant with VNTR findings. WGS sequence type (ST) (ST1123, ST395, ST27, ST155, ST298) and VNTR profiles were all different from each other for each available isolate from patient and environmental sources. Both typing techniques supported the hypothesis that the *P. aeruginosa* contamination in the sink was specific for each sink and did not originate from one common water source and there had not been transmission to or from patients 2 and 3. We found that WGS results for the suspected NICU *P. aeruginosa* outbreak were available before conventional reference laboratory results (10 *versus* 12 days).

Table 3.43. NICU *P. aeruginosa* MLST results

Sample	ST
Patient 2 eye swab	1123
Patient 2 stool sample	1123
Patient 3 bacteraemia	395
Milk store decontamination sink drain	27
NNU Milk store swab sink tap	155
NNU Visitor's toilet left basin cold pre flush	298
NNU Visitor's toilet left basin mixed pre flush	298
NNU Visitor's toilet left basin mixed post flush	298

NNU, neonatal unit.

3.2.4.2 *Klebsiella pneumoniae* NICU

3.2.4.2.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of *K. pneumoniae* in a tertiary hospital NICU. An overview of the epidemiological and IPC investigation carried out for this suspected outbreak is shown in Table 3.44.

Table 3.44. Overview of the epidemiological and IPC investigation carried out for suspected *K. pneumoniae* NICU outbreak

Healthcare facility/hospital	NICU
Microorganism	*K. pneumoniae*
Description of suspected outbreak	• Towards the end of the previous *P. aeruginosa* outbreak described in section 3.5.1 the IPCT identified that two patients had been identified to have of *K. pneumoniae* infection within a 13 day period. Patient 1 was premature, had positive blood cultures and died after receiving antibiotics. • The second patient was premature, identified to have a positive endotracheal aspirate and blood cultures and their clinical condition improved following treatment with antibiotics. • Concern was raised that transmission had occurred as both patients had similar antibiograms apart from tigecycline • The environment was sampled and a swab from the milk store decontamination sink drain was found to have a profuse growth of *K. pneumoniae*.
IPCT hypothesis	Transmission of *K. pneumoniae* in the NICU between the environment and patients.
Mode of spread	Contamination of sink drains, hand contact and contamination of equipment
Case definition	Any patient identified to have a clinical sample which has with *K. pneumoniae*
Total number suspected patients	2
Outbreak specific control measures	• Reinforcement of hand hygiene practice. • Flushing and appropriate discarding of breast milk.
HIIAT Score	Red
Outbreak confirmed/not confirmed	Ruled out
WGS influence on IPC management	Routine typing took 7 days to deliver results. WGS took 50 days and results did not directly influence outbreak management but gave greater detail of the diversity amongst isolates.

Impact of WGS	**Greater discrimination** -Findings from VNTR and WGS were concordant and both supported ruling out an outbreak.

3.2.4.2.2 Routine microbiology results

Antibiograms of the *K. pneumoniae* isolates are shown below in table 3.45

Table 3.45. Antibiograms of the NICU *K. pneumoniae* isolates

P	Sample	Date	AMX	AMC	ATM	CIP	SXT	CRO	GEN	MEM	TZP	TGC
1	Blood cultures	22/11/15	R	S	S	S	S	S	S	S	S	I
1	Blood cultures	23/11/15	R	S	S	S	S	S	S	S	S	I
1	Blood cultures	27/11/15	R	S	S	S	S	S	S	S	S	I
2	Blood culture	11/12/15	R	S	S	S	S	S	S	S	S	S
2	ETA	10/12/15	R	S	S	S	S	S	S	S	S	S

P, patient; AMX, amoxicillin; AMC, amoxicillin-clavulanic acid; ATM, aztreonam; CIP, ciprofloxacin; SXT, trimethoprim-sulfamethoxazole; CRO, ceftriaxone; GEN, gentamicin; MEM, meropenem; TZP, piperacillin-tazobactam; TGC, tigecycline; R, resistant; I, intermediate; S, sensitive; I, intermediate; ETA, endotracheal aspirate.

3.2.4.2.3 Reference Laboratory Results

VNTR was carried out and all profiles were different as seen in Table 3.46.

Table 3.46. VNTR profiles of *K. pneumonia*e isolates

Sample	VNTR Profile
Patient 1 3 sets of blood cultures	4, 3, 6, 1, 2, 1, 4, 3, 1
Patient 2 Blood culture and endotracheal aspirate	10, 6, -, -, 2, 1, 3,26, 1
Sink Drain Milk Store	3, 5, -, -, 2, 2, 4, 3, 1

3.2.4.2.4 Whole Genome sequencing results

The comparison of routine typing with WGS can be seen in 3.47. VNTR profiles differed and WGS revealed isolates belonged to ST461, ST2251, and ST359. Findings from VNTR and WGS were concordant and both supported ruling out an outbreak.

Table 3.47. *K. pneumoniae* NICU isolates MLST results

Sample MLST	MLST
Patient 1 blood cultures	461
Patient 2 Blood culture and endotracheal aspirate	2251
Sink Drain Milk Store	359

3.2.4.3 *Klebsiella pneumoniae* NICU

3.2.4.3.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of *K. pneumoniae* in a tertiary hospital NICU. An overview of the epidemiological and IPC investigation carried out for this suspected outbreak is shown in Table 3.48.

Table 3.48. Overview of the epidemiological and IPC investigation carried out for suspected NICU *K. pneumoniae* outbreak

Healthcare facility/hospital	NICU
Microorganism	*K. pneumoniae*
Description of suspected outbreak	• Patient one was initially treated with antibiotics for a *K. pneumoniae* bacteraemia. • Fifty four days later a further neonate (patient 2) was identified to be colonised with *K. pneumoniae* with positive swab of the groin area. • Four days after this, patient 3 was identified to have a urine that isolated *K. pneumoniae*. This patient was treated with gentamicin and amoxicillin for 24 hours. • A total 14 babies underwent screening of faeces, breaks in skin, urine and nasal/throat swabs. This revealed two further patients. Patient 4 was found to have had positive faeces for *K. pneumoniae* and patient 5 had positive faeces and nasal swab.
IPCT hypothesis	Transmission had occurred as two patients were identified to have *K. pneumoniae* within a four day period and antibiograms of isolates from patients two and three were highly similar. Transmission could be due to breaches in hand hygiene practice.
Mode of spread	Hand contact
Case definition	Any patient identified to have a clinical sample which has *Klebsiella pneumoniae.*
Total number suspected patients	5
Outbreak specific control measures	• IPCT visited the unit, hand hygiene training and audits were carried out. Audits later found that hand hygiene was good and equipment was clean. • Contact precautions and cleaning with chlorine based agents was instigated
HIIAT Score	Red
Outbreak confirmed/not confirmed	Confirmed
WGS influence on IPC management	Routine Typing took 8 days. WGS results were available in 11 days. WGS results did not directly influence outbreak

	management
Impact of WGS	**Greater discrimination**- WGS supported VNTR confirming transmission between patient two and three (isolates ST252).

3.2.4.3.2 Routine microbiology results

Antibiograms of the *K. pneumoniae* isolates are shown in Table 3.49.

Table 3.49. Antibiograms of the NICU *K. pneumoniae* isolates

P	Date	Sample	ATM	CIP	CAZ	GEN	MEM	TZP	TOB	AMX	AMC	TGC	SXT
1	13/08/17	Blood culture and line tip	S	S	S	R	S	R	R	R	R	S	R
2	05/10/17	Wound swab groin	S	S	S	S	S	I	S	R	S	R	S
3	09/10/17	Urine	S	S	S	S	S	I	-	R	S	-	S
4	27/10/17	Faeces	S	S	S	S	S	S	S	R	S	S	S
5	04/11/17	Faeces and nasal swab	S	S	S	S	S	S	S	R	S	S	S

P, patient; ATM, aztreonam; AMK, amikacin; CIP, ciprofloxacin; CAZ, ceftazidime; GEN, gentamicin; MEM, meropenem; TZP, piperacillin-tazobactam; TOB, tobramycin; AMX, amoxicillin; AMC, amoxicillin-clavulanic acid; TGC, tigecycline; SXT, trimethoprim-sulfamethoxazole; ; R, resistant; I, intermediate; S, sensitive.

Timeline of bed movements of *K. pneumoniae* positive patients can be seen in Figure 3.22 below.

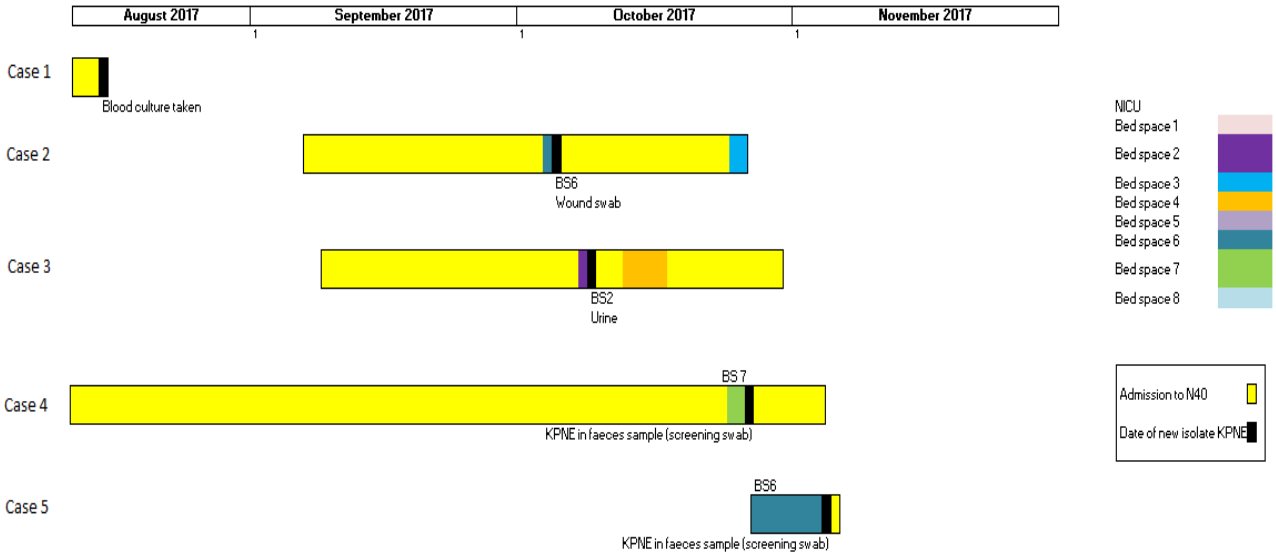


Figure 3.22. Timeline of bed movements of *K. pneumoniae* positive patients during the NICU outbreak. Produced by NHS Tayside IPC department using data from ICNet outbreak surveillance software.

3.2.4.3.3 Reference laboratory typing

Isolates underwent routine typing by VNTR (Table 3.50). VNTR profiles were found to be identical for isolates from patients 2 and 3 (5, -, 8, 3, 2, 2, 4, 5, 3, 2, 4) supporting that transmission had taken place.

Table 3.50. VNTR results of NICU *K. pneumoniae* isolates

P	Sample	VNTR
1	Blood culture and line tip	5,4,6,IS,2,2,4,3,5,3,8
2	Wound swab groin	5,-, 8,3,2,2,4,5,3,2,4
3	Urine	5,-,8,3,2,2,4,5,3,2,4,
4	Faeces	3,5,1,1,1,1,1,2,4,2,3
5	Faeces and nasal swab	NT

P, patient; NT, not tested.

3.2.4.3.4 Whole genome sequencing results

WGS results can be seen below in Table 3.51, patient two and three had isolates that belonged to ST252.

Table 3.51. MLST results for NICU *K. pneumoniae* isolates

P	Sample	ST
1	Blood culture and line tip	ST37
2	Wound swab groin	ST252
3	Urine	ST252
4	Faeces	Unknown ST
5	Faeces and nasal swab	NT

P, patient; NT, not tested.

3.2.4.4 *Klebsiella oxytoca*

3.2.4.4.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of *Klebsiella oxytoca* NICU in a tertiary hospital. An overview of the epidemiological and IPC investigation carried out for this suspected outbreak is shown in Table 3.52.

Table 3.52. Overview of the epidemiological and IPC investigation carried out for suspected NICU *K.oxytoca* outbreak

Healthcare facility/hospital	NICU
Microorganism	*K. oxytoca*
Description of suspected outbreak	• Five patients in NICU were found to isolate *K. oxytoca* over a three month period. The antibiogram was the same for all samples in this first cluster and there was concern that this could be an outbreak and IMTs were held. • PFGE was carried out as part of routine typing and DNA was found to degrade for all isolates. • It was noted that the unit had been extremely busy however staffing levels had been maintained to ensure 1:1 nursing care where this was necessary and the unit had been compliant with the British Association of Perinatal Medicine (BAPM) guidance for staffing in neonatal units.(215) Where issues had arisen and this had not been achievable, the unit has closed to admissions. • The existing ward environment did not meet the recommended space between beds but work was being undertaken to reduce items in the clinical area to a minimum. After a period of 69 days in which no further new isolates were identified a sixth baby isolated *K. oxytoca* from nasopharyngeal aspirate. The DNA for this isolate was also found to degrade on testing at the reference laboratory, therefore typing was not possible. • A seventh neonate was found a further 32 days later.
IPCT hypothesis	• Transmission of *K. oxytoca* on the NICU due to non-compliant hand hygiene practices.
Mode of spread	Contact
Case definition	Any patient identified to have a clinical sample which has *K. oxytoca*
Total number suspected patients	7
Outbreak specific control	• Hand hygiene training and audits were carried out.

measures	• Environmental swabs were taken from various sites in the unit. These swabs grew various skin organisms but not *K. oxytoca.* • Enhanced cleaning using chlorine based agents was carried out by domestic and nursing staff throughout the unit.
HIIAT Score	Yellow
Outbreak confirmed/not confirmed	Outbreak confirmed.
WGS influence on IPC management	Reference laboratory results took 20 days. WGS results were available in 11 days and gave the IPC team assurance that this was an outbreak and IPC measures should be kept in place as they were appropriate.
Impact of WGS	**Greater discrimination**- PFGE found all isolates to have DNA that degraded apart from the seventh patient which was confirmed by PFGE to be unique. WGS successfully determined that all isolates (apart from patient seven) were *K. oxytoca* ST178 all closely related within 2 SNPs difference. In addition, WGS identified patient 7 isolate to be *Klebsiella michiganensis.*

3.2.4.4.2 Routine microbiology results

The antibiogram of all isolates of *K. oxytoca* showed no variation in sensitivities (Table 3.53).

Table 3.53. Antibiograms of the NICU *K. oxytoca* isolates

P	Sample	Date	ATM	AMK	CIP	CST	CAZ	GEN	MEM	TZP	TOB	AMX	AMC	TGC	SXT
1	NPA	13/04/18	S	S	S	S	S	S	S	S	S	R	S	S	S
2	WS and NPA	23/05/18	S	S	S	S	S	S	S	S	S	R	S	S	S
3	ETA	30/05/18	S	S	S	S	S	S	S	S	S	R	S	S	S
4	ETA x 2	11/06/18	S	S	S	S	S	S	S	S	S	R	S	S	S
5	WS	29/06/18	S	S	S	S	S	S	S	S	S	R	S	S	S
6	NPA	06/09/18	S	S	S	S	S	S	S	S	S	R	S	S	S
7	NPA	08/10/18	S	S	S	S	S	S	S	S	S	R	S	S	S

P, patient; ATM, aztreonam; AMK, amikacin; CIP, ciprofloxacin; CST, colistin; CAZ, ceftazidime; GEN, gentamicin; MEM, meropenem; TZP, piperacillin-tazobactam; TOB, tobramycin; AMX, amoxicillin; AMC, amoxicillin-clavulanic acid; TGC, tigecycline; SXT, trimethoprim-sulfamethoxazole; R, resistant; I, intermediate; S, sensitive; NPA, nasopharyngeal aspirate; WS, wound swab; ETA, endotracheal aspirate.

3.2.4.4.3 Reference laboratory typing

Routine typing by PFGE found all isolates to have DNA that degraded apart from the seventh patient which was confirmed by PFGE to be unique to all isolates that had previously been sent from the hospital to the reference laboratory (Table 3.54)

Table 3.54. PFGE results of the NICU *K. oxytoca* isolates

P	Sample	PFGE result
1	NPA	DNA degraded
2	WS	DNA degraded
3	ETA	DNA degraded
4	ETA	DNA degraded
5	WS	DNA degraded
6	NPA	DNA degraded
7	NPA	Unique

P, patient; NPA, nasopharyngeal aspirate; WS, wound swab; ETA, endotracheal aspirate.

3.2.4.4.4 Whole genome sequencing results

WGS successfully determined that all isolates (apart from patient seven) were *K. oxytoca* ST178. The isolates that were ST178 were all closely related within 2 SNPs difference. WGS identified patient 7 isolate to be *Klebsiella michiganensis.*

3.2.4.5 *Enterobacter asburiae* NICU

3.2.4.5.1 Description of suspected outbreak

The IPCT were alerted to a suspected outbreak of *Enterobacter asburiae* in a tertiary hospital NICU. An overview of the epidemiological and IPC investigation carried out for this suspected outbreak is shown in Table 3.55

Table 3.55 Overview of the epidemiological and IPC investigation carried out for suspected NICU *E. asburiae* outbreak

Healthcare facility/hospital	NICU
Microorganism	*E.asburiae*
Description of suspected outbreak	• *E. asburiae* was isolated from two NICU patients within 7 days of each other. Patient 1 was an extremely premature patient and *E. asburiae* was isolated from three ETAs however they showed no signs of infection. • Patient two had a bacteraemia due to *E. asburiae* and was treated with antibiotics (vancomycin, gentamicin, metronidazole, cefotaxime). • Antibiograms were similar apart from tobramycin sensitivities • Both babies were nursed in incubators in the intensive care, next to each other.
IPCT hypothesis	Transmission of *E. asburiae* between patients in NICU may have occurred.
Mode of spread	Hand contact
Case definition	Any patient identified to have a clinical sample which has *E. asburiae*
Total number suspected patients	2
Outbreak specific control measures	• The IPCT visited the ward and advised that contact precautions should be used. • Chlorine based cleaning was advised. • Hand hygiene audits, equipment checks and observation of practice were carried out and no issues of concern were identified.
HIIAT Score	Green
Outbreak confirmed/not confirmed	Ruled out
WGS influence on IPC management	Routine typing took 16 days. WGS took 24 days. WGS results did not directly influence IPC management when investigating the suspected outbreak.
Impact of WGS	**Greater discrimination** -WGS provided greater granularity to PFGE and showed that the two isolates differed in excess of 50,000 SNPs.

3.2.4.5.2 Routine microbiology results

Antibiograms were similar apart from tobramycin sensitivities (Table 3.56)

Table 3.56 Antibiograms of NICU *E. asburiae* isolates

P	Sample	Date	ATM	AMK	CIP	CST	CAZ	GEN	IPM	MEM	TZP	TOB	AMX	AMC	TGC	SXT
1	ETA	07/02/19	S	S	S	S	S	S	-	S	S	I	R	R	S	S
2	BC	14/02/19	S	S	S	S	S	S	S	S	S	S	R	R	S	S

ATM, aztreonam; AMK, amikacin; CIP, ciprofloxacin; CST, colistin; CAZ, ceftazidime; GEN, gentamicin; IPM, imipenem; MEM, meropenem; TZP, piperacillin-tazobactam; TOB, tobramycin; AMX, amoxicillin; AMC, amoxicillin-clavulanic acid; TGC, tigecycline; SXT, trimethoprim-sulfamethoxazole; R, resistant; I, intermediate; S, sensitive; ETA, endotracheal aspirate; BC, blood culture.

3.2.4.5.3 Reference laboratory typing

PFGE was used for routine bacterial typing and found both isolates to be unique.

3.2.4.5.4 Whole genome sequencing results

WGS supported these findings distinguishing the two isolates by finding that they differed in excess of 50,000 SNPs.

3.3 Bacterial isolates sequenced and accession numbers

A supplementary table of samples sequenced and their accession numbers available at the time of writing can be seen in Appendix 5. Further sequenced isolates will be submitted to the European Nucleotide Archive as collections are collated and they will be able to be accessed from this.

4 Discussion

4.1 Overview of study

The first and main objective of this work was to establish a WGS service for the investigation of suspected HAI outbreaks to test the hypothesis that this is achievable in the UK.

- Since 2014 I have successfully applied WGS to 21 different outbreaks in NHS Tayside and NHS Grampian. As part of this, I have used SPC charts to identify outbreaks and over 400 isolates were sent for typing from a wide range of outbreak investigations. These occurred in a varied spectrum of specialities in both community and secondary care hospital settings.
- These outbreaks have involved a wide range of Gram positive and Gram negative organisms including a variety of resistance mechanisms for example MRSA, ESBL, *optrA* gene positive enterococci, CPE and VREfm.
- Community healthcare outbreaks included rehabilitation hospitals, clinics, midwife units and care homes whilst secondary care outbreaks included wards such as renal, maternity, orthopaedics, and high risk clinical areas such as NICU and ICU.
- Overall, I have demonstrated that a WGS service for the investigation of suspected HAI outbreaks can be established in the UK and that this is achievable across a range of NHS settings and outbreak scenarios.

Objective 2 involved identifying the challenges of implementing this new technology. The challenges that I identified fall into six different groups and I will discuss these in more detail later in this chapter:

1. Infrastructure
2. Performance and quality assessment of data and processing
3. Pipelines and management of reference databases
4. When should WGS be used?
5. Clinical interpretation of results and meaning
6. When to use increased WGS discriminatory power in outbreak investigations

My third objective was to identify any clinical benefits that can occur from using WGS for suspected outbreak management. I identified that there was a wide range of instances in which the use of WGS was more beneficial to staff investigating outbreaks than standard typing techniques, supporting my second hypothesis that utilising WGS for the management of outbreaks will result in clinical benefits for staff and patients. These clinical benefits will be discussed in detail in this chapter and will be considered under the following headings:

1. WGS can provide results with greater granularity than routine typing methods
2. Genomic analysis can enhance the detection of 'alert organisms'
3. WGS could replace the need for multiple tests
4. Genomic analysis can be used to rule out outbreaks
5. WGS can be utilised to investigate new resistance mechanisms

After discussing the clinical benefits of WGS I will then present robust evidence that supports the utility of WGS in HAI investigations. I have synthesised my findings to make recommendations of how best to establish a WGS service. I also present a clinical decision aid that I developed to assist users on how to utilise WGS in HAI outbreaks investigations in real-time.

4.2 Challenges of establishing a WGS service

Recommendations for the solutions to the various challenges that must be resolved when translating the promise of WGS into clinical practice can be seen in Figure 4.1. Each of these will be discussed in detail in this chapter.

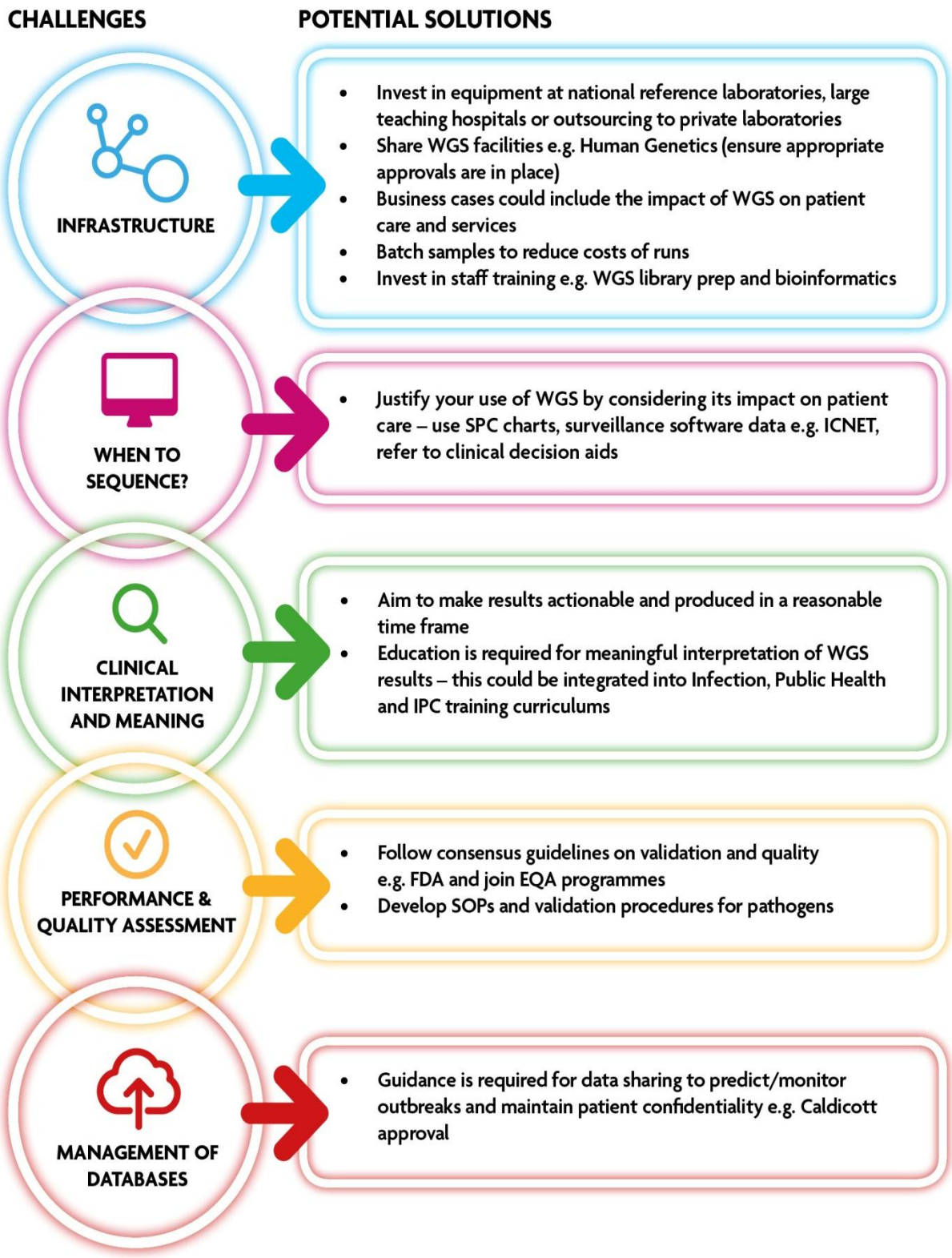


Figure 4.1. Challenges of sequencing in a clinical environment and potential solutions. These included infrastructure, when to sequence, clinical interpretation and meaning, performance/quality assessment and management of databases.

4.2.1 Infrastructure

The first barrier to integration of WGS into conventional outbreak analysis in any clinical microbiology laboratory is infrastructure. Options include placing and investing in equipment at national reference laboratories, at hub sites such as large teaching hospitals or outsourcing to private laboratories. Developments in NGS has enabled clinical microbiology laboratories to come a step closer to performing low cost WGS themselves using simple benchtop technology and user friendly library preparation protocols. Workflows have been created in some institutions combining microbiology, bioinformatic and epidemiological methods which are used to produce data which is then analysed and fed to public health authorities for action.(216) Thought must be given to producing an actionable result within a useful timeframe and I found clinical impact was greatest when turn-around times (TATs) of WGS were similar to that of routine typing methods therefore I advise these should at least be aimed for. Benefits of placing facilities closer to patients include streamlining of work (replacement of multiple tests and reduction in sending samples away for confirmatory testing) which could result in reduced TATs. However, if WGS is incorporated into local teaching hospital facilities there would need to be investment in appropriately trained staff, equipment, and data analysis. To increase the success of a business case the potential impact of WGS on patient management should be included. This could include impact on antimicrobial stewardship, IPC measures, outbreak investigation and subsequent impact on services. To further strengthen a business case WGS facilities could be shared with other departments such as Human Genetics or a University department as long as appropriate approvals are in place. Conversely centralising WGS services to reference laboratories could improve cost effectiveness as isolates could be pooled resulting in cheaper individual tests however travel time and costs for this would need to be considered. Consideration also should to be given to the setting up of pipelines, cluster analysis and investment in staff training e.g. WGS library prep, bioinformatics and using programmes such as Pathogenwatch.(217)

4.2.2 Performance and quality assessment of data and processing

Regardless of location laboratories need to assure the quality of NGS and validate their methods. It is recommended that laboratories undergo UKAS accreditation and staff should follow consensus guidelines on validation and quality e.g. FDA consideration (number 4). They will also be required to develop internal standard operating and validation procedures for all types of pathogens and join EQA programmes for wet prep and digital validation sets.(218)(219) Figure 4.2 gives examples of performance and quality assessments during the process of WGS for clinical microbiology users.

Figure 4.2. Recommendations for performance and quality assessment of data and processing in WGS service for outbreak detection. Quality assessments of data are displayed in orange coloured boxes.

4.2.3 Pipelines and management of reference databases

The analysis of high-throughput sequencing (HTS) data can be challenging and traditionally bioinformatic support has always been required. A suitable pipeline should be selected based on what is required from WGS. This depends on whether the aim is to focus on pathogen identification, characterisation, virulence factors, antibiotic resistance or whether metagenomic approaches are being untaken. De Almeida *et al.* define pipelines as: "*a set of scripts, which can be written in various programming languages and its function is to contain tasks that modify the input file"*.(220). A number of pipelines have recently emerged for instance SUPRI (sequenced based ultra-rapid pathogen identification), pathospere, and galaxy software can be used for pathogen identification.(220) Microbial identification and characterization through read analysis (MICRA) is an automatic pipeline with a web interface for fast characterisation of microbial genomes through read analysis.(221) By mapping against reference genomes certain genes can be identified and there is an option for prediction of antibiotic susceptibility and resistance number. TORMES is an automated pipeline for whole bacterial genome analysis designed for analysis of bacteria from HTS and Illumina platforms have user friendly non-bioinformatician automated WGS analysis steps.(222) PathoScope 2.0 can be used for diagnostic metagenomic approaches to identify bacteria in clinical samples.(223) There are other useful resources available such as the Virulence factor database first published in 2005 and the BLAST search tool which can be used to record and find VF-related genes.(224)(220) For antibiotic resistance a number of tools are available including ResFinder, ARG-ANNOT (Antibiotic Resistance Gene-ANNOTation) and CARD (Comprehensive Antibiotic Resistance Database).(225)(220)

At present there is no official centralised cloud or virtual computing system for clinical laboratories to carry out pathogen genomic analysis. Local clusters need to be sustainable and have capacity for demand, Potentially this could be arranged with data centres such as the ENA or Genomics England data centre.(186) In addition, guidance for sharing genomic data needs to be developed so that privacy is maintained and genomic sequences can be uploaded onto databases that can be shared and used as early warning system for outbreaks. The FDA-ARGOS (FDA dAtabase for Regulatory-grade microbial Sequences) is a public microbial reference database developed by the FDA (Food and Drug Administration**)**, Department of Defence (DoD), University of Maryland and the National Center for Biotechnology Information (NCBI) can also be used for *in-silico* performance validation.(226)

4.2.4 When should WGS be used?

IPCTs can have pressing questions that needed to be answered when faced with a suspected outbreak; does the IPCT need to convene a meeting with staff in the hospital; is there likely to be a common source; should the environment be sampled; which specific IPC interventions need to be implemented; should patients, staff or relatives be screened; should the unit be closed? Firstly, they need to confidently assess whether it is likely that transmission has occurred. This can be challenging as the prevalence of causative pathogens and clusters can vary with patient population and over time.(65)(227) IPCTs are often on the back foot as outbreak investigation may involve retrospective investigation of transmission events.(228) An early response is essential to identify transmission pathways and targeted measures to prevent further transmission. In order to do this, we developed a process of outbreak identification that was automated with set "action line" rules using epidemiological data, SPC charts and ICNet outbreak surveillance software to inform when an outbreak is likely and isolates should undergo WGS. Considering our findings, I have suggested various approaches to using WGS for outbreak detection (Figure 4.3).

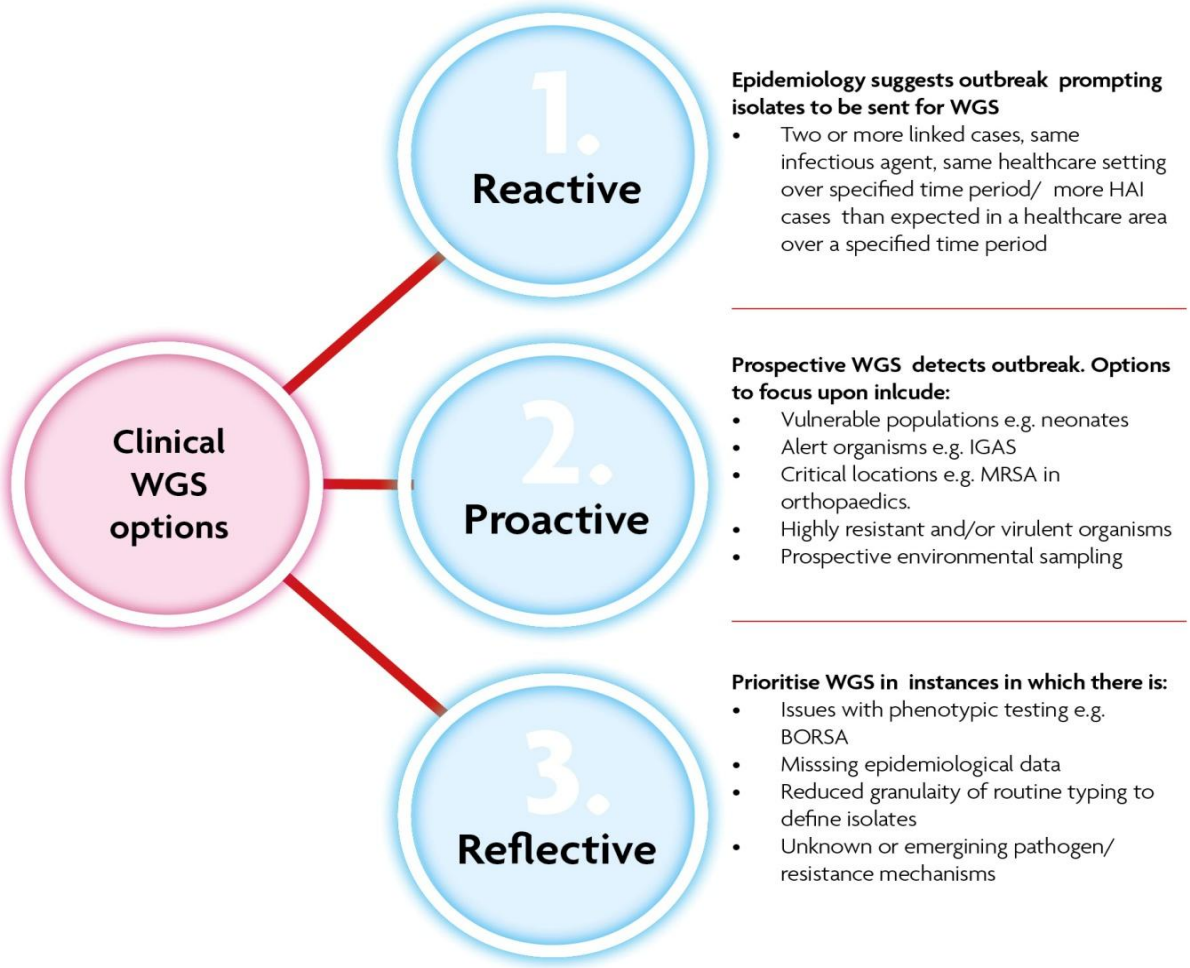


Figure 4.3. Recommendations for utilising WGS for HAI outbreak detection. Potential options may be reactive, proactive or reflective.

The first approach involves a reactive automated response to epidemiology that suggests there is an outbreak. In this instance phenotypic methods may suggest there is an outbreak and isolates could then be sent for WGS. An alternative to this is a proactive approach in which WGS is used to detect an outbreak regardless of epidemiology. In these instances users of WGS could prospectively sequencing select populations of patients who may be vulnerable to infection who may be from a critical location such as NICU or focus upon sequencing an alert, highly resistant or virulent organism from patient samples or the environment. Sequencing of a specific group of microorganisms that cause HAI in a complete geographic area can give a high-resolution view of the pathogen population that can pinpoint the genetic basis of resistance and spread of the pathogen. This would represent a shift in identifying outbreaks. A reflective approach to outbreak investigations could include using WGS in defined instances when there is missing epidemiology data or in scenarios in which phenotypic testing lacks granularity.

4.2.5 Clinical interpretation of results and meaning

Time and resources also need to be allocated to develop staff skills in carrying out WGS and also for the interpretation of results. It would be essential that training would become integrated in curriculums for microbiology, infection, public health, and IPC training. There are also residential courses such as Genomics and Clinical Microbiology and online virtual courses such as Bacterial Genomes: Disease Outbreaks and Antimicrobial Resistance that can be accessed. (229)(230) User friendly software such as Microreact is available can be used to detect outbreaks allowing NHS staff to upload, visualise and explore microbiology data such as antibiograms and dendrograms linked to geographic locations e.g. hospital wards. It has been developed by the Centre for Genomic Pathogen Surveillance at Imperial College London and the Wellcome Genome Campus. This software has successfully been implemented to detect and manage a number of outbreaks: tracing the spread of carbapenem resistance *K. pneumoniae* across Europe; West African Ebola epidemic 2013-2016; Zika virus outbreak; *Vibrio cholerea* outbreak; high level azithromycin resistant *N. gonorrhoeae* outbreak in the UK. (230)(63) Results should be actionable, understandable and produced in a reasonable time frame. It is important to share the necessary results and clinical interpretation of WGS with members of the outbreak team. During this MD results were sometimes shared during outbreak meetings or separately with members of the IPC team when relevant. In the future it would be beneficial to feedback results as part of a combined IPC team/genomic meeting. Phylogenetic trees were plotted for the investigation of certain outbreaks investigated in this MD. They were created when it was beneficial to display the relatedness of bacterial isolates in detail. This was not required for all outbreaks particularly when there was considerable epidemiological data available. We did consider plotting phylogenetic trees with annotated antibiotic resistance and genes however this was not required for the investigation of the outbreaks and we would have had limited time to carry this out for each outbreak. Figure 4.4 shows a WGS example report.

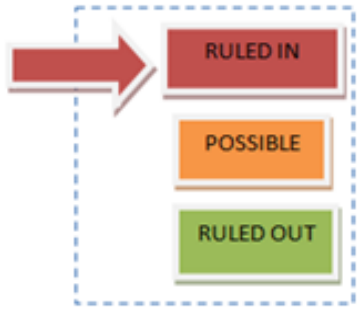
University of St Andrews

Whole Genome Sequencing report

Date isolates submitted: 20/11/17

Date reported: 29/11/17

Suspected outbreak rating score



Five MRSA samples were sequenced one from each of case 2,3,4,5,6. All of these isolates were EMRSA15. According to MLST these samples all belonged to CC22, four of the samples to ST22 and one was a single locus variant with a novel allele. The sequence data was mapped to an EMRSA-15 reference. Three of the samples (patients 3, 4, 5) all clustered very closely together and differed by 9 SNPs, with patient 5 and patient 3 being closest differing by 3 SNPs. The next nearest genetically related isolate was patient 2 which is 87 SNPs away from this cluster. Genome data alone suggestS that patient 5, patient 3, patient 4 are closely enough genetically to be part of an outbreak.

Phylogenetic tree

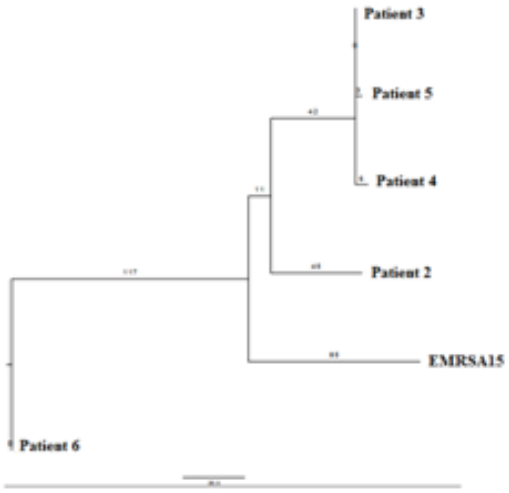


Figure 4.4. An example of a WGS report. Created by Benjamin Parcell to be communicated to clinical microbiologists and IPC teams. The report gives a score on the likelihood of an outbreak and an explanation of the phylogeny and its clinical implications.

4.2.6 When to use increased WGS discriminatory power in outbreaks

As part of objective 2 – identifying the practical challenges for implementing a WGS service for suspected outbreak detection I have also aimed to determine the instances in which increased WGS discriminatory power and phylogenetic analysis is required. There continues to be significant cost to sequencing and using WGS for the investigation of all outbreaks is not likely to be feasible for the foreseeable future. Figure 4.5 below summarises my experience of when WGS discriminatory power was particularly required in the outbreak situations I encountered.

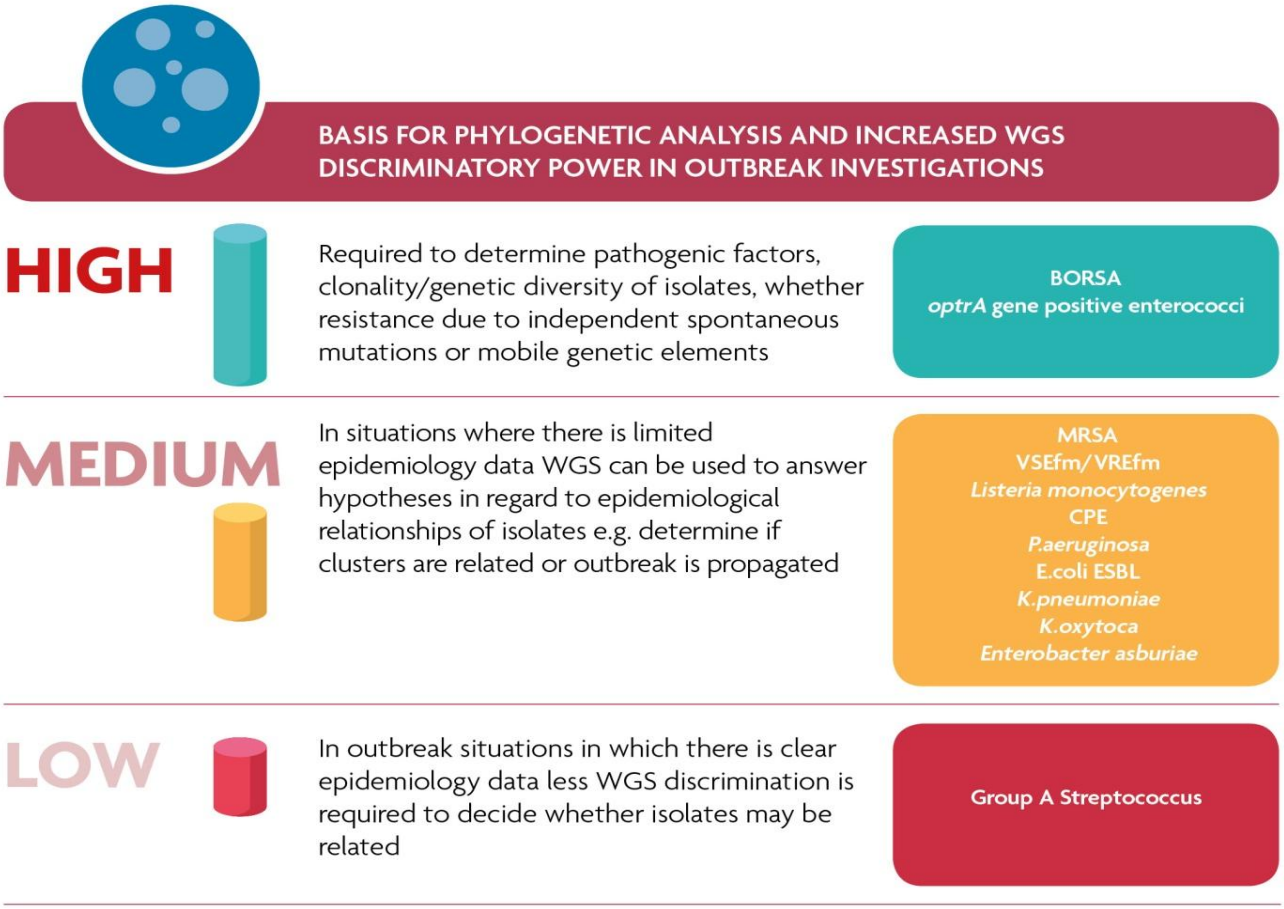


Figure 4.5. Basis for phylogenetic analysis and increased WGS discriminatory power in outbreak investigations based on the findings from this MD.

I found that in instances in which we were faced with new or unusual resistance mechanisms we required greater depth of sequencing in terms of investigating mobile genetic elements. It also gave us the opportunity to potentially develop a PCR test for BORSA screening. The majority of outbreaks required a medium level of WGS discriminatory power to determine if isolates were related. In instances in which the epidemiology demonstrated there was likely to be an outbreak less WGS discrimination was required.

4.3 Lessons learned and clinical benefits of WGS in NHS outbreak investigations

To test my second hypothesis that utilising WGS for the management of outbreaks will result in clinical benefits I have reviewed the impact of using WGS results compared to conventional typing results in suspected healthcare outbreaks from a microbiologist, IPCT and clinical care team perspective. I have observed a number of clinical benefits of translating genomics into clinical practice and these can be divided into the following 5 sections:

1. WGS can provide results with greater granularity than routine typing
2. Genomic analysis can enhance the detection of 'alert organisms'
3. WGS could replace the need for multiple tests
4. Genomic analysis can be used to rule out outbreaks
5. WGS can be utilised to investigate new resistance mechanisms

An overview of the clinical benefits of incorporating WGS into outbreak investigation can be seen in the figures 4.6 and 4.7 below.

MRSA Elderly Rehabilitation Hospital	X					M
VSEfm/VREfm	X		X			M
optrA gene positive enterococci	X			X	X	H
Listeria monocytogenes	X		X			M
GAS Ward	X					L
GAS Maternity unit	X			X		L
GAS Midwife unit	X			X		L
GAS Care home	X					L
BORSA Dermatology	X	X			X	H
CPE Renal ward	X	X				M
P.aeruginosa CF clinic	X					M
E.coli ESBL Community Hospital	X					M
E.coli ESBL Residential carehome	X			X		M
E.coli ESBL Household	X					M
E.coli ESBL Maternity	X			X		M
P.aeruginosa ICU	X	X				M
P.aeuginosa NICU	X					M
K.pneumoniae NICU	X					M
K.oxytoca NICU	X					M
Enterobacter asburiae NICU	X					M

HIIA Tool	Green	Manage within the NHS Board. Inform CPHM.
	Amber	Report to SGHD. Engage with CPHM, Consider issuing press statement
	Red	Report to SGHD. Engage with CPHM. Report to HPS. Issue press statement

WGS Whole Genome Sequencing; M Medium; H High; L Low; ICU Intensive Care Unit; MRSA Meticillin-resistant *Staphylococcus aureus;* ESBL Extended-spectrum beta-lactamase; CF Cystic Fibrosis; NICU Neonatal Intensive Care Unit; BORSA Borderline oxacillin-resistant *Staphylococcus aureus;* VSEfm vancomycin-sensitive *Enterococcus faecium;* VREfm Vancomycin-resistant *Enterococcus faecium;* CPE Carbapenemase producing *Enterobacteriaceae;* GAS Group A Streptococcus; HIIA Hospital Infection Incident Assessment; SGHD Scottish Government Health and Wellbeing Directorate; CPHM Consultant in Public Health Medicine

Figure 4.6. Overview of the clinical benefits of incorporating WGS into outbreak investigations based on the findings from this MD. Chart shows HIIA tool severity score.

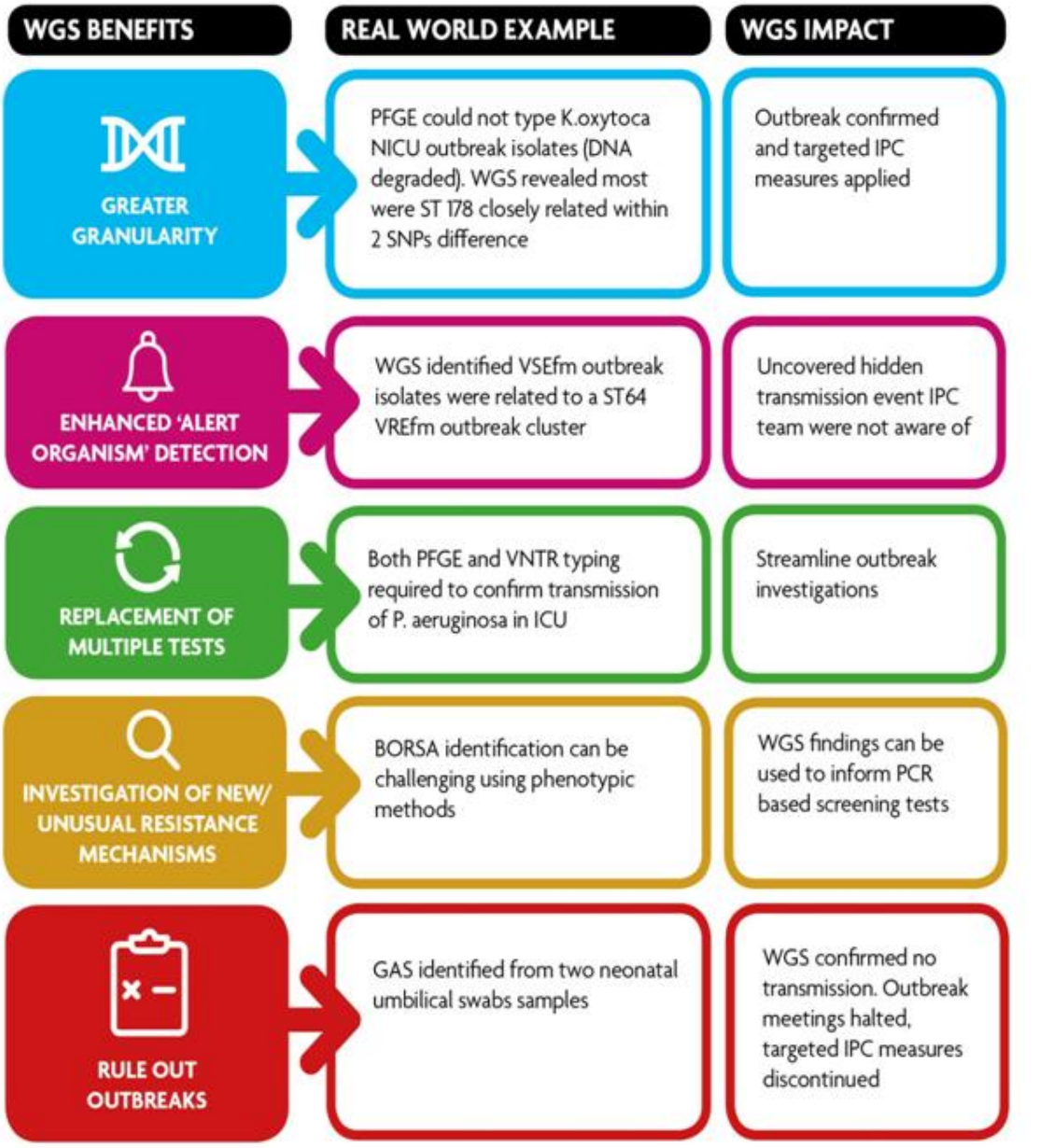


Figure 4.7. Specific real-world examples involving clinical benefits of translating genomics into clinical practice based on data generated from this MD.

4.3.1 WGS can provide results with greater granularity than routine typing

The utilisation of genomics for outbreak management can provide results with improved resolution compared to traditional typing methods. As part of the work of this thesis WGS was used to assist the outbreak team (IPCT and Public and Environmental Health Team) in a hospital outbreak investigation of two patients of invasive *L. monocytogenes* infection. In this investigation both patients were immunocompromised and had been admitted to hospital at various points. *L. monocytogenes* is a Gram-positive notifiable foodborne pathogen that can cause gastroenteritis. Listeriosis has a crude mortality rate of 20 % and in this scenario it was essential that the source of the *Listeria* was identified in order to prevent further transmission to patients, staff and the public.(231) The HPT conducted interviews with family members and one patient was identified to have a history of leaving food unrefrigerated but there were no obvious links identified between the patients. Environmental Health carried out hospital kitchen inspections but no areas of concern were identified. Routine typing identified the *Listeria* as serotype 4 (clonal complex 1 and multilocus sequence typing (MLST) 1). At this point in the investigation considering the epidemiology the outbreak team were unsure whether there had been transmission since this was one of the more common serotypes found in clinical isolates (1 in 6 invasive *Listeria* isolates are clonal complex 1). As part of this investigation both isolates of *Listeria* underwent WGS and were mapped to the reference chromosome of strain F2365 and SNPs called against this. The two isolates were found to be identical and therefore highly likely to be epidemiological linked. In this example WGS revealed additional information that prompted further action. The outbreak team had greater confidence to repeat kitchen inspections and as a result of this identified that the handling of salads and meat did not meet national recommendations and subsequently hospital catering facilities were temporary closed until remediation action was undertaken.

In the case of the largest NICU outbreak involving seven patients it was not possible to conventionally type the *K. oxytoca* outbreak isolates by PFGE as the DNA was found to degrade. One could infer that the unusual behaviour displayed by this microorganism suggests that it is the same strain. Given that the majority of suspected outbreak isolates DNA autodegraded with PFGE it would be easy to assume that this was a characteristic of an outbreak strain supporting transmission. It should be noted however that there are reports from various sources that the phenomenon of DNA autodegradation should not be considered

a clonal trait. (232) (233) DNA degradation with PFGE is not only restricted to *K. oxytoca.* There have been reports of this finding occurring with various bacteria such as *Vibrio parahaemolyticus, C. difficile*, *Ralstonia, Pseudomonas* and *Serratia.*(234)(235)(236)(233)(232) In some instances, the bacterial DNA is broken up by Tris-dependent endonuclease activity during sample preparation for PFGE leading to the terms "thiourea-dependent" or "untypeable" strains being used.(232)Various modifications have been utilised in an attempt to overcome this issue. These include adding thiourea to the electrophoresis buffer, using a lysis buffer of EDTA and SDS with plug washing with diethylpyrocarbonate (DEPC) and high temperature inactivation of DEPC using HEPES buffer electrophoresis, using formaldehyde for an inactivation step for DNase and embedding in thinner plug molds than usual, varying the incubation times for digestion, increasing Proteinase K and using different proteases.(237)(232)(238)(233) These modifications may still be unsuccessful and time consuming and thiourea is not used routinely as it is toxic. We propose that genomics should be employed in place of PFGE. We found that when pathogen sequencing was applied it revealed in a single step that the majority of *K.oxytoca* isolates were ST178 giving additional information that the isolates were closely related within 2 SNPs difference.

4.3.2 Genomic analysis can enhance the detection of 'alert organisms'

In this next example we demonstrate the benefits of using genomic analysis to detect alert organisms rather than solely accepting results from phenotypic tests. VREfm has become a leading cause of HAI. As with many other MDROs there are limited treatment options for infected patients and these may be less effective than standard treatments in some patients require monitoring due to toxic side effects. These infections have been associated with increased length of stay, higher healthcare costs and high mortality rates.(239) Management of HAI outbreaks of VREfm are challenging as patients may develop gastrointestinal colonisation which is unlikely to be eradicated by decolonisation and *Enterococcus* can survive in contaminated healthcare environments. (240) We applied WGS to the investigation of a seemingly small separate series of VREfm outbreaks on different wards. Initially two patients had been identified to be positive for VREfm in urine cultures whilst admitted to the same ward of an orthopaedic rehabilitation hospital. Antibiograms were identical and further screening samples were taken looking for faecal carriage. Of note, it was identified that one patient was colonised with two different strains of VREfm one identified from rectal swab

and another from urine, each strain related to an entirely different outbreak cluster in the main hospital. This demonstrated the importance of repeated and sequential patient sampling from different body sites including multiple colonies for sequencing during *Enterococcus* outbreak investigations. Over a two-year period further positive patients were identified on a SHDU and Renal ward of the main hospital, overall 11 patients were involved. Routine typing by PFGE of all the VREfm isolates from all patients identified that there were 5 clusters in total (three ST80 clusters, one ST64 culture and one ST203 cluster). When WGS was applied it revealed that in fact there was only one ST80 cluster in total with subtle interlinking between the ST80 isolates rather than numerous discrete clusters. Remarkably WGS and PFGE identified that two VSEfm isolates from two separate patients previously identified during a separate VSEfm outbreak in ICU the year before were related to the ST64 cluster. This was an unexpected finding uncovering a hidden transmission event that the IPCT were not aware of. The four isolates were differentiated by only 21 SNP sites suggesting a common course. The investigation also revealed that interhospital transmission had occurred between local hospitals and also a regional hospital carrying out renal transplants. Our results were produced as part of a real-time VREfm outbreak investigation and they support findings from a retrospective study in the UK in which Raven *et al.* applied WGS to the genetic characterisation of 293 VREfm bacteraemia isolates. In this study the researchers found the majority of VREfm bacteraemia isolates were hospital-acquired or healthcare-associated and over 50% of isolates were highly related. In 32% of patients complex transmissions had occurred over a number of years and across various wards, there was a mixture of vancomycin-resistant and –susceptible antibiotic profiles and this was due to isolates having lost or gained transposons carrying the gene encoding vancomycin resistance (*vanA*). As part of this investigation a vancomycin-sensitive *E. faecium* genome was scrutinised and a plasmid was identified in which *vanA* had been lost as part of a 25-kb deletion. (241) In a similar study carried out in Australia, Howden *et al.* analysed 61 *E. faecium* isolates (mainly from blood cultures) of which 36 were VREfm over a 12 year period specifically examining the *vanB* gene containing the Tn*1549* transposon. Researchers also examined five *vanB*-positive faecal anaerobic commensal bacteria from faeces. They found 8/36 VREfm had acquired the transposon via independent insertion events indicating *de novo* generation of VREfm rather than cross transmission. They also identified that the Tn*1549* sequences in 15 VREfm were the same as those found in one of the gut anaerobes and postulated that VREfm generation could occur during antibiotic therapy. (242) Taken together these findings indicate that resistance in enterococci is not stable, using this as a marker for transmission is not

reliable and could be compared to taking aim at a moving target. This suggests that a paradigm shift is required when investigating VRE outbreaks. We recommend that IPCTs should have a low threshold for including genomic approaches and should therefore consider including VSEfm and VREfm isolates in their analysis. It may be more appropriate to consider enterococci clone transmission in healthcare facilities and communities rather than focusing on resistance as a target to identify and manage outbreaks.

4.3.3 WGS could replace the need for multiple tests

CPE are notoriously challenging to detect in the clinical laboratory and multiple methods are routinely used to detect them. In an outbreak involving three renal patients infected with bla_{KPC} positive ST258 *Klebsiella pneumoniae* WGS negated the need for screening by multiplex PCR assays in two different laboratories. It could have also been used to replace a PCR assay identifying different capsular types, regulators of mucoviscosity and VNTR analysis. In addition to this, WGS revealed that one of samples was mixed with an ST3 *E. coli* infection which had not been identified by the routine clinical microbiology laboratory. Figure 4.8 shows the integration of WGS into CPE outbreak investigation

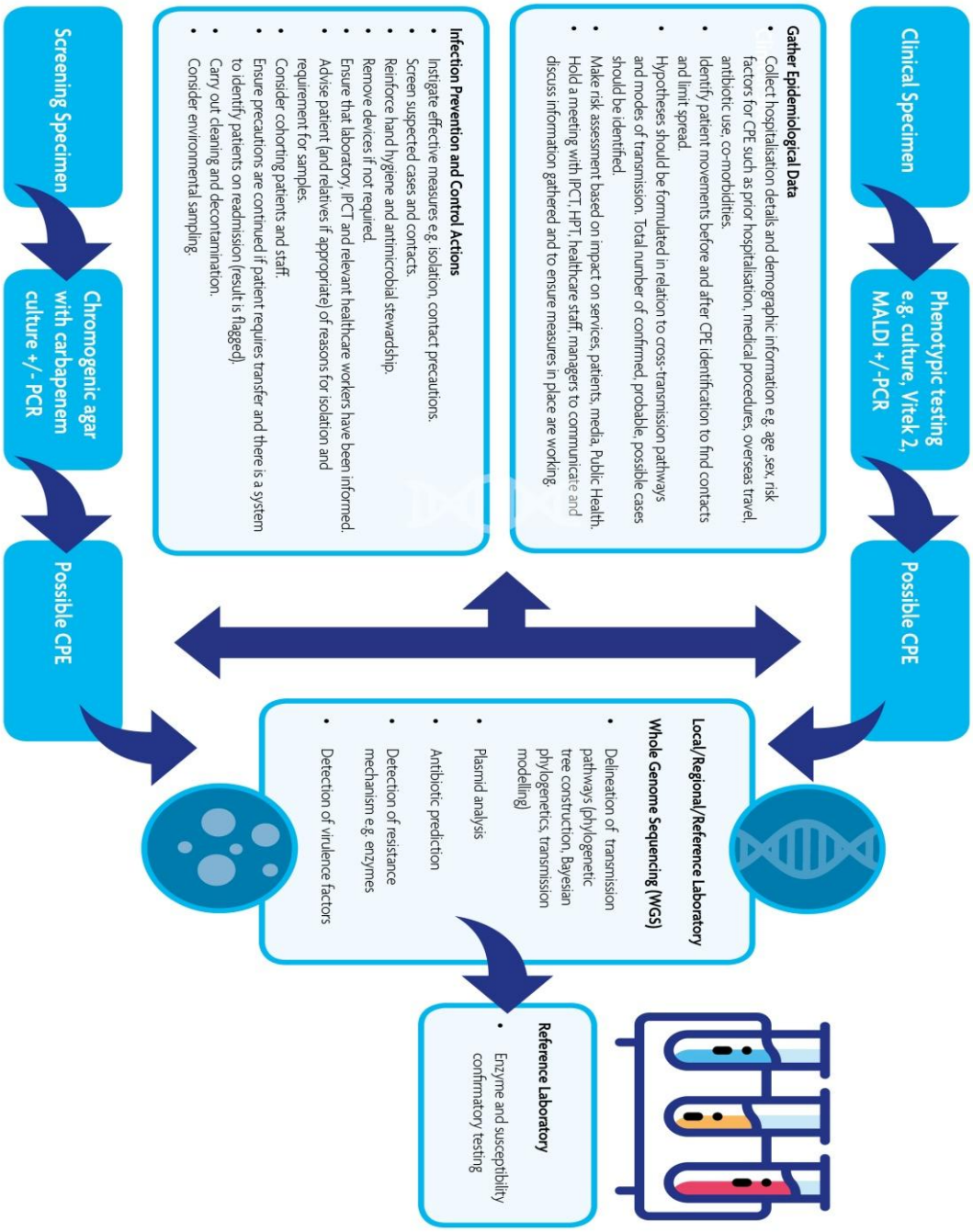


Figure 4.8. Integration of WGS into CPE outbreak investigations based on findings generated by this MD.

WGS costs are dramatically decreasing and a recent paper by Mellman *et al.* published in 2016 examined the cost of sequencing 1195 MDR isolates (MRSA, VREfm, MDR *E. coli*, and MDR *P. aeruginosa*) at a German University Hospital.(243) The authors took into account all costs for sequencing consumables, included value-added tax (VAT), depreciation over a 3 year period for their MiSeq instrument, and labour costs. They estimated that the cost of sequencing was € 202.49 Euros per isolate (£178.15) and found they saved in excess of €200,000 due to the impact of WGS on patient flow throughout the hospital.(243) I examined the cost of routine microbiology testing for one isolate (excluding staff time and WGS) in our outbreak and calculated that the overall costs for tests alone was £358.94. Integrating WGS technology into routine microbiology clinical workflow may improve turn-around times (TATS) for results and replace the need for numerous tests saving money however allocation of resources would be required to build an infrastructure to support this. There would need to be financial investment to buy equipment and train staff in sample preparation and bioinformatic interpretation. The integration of WGS technology should also be considered in the context of the cost of an outbreak for instance an outbreak of New Delhi metallo-β-lactamase (NDM)-producing *K. pneumoniae* affecting 40 patients in five hospitals across three sites in London cost €1.1m over 10 months (£876, 295.00) when costs due to staff time, bed closures, elective surgical missed revenue, anti-infective costs, enhanced CPE screening, contact precautions, ward-based monitoring of hand hygiene and environmental practice, and environmental decontamination were taken into consideration.(244). We found that WGS agreed with routine typing methods confirming transmission by ST258 but also revealed that one patient had mixed infection with *E. coli*. This supported the epidemiology and our hypothesis that this was an outbreak due to clonal spread. There is increasing evidence that WGS can be used in clinical practice to provide greater information than routine typing in CPE outbreaks. For instance in one investigation authors reported that they found WGS to be more valuable than PFGE as a typing tool as it identified numerous clades that were involved in the causing the outbreak which suggested that multiple transmission events had occurred.(245) In another outbreak authors were faced with limitations in terms of epidemiology as they could not infer whether a patient had acquired a new strain of CPE or whether transmission had occurred from an index patient. Authors explained that PFGE or repetitive PCR results did not give enough granulation to assist in determining which event had occurred. This was due to the genetic similarity of ST258 in United States hospitals and the long timeframe between patient presentations. They overcame this by constructing a putative transmission map with WGS data and overlapping patient hospital data.(246) Other

researchers have reported using WGS for the investigation of CPE and in some instances they have used Bayesian analyses of the genomic data to support epidemiological findings, develop CPE guidelines and improve outbreak management and surveillance.(216)

Capsules can resist complement-mediated lysis or phagocytosis, neutralising effect on antibodies and increase survival in the environment. In our CPE outbreak, the strain was tested by the reference laboratory for serotype specific targets for capsular types K1, K2, K5, K54 and K57 which are associated with invasive disease and was found to be negative. They did not detect mucoviscosity regulators rmpA and rmpA2 or the wcaG (a capsular fucose synthesis gene). Researchers are increasingly examining virulence factors as part of outbreak investigations to understand pathogenicity with a view to identifying measures to reduce transmission. Researchers have examined *K. pneumoniae* capsular type, iron uptake systems and fimbrial adhesions, which play a role in *K .pneumoniae* colonisation using tools such as the Pasteur MLST *K. pneumoniae* database. As in our study, Rimoldi *et al.* did not find the presence of any of the well-known hypervirulent capsular serotypes (i.e. K1, K2 and K5) suggesting that other virulence factors may play a more important role in pathogenicity during KPC outbreaks.(247). Using WGS it is also possible to identify new genetic variations underlying bacterial virulence. Authors have reported that they have found *K. pneumoniae* to contain aerobactin, enterobactin (*ent* operon), and yersiniabactin (*irp* and *ybt* genes) which play a role in iron uptake systems.(247)(248) Rimoldi *et al.* found that all clinical KPC isolates were positive for mrk operon, which encode type 3 fimbrial adhesins involved in biofilm formation on surfaces such as catheters. In their study 39.7% of patients had complicated urinary infections reinforcing the advice from national guidance that invasive devices such as urinary catheters should be removed when not required.(247) Similar findings were also found in another study tracking healthcare associated *K. pneumoniae* outbreaks in which genes for type 1 and type 3 fimbriae were positive in all isolates.(248) It has been speculated that pathogenicity determinants detected by WGS may one day be linked with clinical outcome data in order to inform antibiotic decisions such as duration of therapy or whether combination therapy is required.(249)

Plasmid replicon content is increasingly being used as a marker for bacterial typing. Short-read technologies (e.g. Illumina) and software such as PlasmidFinder-1.3 can be used to define plasmid content and track the diversity of plasmids carrying carbapenemases and other resistance genes and software such as PROVEAN can predict impact on protein function.

(247)(250)(251) Some authors have found there are benefits to using long-read sequencing technology such as Pacific Biosciences single-molecule sequencing as it can avoid some of the problems associated with short-read Illumina sequencing e.g. repetitive elements such as insertion sequences which can accompany mobile resistance elements and confound short-read assemblers and read-aligners.(216) Researchers from the UK have used short-read (Illumina) sequences for MLST and plasmid/Tn4401 fingerprinting, and long-read (MinION) sequence assemblies for plasmid identification in an outbreak involving KPC-positive *Klebsiella* and distinct KPC-producing *Enterobacterales*. By using this approach they found KPC isolates demonstrated bla_{KPC} dissemination via horizontal transposition (Tn4401a), plasmid spread (pKpQIL-D2) and clonal spread (*K. pneumoniae* ST661).(252) It has been suggested that outbreaks of plasmid transmission compared to clonal outbreaks are more likely to respond to antibiotic stewardship interventions (rather than other IPC measures).(253) With this in mind examining plasmid replicon content may become a standard of care in WGS outbreak investigation as it could be used to assist IPC teams informing them which measures to use or not to use.

Detection of CPE can be challenging in the laboratory. Interpretation of culture-based tests that detect carbapenemase release in agar media can be subjective and false-positive results can occur with certain organisms.(254) ST258 has a characteristic antibiogram with susceptibility generally only to colistin, gentamicin and tigecycline. We found that KPC isolates taken from patients in this outbreak reflected this antibiogram pattern.(253) Treatment options for CPE are unfortunately few and there are concerns that further resistance can develop. Patient C developed tigecycline resistance whilst on treatment and this has been described previously.(254) (255) WGS has the potential to predict antimicrobial susceptibilities and has been found to correlate well with phenotypic antibiotic sensitivity testing. (256)(257) In regard to tigecycline resistance researchers have used ResFinder-2.1 software to examine AcrAB-TolC efflux pump-dependent and pump-independent tigecycline resistance mechanisms (acrR, ramR, marR, soxR, lon, rpsJ and rpoC) in Italian KPC hospital isolates which they had taken over a 3 year period. They found that the most frequently observed mutations were SNPs. However gene sequences varied in terms of their presence, mutations and deletions and the authors were not able to correlate this with tigecycline resistance due to the presence of a common mutation in susceptible, intermediate and resistant strains.(247)

The *P. aeruginosa* outbreak in an adult ICU can be used to illustrate how WGS can streamline outbreak investigations as WGS could have been used in place of two typing techniques (VNTR and PFGE). The conventional methods, antibiotic susceptibility, VNTR and PFGE, grouped the isolates in different ways. Both VNTR and PFGE correctly identified the two patients who were part of the outbreak but identified several false positive environmental links. WGS alone provided the necessary resolution to identify the transmission pathway, demonstrating unequivocally spread between single water supply to patients, and eliminating other potential transmission events and sources.(69) An overview of conventional typing and genomic analysis and the timeline is illustrated in Figure 4.9

.

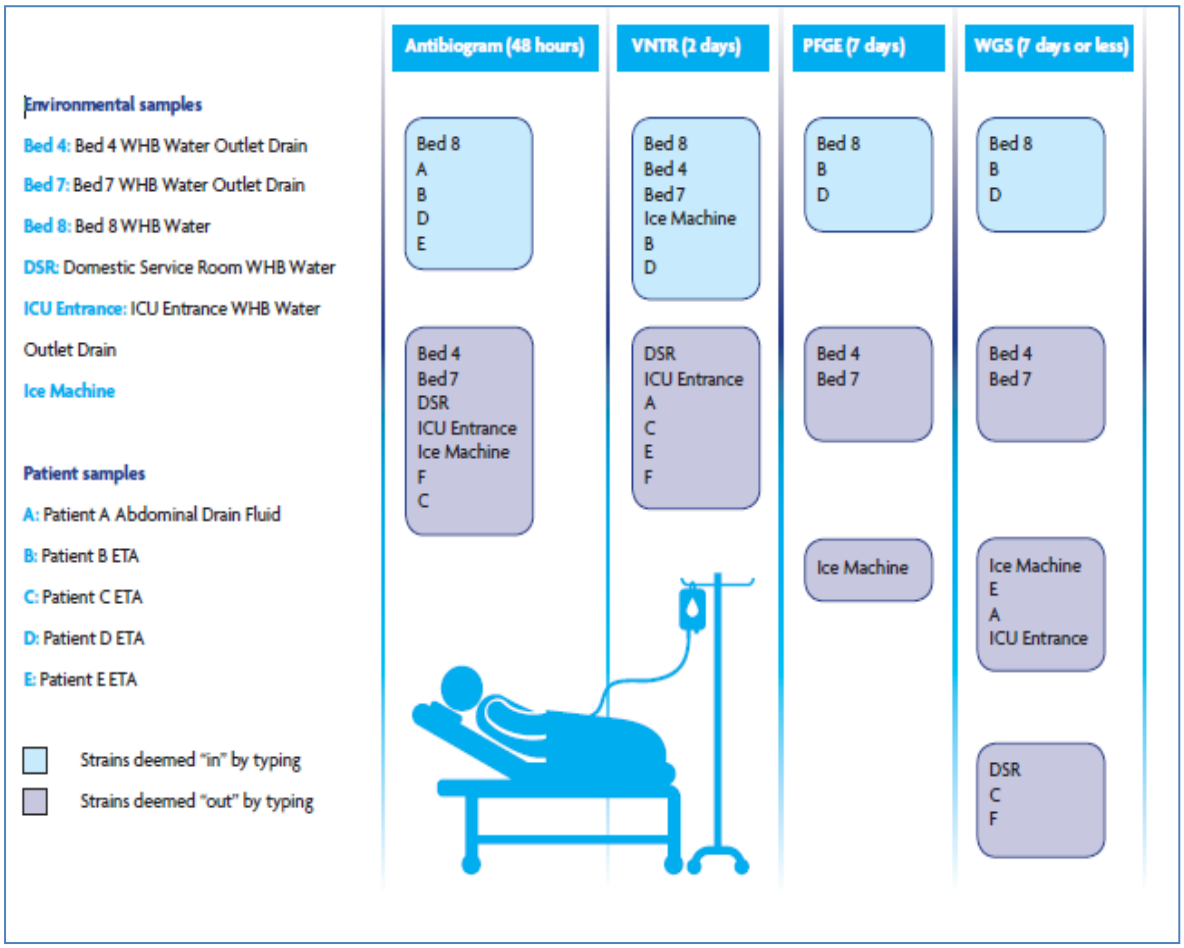


Figure 4.9. An overview of conventional typing, genomic analysis and turn-around times for an outbreak of *P.aeruginosa* in an ICU. This figure was created using data generated by this MD and information collected from the IPC team in NHS Tayside.

4.3.4 WGS for ruling out outbreaks negating the need for outbreak meetings

In several instances we found that WGS could also be used by the IPCT to rule out outbreaks averting disruption to services. WGS in these instances mainly supported routine typing results. An illustration of this can be given with the following outbreak in which two GAS were identified from babies' umbilical swabs. GAS is a virulent organism able to produce toxins resulting in sepsis and death. Suspicion was raised that this could be an outbreak as samples were taken a day apart. An IPCT investigation revealed that the babies had been born a week apart and admitted to the same labor suite. Patients were both treated and, in this instance, *emm* typing revealed they were 89.0 and 6.4. WGS showed the two GAS isolates were different (ST101 and ST382) both differing in all 7 alleles and IPC measures were discontinued.

4.3.5 WGS can be utilised to investigate new resistance mechanisms

The following example illustrates the usefulness of WGS as a first line tool for the investigation of emerging resistance mechanisms. The *optrA* is an ABC transporter gene encoding resistance to oxazolidinones (linezolid and tedizolid) and phenicols such as chloramphenicol and florfenicol via active efflux. In 2015, Wang and colleagues first reported evidence that this resistance gene could be present in *E. faecium* and *E. faecalis* in humans and animals using whole-plasmid sequencing.(258) National Resistance Alerts have been issued highlighting the risk to public health as this new resistance mechanism is plasmid-mediated and could potentially transfer to other strains, species and genera present on the skin and gut of humans and animals. Alerts have advised that standard IPC measures alone are not effective enough to prevent spread and patients with linezolid-resistant enterococci or staphylococci should be isolated to prevent onwards transmission.(259)(260)(261). In 2016, an IPCT investigation was undertaken as Public Health England's Antimicrobial Resistance and Healthcare Associated Infections (AMRHAI) Reference Unit confirmed a patient had a urine sample which was positive for *optrA* linezolid resistant *Enterococcus*. A retrospective search of the hospitals historical laboratory culture results identified two further patients to have *E. faecalis* isolates that were linezolid resistant – identified in 2014 and 2015. These isolates were initially reported as *cfr* gene negative – which was the plasmid-borne mechanism of resistance known at that time. We applied WGS

to investigate these isolates and found that both were positive for the *optrA* gene. A review of the epidemiology revealed that all patients had urinary tract infections with one patient admitted to hospital and two patients managed in the community. WGS was useful in this situation as it identified that that the isolates were of 3 distinct sequence types (ST480, ST19, and ST330). Applying WGS to the investigation of these isolates assisted the IPCT in their investigation as it supported their hypothesis that linezolid resistance had emerged independently and was not due to a circulating clone. Interestingly the patient who had an ST19 isolate was a beef cattle farmer. A recent analysis of linezolid resistant *Enterococcus* isolates from patients with links to Polish hospitals showed that resistance was predominantly caused by independent *de novo* resistance mostly representing high-risk enterococcal clonal complexes (HiRECCs). Mutation of the 23S rRNA gene was the most common determinant of resistance. The *optrA* gene was found in two epidemiologically unrelated non-HiRECC ST116 *E. faecalis*. The researchers identified that the *optrA*-carrying plasmid was transferable from *E. faecalis* to both *E. faecium* and *E. faecalis* recipients. They highlighted the importance of one health surveillance postulating that ST116 may act as a vehicle for antimicrobial resistance movement between environment and hospital since these isolates have been found from both human infections as well as food-producing animals and retail meat.(262) Being able to determine whether resistance genes are likely to transfer is also useful for IPCT and HPT when justifying the implementation of measures to prevent and control infection. Routine Microbiology laboratories at present cannot confirm the *optrA* gene resistance and this example also demonstrates the importance of implementing enhanced IPC procedures when resistance mechanisms are unknown.

4.4 Evidence that supports the utility of WGS in HAI investigations

Twenty one suspected outbreaks were examined in this study. To assess the utility of WGS in HAI outbreak investigations I analysed WGS output versus standard reference laboratory typing to determine additional information provided by WGS. I then compared WGS sample preparation/result generation time with routine reference typing to provide evidence that supports the utility of WGS for outbreak investigations (Appendix 6 and Appendix 7).

WGS notably had direct impact on IPC management when results were produced in real-time throughout an outbreak. When results were presented to the clinical team and IPC team

during a BORSA outbreak meeting there was greater engagement of clinical staff and results informed a range of layered mitigations such as screening and environmental testing.

WGS results were produced before routine reference typing results in 4 suspected outbreaks. This impacted on IPC management each time and this occurred in different ways. In one example rapid WGS results during an ESBL producing *E.coli* outbreak in a community hospital influenced the IPC team to introduce a range of outbreak measures including screening of patients. In one suspected outbreak of ESBL producing *E.coli* in a Maternity unit rapid WGS was used to rule the outbreak out saving time from holding an outbreak meeting. In this instance WGS reassured the clinical staff transmission had not taken place. In a suspected outbreak of *P.aeruginosa* in NICU WGS allowed the IPC team to open the ward earlier. In another instance WGS of *K.oxytoca* in NICU assured the IPC team there was an outbreak and IPC measures should remain.

The production of WGS results was slower than routine typing in seven suspected outbreaks. The results were found to be valuable by the IPC team in three of these instances. In a suspected outbreak of *L.monocytogenes* WGS confirmed that transmission had taken place between two patients and the IPC team decided to repeat kitchen inspections. In another instance the IPC team found WGS results to be useful as they supported their decision to do intensive patient screening during a CPE outbreak. Finally WGS results gave greater assurance an outbreak meeting was not required saving staff time when WGS was used to investigate an ESBL producing *E.coli* outbreak in a residential care home.

WGS results were significantly delayed and not produced in real-time due to batching in eight of the suspected outbreaks. In each of these instances there was no direct impact on IPC management or clinical care which demonstrates the need for rapid WGS to inform IPC management of outbreaks.

4.5 Methodology and recommendations of how best to establish a WGS service for the investigation of HAI outbreaks

The overarching objective of this study was to establish a WGS service for the investigation of suspected HAI outbreaks. As part of this process I have identified practical barrier to this and the instances in which increased WGS discriminatory power is required (objective 2), along with the clinical benefits of WGS in this manner. Meeting these objectives has informed the development a clinical decision aid for staff such as IPCNs, experts in Microbiology and Infectious diseases and Public Health on how best to utilise a WGS service (Figure 4.10)

Figure 4.10. Clinical decision aid on how best to utilise a WGS service for the investigation of HAI outbreaks in real time based on data generated from this MD.

Initially an outbreak or healthcare infection incident should first be determined to have occurred (step 1). It is traditionally accepted that an HAI outbreak is defined as: *"two or more linked cases with the same infectious agent associated with the same healthcare setting over a specified time period";* or a *"higher than expected number of cases of HAI in a given healthcare area over a specified time period"*.(72) The speed at which outbreaks emerge differs by organism and per ecological/ clinical setting. It is recognised that reductions in staff numbers and application of hand hygiene can serve as an early indictor prior to events. Clustered organisms can arise sufficiently far apart in time or location that they are not identified therefore local surveillance systems should be set to have a trigger/ threshold that prompts IPC action. When a trigger point is reached this may indicate an outbreak but some triggers may be due to natural variation in alert organism incidence. With triggers, it is easier to detect and respond to issues early. Prevalence of alert organisms varies depending on the patient population and healthcare risk factors and so the same trigger levels cannot be set for all clinical areas. Benefits of setting trigger levels include that they signal problems early, the signal can be detected at any time, it can result in a set of actions and be used to communicate and feedback information to clinical staff.(65) Figure 4 illustrates the optimal protocol for sequencing assembly and mapping that we found during this study and demonstrates how WGS can be incorporated into conventional outbreak analysis workflows for *in-silico* outbreak investigation. We suggest that the process of outbreak identification should be automated by developing systems that function without human coordination with set "action line" rules. This can be brought about by collecting epidemiology data, using SPC charts and implementing outbreak surveillance software such as the ICNet.

Step 2 involves determining whether sequencing will impact on patient care. Alert organisms include bacteria that are deemed to pose increase risk to patients. This can be due to virulence factors such as toxins e.g. GAS or resistance mechanisms that render them more difficult to treat e.g. VRE, MRSA, ESBL positive *Enterobacterales*, CPE. IPCTs should also have a lower threshold for acting on certain organisms in high risk areas for instance *P. aeruginosa* in NICU. We have shown there are clinical benefits to applying WGS to the routine investigation of outbreaks involving alert organisms. Additionally, it is a valuable tool for the IPCTs when these occur in high risk clinical areas and prospective sequencing of vulnerable populations such as ICU patients can be used for surveillance. Other instances where there is value include situations in which there are issues with phenotypic testing or lack of

granularity with traditional typing methods and when resistance mechanisms are unknown or emerging.

IPCTs want clear information and rapid identification of outbreaks. In instances where outbreaks are confirmed optimal intervention strategies and early intervention can be put in place before the organism becomes endemic. Additionally IPCTs also require a rapid conclusion that there is not an outbreak so that the outbreak group meetings can be stood down and targeted measures discontinued freeing hospital services from closure and disruption. Step 3 involves the interpretation of WGS results. Care needs to be taken in relation to this in particular in regard to the meaning of SNP differences. The molecular evolution rate also known as the "molecular clock'" of bacteria can significantly vary throughout bacteria genomes. This is due in some bacteria to mutation being highly regulated by DNA repair mechanisms. In highly expressed genes there can be lower rates of mutations. There can be difference within species and this can also differ temporally due to life cycles for instance *M. tuberculosis* can be active or latent with different rates of evolution. (263) In *Clostridium* endospores can be produced, when this occurs there is likely to be less mutations compared to dividing cells. (263) Researchers have also described that the "molecular clock" for *S. aureus* may also vary depending on whether bacteria are present causing infection or carriage within hosts. (264) Hypermutation can also occur in which DNA repair pathways are modulated resulting in an increased in substitution rate. This can be an advantage to bacteria as they can gain resistance to antibiotics. (263)The application of WGS to the investigation of a NICU MRSA outbreak identified that one isolate had a hypermutator phenotype resulting in a higher than expected number of SNPs. The authors concluded that for this reason it would be difficult to impose thresholds for the number of SNPs between isolates to confirm recent transmission.(161) These factors need to be considered in the context of pathogen genome stability and the environment. For instance, the literature reports that various SNP differences have been found during the investigation of *Listeria* outbreaks and these can range from as small as 7SNPs to 42 SNPs. (265).(266) Public Health England (PHE) have themselves observed that *Listeria* isolates in outbreaks linked to a single food premises can be as much as 20 SNPs apart which is in contrast to findings in verocytotoxigenic *E. coli* and *Salmonella* incidents where only isolates within a 5 SNP cluster would be considered to be linked "this is because *Listeria* spp. can remain as an environmental contaminant in premises over a long period of time (years)". (Gill Hawkins, Health Protection Scotland, Personal Communication 1 September 2016).

It is also important to consider how long a patient may be infected and their medical conditions. In a study investigating genotypic diversity of *P.aeruginosa* in a CF patient investigators examined 44 morphologically identical isolates from a sputum sample.(267) WGS was carried out on 22 isolates and researchers found high levels of isolate diversity (5 to 64 SNPs). Recombination was mainly responsible for this diversity not spontaneous mutation. The authors advised that within-patient diversity for pathogens needs to be defined to accurately inform contact transmission networks or researchers should consider multiple colony or whole population sequencing from every sample. They warned that using a cut-off of 5 SNPs or for MRSA 11 SNPs may not always be an accurate approach.(267)

The investigation and management of outbreaks can be challenging. To prevent spread investigators must be confident that their hypotheses in relation to cross-transmission pathways and modes of transmission are correct and IPC measures in place are effective (step 4). WGS can be used to aid IPCTs in their decisions and this can be enhanced by putative transmission mapping and by Bayesian analyses of genomic data. The application of WGS has led researchers to develop greater insights into the dissemination of AMR (horizontal transposition, plasmid spread and clonal spread). It can also be utilised to develop outbreak-specific multiplex PCRs to control the dissemination of resistant organisms for instance ESBL producing *K. pneumoniae*.(268)

4.6 Strengths of the study

A major strength of this study is that we have applied WGS to a significant number of outbreaks (21 in total) over a five-year period. The outbreaks involve a wide range of Gram positive and Gram negative bacteria possessing a wide range of antimicrobial resistance mechanisms e.g. MRSA, ESBL, *optrA* gene positive enterococci, CPE, BORSA and VREfm. The outbreaks also involved a range of patients within different healthcare settings. The high number and variety of outbreaks supported me in meeting the study's objectives and testing the hypotheses. This resulted in the successful establishment of a WGS service that could be utilised to react to a wide variety of suspected outbreak scenarios. I was involved in managing all of the outbreaks which provided continuity in interpreting the results of WGS and determining clinical impact and challenges. From this, I was able to identify the different

practical barriers to WGS implementation, WGS clinical benefits and determine when increased WGS discriminatory power was required in different outbreak investigations.

4.7 Limitations

We did not utilise WGS for the investigation of suspected outbreaks of *C. difficile*. In addition to this, during outbreak investigations staff and the environment were not routinely screened and in some instances patient samples were not available for analysis.

4.8 Future work

In recent years WGS has been implemented as a standard typing technique by reference laboratories for pathogens such as *E. coli, Shigella Salmonella, Listeria, Campylobacter, S. aureus, Salmonella* and *Mycobacteria*.(269)(79)(80)(81) However, it has not yet been widely adopted across NHS clinical laboratory microbiology services for outbreak detection for reasons such as high costs and lack of expertise. As part of this research I have demonstrated there is clinical value for its use. Recommendations that local clinical microbiology departments should be integral to IPC programmes assisting in local surveillance and arranging molecular typing of pathogens date back to 1998 and since then there has been significant developments in surveillance systems.(270) Whilst carrying out this MD other researchers have also published work calling for a paradigm shift in which genomic surveillance is used as part of outbreak investigations.(271) Further research is required assessing the best ways to apply this potentially disruptive technology into the system of outbreak detection. Considering the importance of ensuring that the system we set up was an end-to-end WGS service which was linked to the diagnostic microbiology and IPC further work could be carried out assessing the incorporation of outbreak web applications which can be used to investigate and/or display pathogen genomic information such as Shiny, Outbreak Tools and Microreact.(61)(62)(63)(64) We found that the value in producing results was also dependent on providing results in clinically relevant time frame and further research could be carried out assessing the role of Pathogenwatch which can provide rapid predictions of clustering and genotype resistance.(217) Further work could also assess colony pick sequencing, a newly described approach which has been shown to overcome the delay associated with performing pure cultures and DNA extraction.(272)

With researchers from the University of Dundee I became a co applicant for a successful application for £2.12M funding from the 2018 DHSC/NIHR AMR Infrastructure Award to enhance bacterial genomics capability and linkage to health informatics at the Department of Population Health and Genomics, School of Medicine, University of Dundee. Since this has been awarded I have been involved in the selection of new technology such as a GridION and giving advice in relation to laboratory refurbishment in order to provide a contiguous facility focusing on a bacteria-specific workflow with a sample entry portal from the NHS Microbiology department. As part of this research we will establish a local bacterial genomics service situated next to the NHS Microbiology department. Using the sequencing approaches described in this MD (reactive, proactive, and reflective) we will sequence clinically important pathogens from suspected outbreaks. In addition, we are currently carrying out genomic study in which WGS will be applied to all *E. coli* bacteraemia isolates taken in 2019. Genomic data will be linked to IPC databases and the Health Informatics Centre electronic health records. This research will increase understanding of strain relatedness, genotypic versus phenotypic resistance and virulence traits and patient risk factors for *E. coli* bacteraemia and the relative contribution of bacterial, patient and healthcare factors, to inform IPC and effective use of current antimicrobials.

I am currently carrying out an evaluation of the role of the IR Biotyper® (Bruker Optics-Daltonics GmbH) in a clinical environment as part of outbreak investigation in real-time. In particular, focusing on determining the discriminatory power, typing speed and cost of this new technology for the characterisation of *E. coli* isolates potentially involved in a hospital outbreak. These findings will be compared to those obtained via either PGFE or WGS.

During the SARS-Cov-2 pandemic I have been managing nosocomial outbreaks of SARS-Cov-2 whilst collecting and analysing clinical and epidemiological data for the COVID-19 Genomics UK (COG-UK) Consortium. This data has been used with genomic data to understand the multiple introductions of SARS-CoV-2 from mainland Europe into Scotland and will be used to examine the epidemiology of the new B.1.1.7 lineage. (83) In the first 5 months COG-UK sequenced over 40, 000 SARS-CoV-2 genomes (approximately 50% of the global total).(273) Additionally a suite of tools have been developed by COG-UK researchers such as Civet (cluster investigation and virus epidemiology tool) and Pangolin (phylogenetic assignment of named global outbreak lineages) both of which are open-source tools allowing

rapid sharing of data.(273)(274)(275) Many of these tools are hosted on the CLIMB-BIG-DATA project (Cloud Infrastructure for Big Data Microbial Bioinformatics), an initiative that allows use of a dynamic bioinformatics platform for academic research groups funded by the Medical Research Council (MRC).(273) This response to the SARS-CoV-2 pandemic demonstrates the significant sequencing capacity that can be created when academic and NHS laboratories combine their efforts to provide genomic insights that can be used to inform public health and IPC management strategies.

5 Conclusion

5.1 Proposals for the future

In this chapter I have discussed the process of successfully establishing a WGS service for the investigation of suspected HAI outbreaks using NHS clinical bacterial isolates and sequencing facilities operated by the Infection and Global Health Research Division, School of Medicine, University of St Andrews. I have discussed the challenges faced when establishing such a service. These include infrastructure; performance and quality assessment of data and processing; pipelines and management of reference databases; when to sequence; clinical interpretation of results and when to use increased WGS discriminatory power in outbreak investigations.

Through the process of setting up a WGS service I have identified that there is a range of possible way in which to incorporate WGS into outbreak investigations. It can be utilised as part of a reactive automated response to epidemiology that suggests there is an outbreak or used proactively to detect an outbreak regardless of epidemiology for instance prospectively sequencing select populations of patients from a critical care location such as NICU. A reflective approach could be used during outbreak investigations and could include incorporating WGS when epidemiology data is missing or in scenarios in which phenotypic testing lacks granularity.

I have also identified a number of clinical benefits and lessons learned from translating genomics into clinical practice: WGS can provide the ultimate discrimination of results enabling IPCTs to carry out outbreak investigations efficiently and effectively. This is particularly important in scenarios in which deciphering transmission has been challenging e.g. outbreaks involving multiple strains, movement of different resistance elements and instances in which there is unclear patient epidemiology. IPCTs can use this technique as an adjunct for the detection of alert organisms rather than routinely accepting phenotypic tests results. Real-time sequencing can streamline clinical microbiology services by reducing unnecessary multiple testing of bacterial isolates. Its value also lies in testing effectiveness of IPC measure and the ruling out of outbreaks which negates the requirement for outbreak meetings and disruption to healthcare services e.g. ward closure. WGS is a valuable tool that can be utilised to generate greater understanding of the development of new resistance

mechanisms and dissemination of resistance elements. Although WGS platforms are rapidly advancing this technique at present has not yet fully crossed the divide between research tool and routine clinical diagnostic microbiology technique. In this thesis, I have illustrated how WGS can be incorporated into conventional outbreak analysis workflows for *in-silico* outbreak investigation.

The real challenge has been utilising this technology in real-time and establishing an end-to-end process in order to have the greatest clinical impact. By identifying the practical barriers we faced and the clinical value of WGS for IPCTs, microbiologists and clinical care teams I have been able to make informed recommendations of how best to establish a WGS service. Using this information, I have developed a clinical decision aid for staff to get the most from WGS facilities. It is likely that integrating WGS into routine microbiology clinical workflows either locally, regionally or as part of reference laboratory work will become more achievable as the technology becomes more portable, and bioinformatic analysis tools become more user friendly and less costly. This work shows that WGS can give us unprecedented ability to make real-time decisions based on hard facts informing the best use of increasingly pressurised healthcare resources to maximise safety for patients, healthcare staff and the public.

6 Publications and presentations

6.1 Publications

Parcell BJ, Oravcova K, Pinheiro M, Holden MTG, Phillips G, Turton JF, Gillespie SH. *Pseudomonas aeruginosa* intensive care unit outbreak: winnowing of transmissions with molecular and genomic typing. J Hosp Infect. 2018 Mar;98(3):282-288. doi: 10.1016/j.jhin.2017.12.005. Epub 2017 Dec 8. PMID: 29229490; PMCID: PMC5840502 (Appendix 8).

Parcell BJ, Gillespie SH, Pettigrew KA, Holden MTG. Clinical Perspectives in Integrating Whole Genome Sequencing into the Investigation of Healthcare and Public Health Outbreaks - Hype or Help? J Hosp Infect. 2020 Nov 9:S0195-6701(20)30509-0. doi: 10.1016/j.jhin.2020.11.001. Epub ahead of print. PMID: 33181280 (Appendix 9).

6.2 Oral presentations

"Establishing a WGS service for outbreaks in the NHS". Oral presentation via Microsoft Teams, 26th August, 2021, Scottish Microbiology Reference Laboratories, Glasgow.

"Rapid Molecular Diagnostics and Sequencing in a Clinical Environment." Oral presentation via Microsoft Teams, 1-2 December 2020, UK Group A *Streptococcus* meeting.

"Clinical Perspectives in Integrating Whole Genome Sequencing into the Investigation of Healthcare and Public Health Outbreaks", Oral presentation at the SHAIPI Scientific Meeting Friday 8 November 2019, Hilton Glasgow Grosvenor Hotel.

"Integrating Research into Clinical Microbiology and Infection Prevention and Control in the NHS". Oral presentation at the national Workshop "PiCLS 2019 symposium" (PhD Students' Association in the School of Life Sciences at the University of Dundee), 14 September 2019, Westpark Conference Centre, Dundee.

"Whole Genome Sequencing Service for the Investigation of Healthcare Associated Infection" SHAIPI Annual Research Symposium - 13 June 2019. Kelvin Hall Conference Centre, Glasgow.

"Whole Genome Sequencing in Neonatal Intensive Care Units" Oral presentation at 29th European Congress of Clinical Microbiology & Infectious Diseases (ECCMID), 13 - 16 April 2019, RAI Amsterdam.

"Whole Genome Sequencing Service for the Investigation of Healthcare Associated Infection" Oral presentation at the SHAIPI Workstream 1 Study Group, 25-27 February 2019, The Burn, Scotland.

"*Pseudomonas aeruginosa* Intensive Care Unit outbreak: winnowing of transmissions with molecular and genomic typing" Oral presentation at the national Workshop "Genomic in the Real World Study day", 31 May, 2016, University of St Andrews,

6.3 Posters and Abstracts

1. **Parcell BJ**, Gillespie SH, Pettigrew KA, Holden MTG. Whole genome sequencing in neonatal intensive care units – sophisticated typing for outbreaks of high consequence, European Congress of Clinical Microbiology & Infectious Diseases (ECCMID) Amsterdam, 2019
2. **Parcell BJ**, Gillespie SH, Pettigrew KA, Holden MTG**.** Translating whole genome sequencing into clinical practice for the real-time investigation of carbapenemase-producing *Enterobacteriaceae* outbreaks, European Congress of Clinical Microbiology & Infectious Diseases (ECCMID) Amsterdam, 2019
3. Mchugh M, **Parcell B**, Pettigrew KA, Toner G, Khatamzas E, Karcher AM, Walker J, Weir R, Meunier D, Hopkins K, Woodford N, Templeton K, Gillespie SH, Holden MTG. Genomic investigation of linezolid-resistant *Enterococcus faecalis* carrying the *optrA* gene, European Congress of Clinical Microbiology & Infectious Diseases (ECCMID)Amsterdam, 2019
4. **Parcell B**, Gillespie S, Holden M. Unravelling the benefits and barriers to utilising whole-genome sequencing in the investigation of outbreaks, Federation of Infection Societies (FIS), Edinburgh 2019
5. **Parcell B**, Pettigrew KA, Oravcova K, McHugh M, Holden MTG, Gillespie SH. Establishing a Whole Genome Sequencing Service for the Investigation of Healthcare Associated Infections, Healthcare Infection Society (HIS), Liverpool, 2018
6. **Parcell B**, Carmichael I, Dalrymple L, Buick-Clarke C, Lee J, Harkins C, Holden MTG, Dickson E, Hearn R, Karcher AM. Use of a Rapid Molecular Test to Detect *Staphylococcus aureus* Environmental Contamination During an Outbreak Healthcare Infection Society (HIS), Liverpool, 2018
7. Dalrymple L, Lee J, Hearn R, Harkins C, Buick-Clarke C, Holden M, Dickson E, , Karcher AM, **Parcell B**. Surveillance of a Borderline Oxacillin-Resistant *Staphylococcus aureus* Outbreak in a Dermatology Unit, Infection Prevention Society (IPS) Conference, Glasgow, 2018
8. Dalrymple L, Lee J, Hearn R, Harkins C, Buick-Clarke C, Holden M, Dickson E, **Parcell B**, Karcher AM. An Outbreak of Borderline Oxacillin-Resistant *Staphylococcus aureus* in a Dermatology Unit, Infection Prevention Society (IPS) Annual Conference, Glasgow, 2018

9. **Parcell BJ**, Oravcova K, Pinheiro M, Holden M, Phillips G, Turton GF, Gillespie S. The Application of Whole Genome Sequencing for the Epidemiological Investigation of an Outbreak of *Pseudomonas aeruginosa* in an Intensive Care Unit. NHS Research Scotland (NRS) Annual Conference Glasgow 2015 and Healthcare Infection Society (HIS), Lyon, 2014

6.4 Award

Healthcare Infection Society (HIS) Early Career Award 2020

7 References

1. Pearson A. Historical and changing epidemiology of healthcare-associated infections. J Hosp Infect. 2009;73(4):296–304.
2. Bennett PN. Alexander Gordon (1752-99) and his writing: Insights into medical thinking in the late eighteenth century. J R Coll Physicians Edinb. 2012;42(2):165–71.
3. Levin BR, Lipsitch M, Bonhoeffer S. Population Biology, Evolution, and Infectious Disease: Convergence and Synthesis. Science. 1999;283:806–9.
4. Last JM. Dictionary of epidemiology. 4th ed. New York: Oxford University Press; 2001. 61 p.
5. Riley LW, Blanton RE. Advances in Molecular Epidemiology of Infectious Diseases: Definitions, Approaches, and Scope of the Field. Microbiol Spectr. 2018;6(6):1–12.
6. Riley LW. Molecular Epidemiology of Infectious Diseases Principles and Practices. 1st ed. Washington: ASM Press; 2004. 1–348 p.
7. Peterson LR, Noskin GA. New technology for detecting multidrug-resistant pathogens in the clinical microbiology laboratory. Emerg Infect Dis. 2001;7(2):306–11.
8. Peterson LR, Brossette SE. Hunting health care-associated infections from the clinical microbiology laboratory: Passive, active, and virtual surveillance. J Clin Microbiol. 2002;40(1):1–4.
9. Goodwin S, McPherson JD, McCombie WR. Coming of age: Ten years of next-generation sequencing technologies. Nat Rev Genet. 2016;17(6):333–51.
10. Long SW, Williams D, Valson C, Cantu CC, Cernoch P, Musser JM, et al. A genomic day in the life of a clinical microbiology laboratory. J Clin Microbiol. 2013;51(4):1272–7.
11. Tagini F, Greub G. Bacterial genome sequencing in clinical microbiology: a pathogen-oriented review. Eur J Clin Microbiol Infect Dis. 2017;36(11):2007–20.
12. Mellmann A, Bletz S, Böking T, Kipp F, Becker K, Schultes A, et al. Real-time genome sequencing of resistant bacteria provides precision infection control in an institutional setting. J Clin Microbiol. 2016;54(12):2874–81.
13. Vogel U, Szczepanowski R, Claus H, Jünemann S, Prior K, Harmsen D. Ion torrent personal genome machine sequencing for genomic typing of Neisseria meningitidis for rapid determination of multiple layers of typing information. J Clin Microbiol. 2012;50(6):1889–94.
14. Reuter S, Harrison TG, Köser CU, Ellington MJ, Smith GP, Parkhill J, et al. A pilot study of rapid whole-genome sequencing for the investigation of a Legionella outbreak. BMJ Open. 2013;3(1):1–6.
15. Genestet C, Tatai C, Berland JL, Claude JB, Westeel E, Hodille E, et al. Prospective whole-genome sequencing in tuberculosis outbreak investigation, France, 2017-2018. Emerg Infect Dis. 2019;25(3):589–92.
16. Selwyn S. Hospital infection: the first 2500 years. J Hosp Infect. 1991;18(SUPPL. A):5–64.
17. Atkinson J, Chartier Y, Lúcia Pessoa-Silva C, Jensen P, Li Y, Seto W-H. Natural Ventilation for Infection Control in Health-Care Settings WHO Library Cataloguing-in-Publication Data: Natural ventilation for infection control in health-care settings. 2009;1.
18. Smith PW, Watkins K, Hewlett A. Infection control through the ages. Am J Infect

Control. 2012;40(1):35–42.
19. World Health Organization (WHO). Report on the Burden of Endemic Health Care-Associated Infection Worldwide. WHO Libr Cat Data. 2011;40.
20. Scottish Government. Health and Social Care Delivery Plan. 2016;(December):1–39. Available from: http://www.gov.scot/Resource/0051/00511950.pdf
21. Cairns S, Gibbons C, Hay A, King H, Llano M, Macdonald L, et al. National Point Prevalence Survey of Healthcare Associated Infections and Antimicrobial Prescribing 2016. Heal Prot Scotl [Internet]. 2017; Available from: http://www.hps.scot.nhs.uk
22. European Centre for Disease Prevention and Control (ECDC). Summary : Point prevalence survey of healthcare- associated infections and antimicrobial use in European hospitals 2011 – 2012. 2013.
23. European Centre for Disease Prevention and Control (ECDC). Annual epidemiological report on communicable diseases in Europe 2010. Eur Cent Dis Prev Control. 2010;75–7.
24. Glance LG, Stone PW, Mukamel DB, Dick AW. Increases in mortality, length of stay, and cost associated with hospital-acquired infections in trauma patients. Arch Surg. 2011;146(7):794–801.
25. World Health Organization (WHO). Health care-associated infections Fact Sheet [Internet]. Available from: https://www.who.int/gpsc/country_work/gpsc_ccisc_fact_sheet_en.pdf
26. Murray E, Holmes A. Addressing healthcare-associated infections and antimicrobial resistance from an organizational perspective: Progress and challenges. J Antimicrob Chemother. 2012;67(SUPPL.1):29–36.
27. Harbarth S, Sax H, Gastmeier P. The preventable proportion of nosocomial infections: An overview of published reports. J Hosp Infect. 2003;54(4):258–66.
28. Zarb P, Coignard B, Griskeviciene J, Muller A, Vankerckhoven V, Weist K, et al. The european centre for disease prevention and control (ECDC) pilot point prevalence survey of healthcare-associated infections and antimicrobial use. Vol. 17, Eurosurveillance. 2012. 1–16 p.
29. Barrasa-Villar JI, Aibar-Remón C, Prieto-Andrés P, Mareca-Doñate R, Moliner-Lahoz J. Impact on morbidity, mortality and length of stay of hospital acquired infections by resistant microorganisms. Clin Infect Dis. 2017;65(May):1–27.
30. Al-tawfiq JA, Tambyah PA. Healthcare associated infections (HAI) perspectives. J Infect Public Health. 2014;7(4):339–44.
31. Gray J. Infection control : beyond the horizon. J Hosp Infect. 2018;89(4):237–40.
32. Vincent JL, Rello J, Marshall J, Silva E, Anzueto A, Martin CD, et al. International study of the prevalence and outcomes of infection in intensive care units. JAMA - J Am Med Assoc. 2009;302(21):2323–9.
33. Kalanuria AA, Zai W, Mirski M. Ventilator-associated pneumonia in the ICU. Crit Care. 2014;18(208):1–8.
34. Ramasethu J. Prevention and treatment of neonatal nosocomial infections. Matern Heal Neonatol Perinatol. 2017;3(1):1–11.
35. Gadallah MAH, Fotouh AMA, Habil IS, Imam SS, Wassef G. Surveillance of health care-associated infections in a tertiary hospital neonatal intensive care unit in Egypt: 1-year follow-up. Am J Infect Control. 2014;42(11):1207–11.
36. Brady MT. Health care-associated infections in the neonatal intensive care unit. Am J Infect Control. 2005;33(5):268–75.
37. RA Polin, S Denson MB. Strategies for prevention of health care-associated infections in the NICU. Pediatrics. 2012;129(4):e1085–93.
38. Isaacs D. Evidence-based neonatal infections. 1st ed. Chichester: John Wiley & Sons

Ltd; 2014. 1–19 p.
39. Gastmeier P, Loui A, Stamm-Balderjahn S, Hansen S, Zuschneid I, Sohr D, et al. Outbreaks in neonatal intensive care units-They are not like others. Am J Infect Control. 2007;35(3):172–6.
40. Francis S, Khan H, Kennea N. Infection control in United Kingdom neonatal units: variance in practice and the need for an evidence base. J Infect Prev. 2012;13(5):158–62.
41. Department of Health. Estates and Facilities Divison. HBN 00-01: General Design Guidance for Healthcare Buildings. 2014;(March):77. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/316247/HBN_00-01-2.pdf
42. British Medical Association (BMA). Tackling healthcare associated infections through effective policy action [Internet]. 2009. Available from: https://cdn.ps.emap.com/wp-content/uploads/sites/3/2009/06/hcais_tcm41-186768.pdf
43. National Health Service Scotland. Summary of Incident and Findings of the NHS Greater Glasgow and Clyde : Queen Elizabeth University Hospital / Royal Hospital for Children water contamination incident and recommendations for NHSScotland. 2019; Available from: https://www.hps.scot.nhs.uk/web-resources-container/summary-of-incident-and-findings-of-the-nhs-greater-glasgow-and-clyde-queen-elizabeth-university-hospitalroyal-hospital-for-children-water-contamination-incident-and-recommendations-for-nhsscotland/
44. The Scottish Government. Protecting Scotlands Future. The Government's Programme for Scotland 2019-20 [Internet]. 2019. Available from: https://www.gov.scot/publications/protecting-scotlands-future-governments-programme-scotland-2019-20/
45. Harris AD. How Important Is the Environment in the Emergence of Nosocomial Antimicrobial-Resistant Bacteria? Clin Infect Dis. 2008;46(5):686–8.
46. Villalba J, Navarro FA, Cortés F. Origin, History, and Meanings of the Word Transmission. Microbiol Spectr. 2017;5(6).
47. Damani N. Manual of Infection Prevention and Control. Oxford: Oxford University Press; 2012. p. 1–11.
48. Centers for Disease Control and Prevention (CDC). Principles of Epidemiology in Public Health Practice An Introduction to Applied Epidemiology and Biostatistics. 3rd ed. 2012. 64–68 p.
49. The Scottish Parliament. SPICe Briefing: Healthcare Associated Infections. Scottish Parliam [Internet]. 2009;(November). Available from: http://www.scottish.parliament.uk/ResearchBriefingsAndFactsheets/Factsheets/SB_11-80.pdf
50. National Services Scotland (NSS). Chapter 1 - Standard Infection Control Precautions (SICPs) [Internet]. National Infection Prevention and Control Manual,. 2019. Available from: http://www.nipcm.hps.scot.nhs.uk/chapter-1-standard-infection-control-precautions-sicps/#reference1
51. National Services Scotland (NSS). Chapter 2 - Transmission Based Precautions (TBPs) [Internet]. National Infection Prevention and Control Manual. 2019. Available from: http://www.nipcm.hps.scot.nhs.uk/chapter-2-transmission-based-precautions-tbps/#a1091
52. Debbie Weston. Fundamentals of Infection Prevention and Control Theory and Practice. Second Edi. Wiley Backwell; 2013.
53. Ayliffe GAJ, Fraise AP, Geddes AM, Mitchell K. Control of Hospital Infection. 1st ed. New York: Oxford University Press Inc; 2000. 40–57 p.

54. Magill SS, Hellinger W, Cohen J, Kay R, Bailey C, Boland B, et al. Prevalence of Healthcare-Associated Infections in Acute Care Hospitals in Jacksonville, Florida. Infect Control Hosp Epidemiol. 2012;33(3):283–91.
55. Pitout JDD. Transmission Surveillance for Antimicrobial-Resistant Organisms in the Health System. Microbiol Spectr. 2018;6(5):1–10.
56. Centers for Disease Control and Prevention (CDC). National Healthcare Safety Network Long Term Care Facility Component Tracking Infections in Long-term Care Facilities [Internet]. National Healthcare Safety Network. 2019. Available from: https://www.cdc.gov/nhsn/index.html
57. (NSS) NSS. Appendix 13 – NHSScotland list of alert organisms / conditions. 2017;(April). Available from: http://www.nipcm.hps.scot.nhs.uk/appendices/appendix-13-mandatory-nhsscotland-alert-organismcondition-list/
58. Mayhall c. G. Hospital Epidemiology and Infection Control. 4th ed. Lippincott Williams and Wilkins; 2011.
59. Baxter. ICNET [Internet]. ICNET website. 2019. Available from: http://www.icnetplc.com/
60. Baker MA, Baker MA, Yokoe DS, Stelling J, Kleinman K, Kaganov RE, et al. Automated outbreak detection of hospital-associated pathogens: Value to infection prevention programs. Infect Control Hosp Epidemiol. 2020;1016–21.
61. Smith CM, Allen DJ, Nawaz S, Kozlakidis Z, Nastouli E, Hayward A, et al. An interactive data visualisation application to investigate nosocomial transmission of infections [version 1; peer review: 2 approved]. Wellcome Open Res. 2019;4:1–12.
62. Jombart T, Aanensen DM, Baguelin M, Birrell P, Cauchemez S, Camacho A, et al. OutbreakTools: A new platform for disease outbreak analysis using the R software. Epidemics. 2014;7(August):28–34.
63. Wellcome Trust. Microreact. https://www.sanger.ac.uk/science/tools/microreact.
64. Gladstone RA, Lo SW, Goater R, Yeats C, Taylor B, Hadfield J, et al. Visualizing variation within global pneumococcal sequence clusters (GPSCS) and country population snapshots to contextualize pneumococcal isolates. Microb Genomics. 2020;6(5):1–13.
65. Health Protection Scotland. Some information on Statistical Process Control (SPC) charts that may be useful for clinical teams. Version 2.0 [Internet]. 2012 [cited 2017 Sep 23]. Available from: http://www.hps.scot.nhs.uk/resourcedocument.aspx?id=2266
66. Dworkin M. Outbreak Investigations around the World: Case Studies in Infectious Disease Field Epidemiology. 1st ed. Jones and Bartlett Publishers; 2010. 1–419 p.
67. Woolhouse M. Quantifying Transmission. Microbiol Spectr ASMP. 2017;1–8.
68. Smith-Palmer A, Oates K, Webster D, Taylor S, Scott KJ, Smith G, et al. Outbreak of neisseria meningitidis capsular group w among scouts returning from the world Scout Jamboree, Japan, 2015. Eurosurveillance. 2016;21(45).
69. Parcell BJ, Oravcova K, Pinheiro M, Holden MTG, Phillips G, Turton JF, et al. Pseudomonas aeruginosa intensive care unit outbreak: winnowing of transmissions with molecular and genomic typing. J Hosp Infect. 2018;98(3):282–8.
70. Booth M, Donaldson L, Cui X, Sun J, Cole S, Dailsey S, et al. Confirmed bacillus anthracis infection among persons who inject drugs, Scotland, 2009-2010. Emerg Infect Dis. 2014;20(9):1452–63.
71. Parcell BJ, McIntyr PG, Yirrell DL, Fraser A, Quinn M, Templeton K, et al. Prison and community outbreak of severe respiratory infection due to adenovirus type 14p1 in Tayside, UK. J Public Heal (United Kingdom). 2015;37(1):64–9.
72. National Services Scotland (NSS). Chapter 3- Healthcare Infection Incidents, outbreaks and Data Exceedance. National Infection Prevention and Control Manual

[Internet]. 2017 [cited 2017 Sep 23]. Available from: http://www.nipcm.hps.scot.nhs.uk/chapter-3-healthcare-infection-incidents-outbreaks-and-data-exceedance/
73. Dalziel C. Literature Review: Healthcare infection incidents and outbreaks in Scotland. 2017;(March):1–15.
74. Mcauslane H, Morgan D, Hird C, Lighton L, Mcevoy M, England H, et al. Communicable Disease Outbreak Management: Operational guidance. 2014; Available from: http://www.gov.uk/phe%5Cnwww.gov.uk/phe
75. Health Protection Team. Infectious Disease Incident Plan NHS Grampian Version 7.1. 2017.
76. Health Protection Scotland (HPS). NIPCM Healthcare Infection Incident Assessment Tool (HIIAT) [Internet]. 2017. Available from: http://www.nipcm.scot.nhs.uk/media/1360/2017-03-15-final-appendix-14-hiiat.pdf
77. Mellmann A, Harmsen D, Cummings CA, Zentz EB, Leopold SR, Rico A, et al. Prospective genomic characterization of the german enterohemorrhagic Escherichia coli O104:H4 outbreak by rapid next generation sequencing technology. PLoS One. 2011;6(7).
78. Harris SR, Cartwright EJP, Török ME, Holden MTG, Brown NM, Ogilvy-Stuart AL, et al. Whole-genome sequencing for analysis of an outbreak of meticillin-resistant Staphylococcus aureus: A descriptive study. Lancet Infect Dis [Internet]. 2013;13(2):130–6. Available from: http://dx.doi.org/10.1016/S1473-3099(12)70268-2
79. Grant K, Jenkins C, Arnold C, Green J, Zambon M. Implementing pathogen genomics A case study. Phe [Internet]. 2018;31. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/731057/implementing_pathogen_genomics_a_case_study.pdf
80. Alistair L. Salmonella and Shigella serology update. Scottish Microbiol Ref Lab. 2019;
81. Teams HP, Government S, Scotland FS, Scotland PH, England PH, Wales PH, et al. Health Protection Scotland Briefing Note Salmonella Enteritidis 25-SNP cluster t25.12 outbreak. 2020;1–2.
82. COG UK Consortium [Internet]. COVID-19 Genomics UK (COG-UK) Consortium. 2020. Available from: https://www.sanger.ac.uk/collaboration/covid-19-genomics-uk-cog-uk-consortium/#:~:text=The COG-UK Consortium is an innovative partnership of,Director of the COVID-19 Genomics UK %28COG-UK%29 Consortium.
83. da Silva Filipe A, Shepherd JG, Williams T, Hughes J, Aranday-Cortes E, Asamaphan P, et al. Genomic epidemiology reveals multiple introductions of SARS-CoV-2 from mainland Europe into Scotland. Nat Microbiol [Internet]. 2021;6(1):112–22. Available from: http://dx.doi.org/10.1038/s41564-020-00838-z
84. Mercante JW, Winchell JM. Current and emerging legionella diagnostics for laboratory and outbreak investigations. Clin Microbiol Rev. 2015;28(1):95–133.
85. Engleberg NC, Barrett TJ, Fisher H, Porter B, Hurtado E, Hughes JM. Identification of a carrier by using Vi enzyme-linked immunosorbent assay serology in an outbreak of S. typhoid fever on an Indian reservation. J Clin Microbiol. 1983;18(6):1320–2.
86. Quinn CP, Semenova VA, Elie CM, Romero-Steiner S, Greene C, Li H, et al. Specific, sensitive, and quantitative enzyme-linked immunosorbent assay for human immunoglobulin G antibodies to anthrax toxin protective antigen. Emerg Infect Dis. 2002;8(10):1103–10.
87. Garcia-Nuñez M, Quero S, Catini S, Pedro-Botet ML, Mateu L, Sopena N, et al. Comparative molecular and antibody typing during the investigation of an outbreak of Legionnaires' disease. J Infect Chemother. 2013;19(5):896–901.
88. Jahkola M, Montenegro MA, Helmuth R. Outbreak of Salmonella typhimurium

Infection Traced. Microbiology. 1990;28(12):2597–601.
89. Sails AD, Swaminathan B, Fields PI. Utility of multilocus sequence typing as an epidemiological tool for investigation of outbreaks of gastroenteritis caused by Campylobacter jejuni. J Clin Microbiol. 2003;41(10):4733–9.
90. European Centre for Disease Prevention and Control (ECDC). Multi-country outbreak of Listeria monocytogenes PCR serogroup IVb ST6. 2017;(6 December 2017):Stockholm: ECDC;2017. Available from: https://ecdc.europa.eu/sites/portal/files/documents/RRA-Listeria-monocytogenes-2017_0.pdf
91. Papapetropoulou M, Vantarakis AC. Detection of adenovirus outbreak at a municipal swimming pool by nested PCR amplification. J Infect. 1998;36(1):101–3.
92. Gust DA, Wang SA, Black CM, Brown TM, St. Louis ME, King KA, et al. A Pseudo-outbreak of Chlamydia trachomatis in a State Residential Facility: Implications for Diagnostic Testing . J Infect Dis. 2002;185(6):841–4.
93. Pizzini, CV, Zancopé-Oliveira RM, Reiss E, Hajjeh R, Kaufman L, Peralta JM. Evaluation of a western blot test in an outbreak of acute pulmonary histoplasmosis. Clin Diagn Lab Immunol. 1999;6(1):20–3.
94. Berrazeg M, Diene SM, Drissi M, Kempf M, Richet H, Landraud L, et al. Biotyping of Multidrug-Resistant Klebsiella pneumoniae Clinical Isolates from France and Algeria Using MALDI-TOF MS. PLoS One. 2013;8(4).
95. Lind L, Sjögren E, Melby K, Kaijser B. DNA fingerprinting and serotyping of Campylobacter jejuni isolates from epidemic outbreaks. J Clin Microbiol. 1996;34(4):892–6.
96. Huang SS, Yokoe DS, Stelling J, Placzek H, Kulldorff M, Kleinman K, et al. Automated detection of infectious disease outbreaks in hospitals: A retrospective cohort study. PLoS Med. 2010;7(2):1–10.
97. El-Adhami W, Roberts L, Vickery A, Inglis B, Gibbs A SP. Epidemiological analysis of a methicillin-resistant Staphylococcus aureus outbreak using restriction fragment length polymorphisms of genomic DNA. 1991;137:2713–20.
98. Shi ZY, Liu PYF, Lau YJ, Lin YH, Hu BS. Use of pulsed-field gel electrophoresis to investigate an outbreak of Serratia marcescens. J Clin Microbiol. 1997;35(1):325–7.
99. Chan RCK, Tam JS, Fok TF, French GL. RNA-electrophoresis as a typing method for nosocomial rotavirus infection in a special-care baby unit. J Hosp Infect. 1989;13(4):367–75.
100. Meugnier H, Fernandez MP, Bes M, Bran Y, Bornstein N, Freney J, et al. rRNA gene restriction patterns as an epidemiological marker in nosocomial outbreaks of Staphylococcus aureus infections. Res Microbiol. 1993;144(1):25–33.
101. Durand G, Javerliat F, Bes M, Veyrieras JB, Guigon G, Mugnier N, et al. Routine whole-genome sequencing for outbreak investigations of Staphylococcus aureus in a national reference center. Front Microbiol. 2018;9(MAR):1–12.
102. Pearce ME, Alikhan NF, Dallman TJ, Zhou Z, Grant K, Maiden MCJ. Comparative analysis of core genome MLST and SNP typing within a European Salmonella serovar Enteritidis outbreak. Int J Food Microbiol [Internet]. 2018;274(February):1–11. Available from: https://doi.org/10.1016/j.ijfoodmicro.2018.02.023
103. Waring AL, Halse TA, Csiza CK, Carlyn CJ, Musser KA, Limberger RJ. Development of a genomics-based PCR assay for detection of Mycoplasma pneumoniae in a large outbreak in New York State. J Clin Microbiol. 2001;39(4):1385–90.
104. Zhao Y, Hu K, Zhang J, Guo Y, Fan X, Wang Y, et al. Outbreak of carbapenem-resistant Acinetobacter baumannii carrying the carbapenemase OXA-23 in ICU of the eastern Heilongjiang Province, China. BMC Infect Dis. 2019;19(1):1–7.

105. Martín-Lozano D, Cisneros JM, Becerril B, Cuberos L, Prados T, Ortíz-Leyba C, et al. Comparison of a repetitive extragenic palindromic sequence-based PCR method and clinical and microbiological methods for determining strain sources in cases of nosocomial Acinetobacter baumannii bacteremia. J Clin Microbiol. 2002;40(12):4571–5.
106. Friedman CR, Stoeckle MY, Johnson WD, Riley LW. Double-repetitive-element PCR method for subtyping Mycobacterium tuberculosis clinical isolates. J Clin Microbiol. 1995;33(5):1383–4.
107. Ertugrul N, Rodriguez-Barradas MC, Musher DM, Ryan MAK, Agin CS, Murphy SJ, et al. BOX-Polymerase Chain Reaction-Based DNA Analysis of Nonserotypeable Streptococcus pneumoniae Implicated in Outbreaks of Conjunctivitis . J Infect Dis. 1997;176(5):1401–5.
108. Bik EM, Gouw RD, Mooi FR. DNA fingerprinting of Vibrio cholerae strains with a novel insertion sequence element: A tool to identify epidemic strains. J Clin Microbiol. 1996;34(6):1453–61.
109. Bifani PJ, Mathema B, Liu Z, Moghazeh SL, Shopsin B, Tempalski B, et al. Identification of a W variant outbreak of Mycobacterium tuberculosis via population-based molecular epidemiology. J Am Med Assoc. 1999;282(24):2321–7.
110. Zhang Y, Rajagopalan M, Brown BA, Wallace RJ. Randomly amplified polymorphic DNA PCR for comparison of Mycobacterium abscessus strains from nosocomial outbreaks. J Clin Microbiol. 1997;35(12):3132–9.
111. Whitney CG, Hofmann J, Pruckler JM, Benson RF, Fields BS, Bandyopadhyay U, et al. The role of arbitrarily primed PCR in identifying the source of an outbreak of Legionnaires' disease. J Clin Microbiol. 1997;35(7):1800–4.
112. Lyoda S, Wada A, Weller J, Flood SJA, Schreiber E, Tucker B, et al. Evaluation of AFLP, A High-Resolution DNA Fingerprinting Method, as a Tool for Molecular Subtyping of Enterohemorrhagic Escherichia coli O157:H7 Isolates . Microbiol Immunol. 1999;43(8):803–6.
113. Caierão J, Paiva JACD, Sampaio JLM, da Silva MG, Santos DR de S, Coelho FS, et al. Multilocus enzyme electrophoresis analysis of rapidly-growing mycobacteria: An alternative tool for identification and typing. Int J Infect Dis. 2016;42:11–6.
114. Towner JS, Rollin PE, Bausch DG, Sanchez A, Crary SM, Vincent M, 1 Lee WF, Spiropoulou CF, Ksiazek TG, Lukwiya M, Kaducu F, Downing R NS. Rapid Diagnosis of Ebola Hemorrhagic Fever by Reverse Transcription-PCR in an Outbreak Setting and Assessment of Patient Viral Load as a Predictor of Outcome. 2004;0(8):4330–41.
115. Laksanalamai P, Jackson SA, Mammel MK, Datta AR. High density microarray analysis reveals new insights into genetic footprints of listeria monocytogenes strains involved in listeriosis outbreaks. PLoS One. 2012;7(3).
116. Riley LW. Laboratory Methods in Molecular Epidemiology: Bacterial Infections. Microbiol Spectr. 2018;6(6):1–23.
117. Tang Y, Procop GW, Persing DH. Tang et al 1997 Molecular diagnostics of infectious diseases. Clin Chem. 1997;43:2021–38.
118. Microbiology in Schools Advisory Committee. Suitable and unsuitable micro-organisms [Internet]. 2010. p. 1–6. Available from: http://www.misac.org.uk/PDFs/MISAC_suitable_unsuitable_micro-organisms.pdf
119. Singh A, Goering RV, Simjee S, Foley S ZM. Application of molecular techniques to the study of hospital infection. Clin Microbiol Rev. 2006;19(3):512–30.
120. Howard J, Whitcombe D. Methods in Molecular Biology Diagnostic Bacteriology Protocols volume 46. New Jersy: Humana Press; 1995.

121. Tenover FC, Arbeit RD, Goering R V. How to Select and Interpret Molecular Strain Typing Methods for Epidemiological Studies of Bacterial Infections: A Review for Healthcare Epidemiologists. Infect Control Hosp Epidemiol. 1997;18(6):426–39.
122. Fratamico PM, DebRoy C, Liu Y, Needleman DS, Baranzoni GM, Feng P. Advances in molecular serotyping and subtyping of Escherichia coli. Front Microbiol. 2016;7(MAY):1–8.
123. Ibrahim GM, Morin PM. Salmonella serotyping using whole genome sequencing. Front Microbiol. 2018;9(DEC):1–8.
124. van Belkum A, Tassios PT, Dijkshoorn L, Haeggman S, Cookson B, Fry NK, et al. Guidelines for the validation and application of typing methods for use in bacterial epidemiology. Clin Microbiol Infect. 2007;13(SUPPL. 3):1–46.
125. Farber JM. An introduction to the hows and whys of molecular typing? J Food Prot. 1996;59(10):1091–101.
126. Foxman B, Riley L. Molecular epidemiology: Focus on infection. Am J Epidemiol. 2001;153(12):1135–41.
127. Elwell LP, Inamine JM, Minshew BH. Common plasmid specifying tobramycin resistance found in two enteric bacteria isolated from burn patients. Antimicrob Agents Chemother. 1978;13(2):312–7.
128. Woodford N, Johnson A. Methods in Molecular Medicine Molecular Bacteriology Protocols and Clinical Applications. Humana Press; 1998. 1–659 p.
129. Quainoo S, Coolen JPM, Sacha A. F. T. van Hijum, D, Martijn A. Huynen, Willem J. G. Melchers B, Willem van Schaik E, Wertheimb HFL. crossm Whole-Genome Sequencing of Bacterial Pathogens : the Future of Nosocomial. 2017;30(4):1015–64.
130. Huber CA, Foster NF, Riley T V., Paterson DL. Challenges for standardization of Clostridium difficile typing methods. J Clin Microbiol. 2013;51(9):2810–4.
131. Wilcox MH, Fawley WN. Molecular diagnostic techniques. Surgery. 2003;37(10):1–6.
132. Sabat AJ, Budimir A, Nashev D, Sá-Leão R, van Dijl JM, Laurent F, et al. Overview of molecular typing methods for outbreak detection and epidemiological surveillance. Eurosurveillance. 2013;18(4):1–15.
133. Andrei A, Zervo MJ. The Application of Molecular Techniques to the Study of Hospital Infection. Arch Pathol Lab Med. 2006;130(May):662–8.
134. Tenover FC, Arbeit RD, Goering R V., Mickelsen PA, Murray BE, Persing DH, et al. Interpreting chromosomal DNA restriction patterns produced by pulsed- field gel electrophoresis: Criteria for bacterial strain typing. J Clin Microbiol. 1995;33(9):2233–9.
135. Vos P, Hogers R, Bleeker M, Reijans M, Lee T Van De, Frijters A, et al. Vos 1995 AFLP.PDF. 1995;23(21):4407–14.
136. Lee OR, Kim M, Yang D. Amplified Fragment Length Polymorphism (AFLP) - an invaluable fingerprinting technique for genomic, transcriptomic and epigenetic studies. 2012;862:75–87.
137. Moura A, Criscuolo A, Pouseele H, Maury MM, Leclercq A, Tarr C, et al. Whole genome-based population biology and epidemiological surveillance of Listeria monocytogenes. Nat Microbiol. 2016;2(2):1–10.
138. Woodford N, Alan P. Johnson. Genomics, Proteomics and Clinical Bacteriology. New Jersey; 2004.
139. Koreen L, Graviss S V., Ramaswamy EA, Naidich S, Musser JM, Kreiswirth BN. Typing Method for Discriminating among. J Clin Microbiol. 2004;42(2):792–9.
140. (CDC) C for DC and P. M Protein Gene (emm) Typing [Internet]. http://www.cdc.gov/streplab/MProteinGene-typing.html. 2018. Available from: https://www.cdc.gov/streplab/groupa-strep/emm-background.html

141. Beall B, Facklam R, Thompson T. Sequencing emm -specific polymerase chain reaction products for routine and accurate typing of group A Streptococci. JClinMicrobiol [Internet]. 1995;34(4):953–8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC228924/pdf/340953.pdf
142. Pulcrano G, Roscetto E, Iula VD, Panellis D, Rossano F, Catania MR. MALDI-TOF mass spectrometry and microsatellite markers to evaluate Candida parapsilosis transmission in neonatal intensive care units. Eur J Clin Microbiol Infect Dis. 2012;31(11):2919–28.
143. Von Graevenitz A. Advances in Applied Microbiology. Vol. 81, The Yale journal of biology and medicine. 2012. p. 169–207.
144. Hennequin C, Thierry A, Richard GF, Lecointre G, Nguyen H V, Gaillardin C, et al. Microsatellite Typing as a New Tool for Identification of Saccharomyces cerevisiae Strains. Society. 2001;39(2):551–9.
145. Singhal N, Kumar M, Kanaujia PK, Virdi JS. MALDI-TOF mass spectrometry: An emerging technology for microbial identification and diagnosis. Front Microbiol. 2015;6(AUG):1–16.
146. Schlebusch S, Price GR, Gallagher RL, Horton-Szar V, Elbourne LDH, Griffin P, et al. MALDI-TOF MS meets WGS in a VRE outbreak investigation. Eur J Clin Microbiol Infect Dis. 2017;36(3):495–9.
147. Bittar F, Cassagne C, Bosdure E, Stremler N, Dubus JC, Sarles J, et al. Outbreak of corynebacterium pseudodiphtheriticum infection in cystic fibrosis patients, France. Emerg Infect Dis. 2010;16(8):1231–6.
148. Williamson YM, Moura H, Woolfitt AR, Pirkle JL, Barr JR, Carvalho MDG, et al. Differentiation of Streptococcus pneumoniae conjunctivitis outbreak isolates by matrix-assisted laser desorption ionization-time of flight mass spectrometry. Appl Environ Microbiol. 2008;74(19):5891–7.
149. Schlebusch S, Price GR, Hinds S, Nourse C, Schooneveldt JM, Tilse MH, et al. First outbreak of PVL-positive nonmultiresistant MRSA in a neonatal ICU in Australia: Comparison of MALDI-TOF and SNP-plus-binary gene typing. Eur J Clin Microbiol Infect Dis. 2010;29(10):1311–4.
150. Bruker. IR Biotyper [Internet]. 2020. Available from: https://www.bruker.com/fileadmin/user_upload/8-PDF-Docs/Separations_MassSpectrometry/Literature/Brochures/1852688_IR_Biotyoer_05-2017_ebook.pdf
151. Fleischmann RD, Adams MD, White O, Clayton RA, Ewen F, Kerlavage AR, et al. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science. 1995;496(17):1–16.
152. Goldberg B, Sichtig H, Geyer C, Ledeboer N, Weinstock GM. Making the Leap from Research Laboratory to Clinic : Challenges and Opportunities for Next-Generation Sequencing in Infectious Disease Diagnostics. 2015;6(6):1–10.
153. Kwong JC, Mccallum N, Sintchenko V, Howden BP. Whole genome sequencing in clinical and public health microbiology. Pathology. 2015;47(3):199–210.
154. Didelot X, Bowden R, Wilson DJ, Peto TEA, Crook DW. Transforming clinical microbiology with bacterial genome sequencing. Nat Rev Genet. 2012;13(9):601–12.
155. Land M, Hauser L, Jun SR, Nookaew I, Leuze MR, Ahn TH, et al. Insights from 20 years of bacterial genome sequencing. Funct Integr Genomics. 2015;15(2):141–61.
156. GM G. A New Genomics-Driven Taxonomy of Bacteria and Archaea: Are We There Yet? J Clin Microbiol. 2016;54:1956–63.
157. Ranjan R, Rani A, Metwally A, McGee HS, Perkins DL. Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. Biochem

Biophys Res Commun [Internet]. 2016;469(4):967–77. Available from: http://dx.doi.org/10.1016/j.bbrc.2015.12.083
158. Google. Google trends [Internet]. 2019 [cited 2019 Dec 15]. Available from: https://trends.google.com/trends/
159. Harris SR, Feil EJ, Holden MTG, Quail MA, Nickerson EK, Chantratita N, Gardete S, Tavares A, Day N, Lindsay JA, Edgeworth JD, de Lencastre H, Parkhill J, Peacock SJ, Bentley SD. Evolution of MRSA During Hospital Transmission and Intercontinental Spread. Science (80-). 2010;327(469–474):1689–99.
160. Rohde H, Qin J, Cui Y, Li D, Loman N, Hentschke M, Chen W, Pu F, Peng Y, Li J, Xi F, Li S, Li Y, Zhang Z, Yang X, Zhao M, Wang P, Guan Y, Cen Z, Zhao X, Christner M, Kobbe R, Danchin A, Gao G SY, Li Y, Yang H, Wang J. Open-Source Genomic Analysis of Shiga-Toxin–Producing E.coli O104:H4. 2011;
161. Köser CU, Holden, Matthew T G, Ellington M, Sanders M, Enright MC, Dougan G, Bentley SD, et al. Rapid Whole-Genome Sequencing for Investigation of a Neonatal MRSA Outbreak. N Engl J Med. 2013;366(24):7–20.
162. Eyre DW, Cule ML, Wilson DJ, Vaughan A, Connor LO, Ip CLC, et al. Diverse Sources of C . difficile Infection Identified on Whole- Genome Sequencing. 2014;369(13).
163. Toleman MS, Reuter S, Coll F, Harrison EM, Peacock SJ. Local persistence of novel MRSA lineage after hospital ward outbreak, Cambridge, UK, 2011–2013. Emerg Infect Dis. 2016;22(9):1658–9.
164. Eppinger M, Pearson T, Koenig SSK, Pearson O, Hicks N, Agrawal S, et al. Genomic Epidemiology of the Haitian Cholera Outbreak: a Single Introduction Followed by Rapid, Extensive, and Continued Spread Characterized the Onset of the Epidemic. 2014;5(6):1–8.
165. Coll F, Harrison EM, Toleman MS, Reuter S, Raven KE, Blane B, et al. Longitudinal genomic surveillance of MRSA in the UK reveals transmission patterns in hospitals and the community. Sci Transl Med. 2017;9(413):1–19.
166. Toleman MS, Watkins ER, Williams T, Blane B, Sadler B, Harrison EM, et al. Investigation of a Cluster of Sequence Type 22 Methicillin-Resistant Staphylococcus aureus Transmission in a Community Setting. Clin Infect Dis. 2017;65(12):2069–77.
167. J. D. Watson and F. H. C. Crick. A structure for deoxyribose nucleic acid. J Am Coll Cardiol. 1953;42(2):373–4.
168. Heather JM, Chain B. The sequence of sequencers: The history of sequencing DNA. Genomics. 2016;107(1):1–8.
169. Holley R.W et al. Structure of a ribonucleic acid. Science (80-). 1965;147:1462–5.
170. Sanger F, Air GM, Barrell BG, Brown NL, Coulson a R, Fiddes C a, et al. Nucleotide sequence of bacteriophage phi X174 DNA. Nature. 1977;265(5596):687–95.
171. F Sanger AC. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol. 1975;94.
172. Maxam AM, Gilbert W. A new method for sequencing DNA. 1977. Biotechnology. 1992;24(2):99–103.
173. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci. 1977;74(12):5463–7.
174. Pål Nyrén AL. Enzymatic method for continuous monitoring of inorganic pyrophosphate synthesis. Anal Biochem. 1985;151:504–9.
175. Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, et al. The complete genome of an individual by massively parallel DNA sequencing. Nature. 2008;452(7189):872–6.
176. Balasubramanian S. Solexa sequencing: Decoding genomes on a population scale. Clin

Chem. 2015;61(1):21–4.
177. Turcatti G, Romieu A, Fedurco M, Tairi AP. A new class of cleavable fluorescent nucleotides: Synthesis and optimization as reversible terminators for DNA sequencing by synthesis. Nucleic Acids Res. 2008;36(4).
178. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, et al. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genomics. 2012;13(1):1.
179. Feng Y, Zhang Y, Ying C, Wang D, Du C. Nanopore-based fourth-generation DNA sequencing technology. Genomics, Proteomics Bioinforma [Internet]. 2015;13(1):4–16. Available from: http://dx.doi.org/10.1016/j.gpb.2015.01.009
180. Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, Bore JA, Koundouno R, Dudas G, Mikhail A, Ouédraogo N, Afrough B, Bah A, Baum JH, Becker-Ziaja B, Boettcher JP, Cabeza-Cabrerizo M, Camino-Sánchez Á, Carter LL, Doerrbecker J, Enkirch T, Garc CM. Real-time, portable genome sequencing for Ebola surveillance. Nature. 2016;530(7589):228–32.
181. Deurenberg RH, Bathoorn E, Chlebowicz MA, Couto N, Ferdous M, García-Cobos S, et al. Application of next generation sequencing in clinical microbiology and infection prevention. J Biotechnol [Internet]. 2017;243:16–24. Available from: http://dx.doi.org/10.1016/j.jbiotec.2016.12.022
182. MacCannell D. Platforms and Analytical Tools Used in Nucleic Acid Sequence-Based Microbial Genotyping Procedures. Microbiol Spectr. 2019;7(1):1–17.
183. Dark MJ. Whole-genome sequencing in bacteriology: State of the art. Infect Drug Resist. 2013;6:115–22.
184. Madigan T, Cunningham SA, Patel R, Greenwood-Quaintance KE, Barth JE, Sampathkumar P, et al. Whole-genome sequencing for methicillin-resistant Staphylococcus aureus (MRSA) outbreak investigation in a neonatal intensive care unit. Infect Control Hosp Epidemiol. 2018;39(12):1412–8.
185. Wilson DJ. Insights from Genomics into Bacterial Pathogen Populations. PLoS Pathog. 2012;8(9).
186. Luheshi L, Raza S, Moorthie S, Hall A, Blackburn L, Rands C, et al. Pathogen genomics into practice. PHG Foundation, Cambridge, United Kingdom. 2015.
187. Hendrik C. Introduction to Bioinformatics in Microbiology. First. Switzerland: Springer; 2018. 1–213 p.
188. Gauthier J, Vincent AT, Charette SJ, Derome N. A brief history of bioinformatics. Brief Bioinform. 2019;20(6):1981–96.
189. Choudhuri Supratim. Bioinformatics for Beginners Genes, Genomes, Molecular Evolution, Databases and Analytical Tools. First. London: Elsevier Inc; 2014. 1–225 p.
190. Koonin E V. Darwinian evolution in the light of genomics. Nucleic Acids Res. 2009;37(4):1011–34.
191. Gregory TR. Understanding Evolutionary Trees. Evol Educ Outreach. 2008;1(2):121–37.
192. Wheeler Q, Assis L, Rieppel O. Phylogenetics: Heed the father of cladistics. Nature. 2013;496(7445):295–6.
193. Woese CR. Bacterial evolution. Microbiol Rev. 1987;51(2):221–71.
194. Ahrenfeldt J, Skaarup C, Hasman H, Pedersen AG, Aarestrup FM, Lund O. Bacterial whole genome-based phylogeny: Construction of a new benchmarking dataset and assessment of some existing methods. BMC Genomics. 2017;18(1).
195. Centers for Disease Control and Prevention C. Pulsed-field Gel Electrophoresis (PFGE). https://www.cdc.gov/pulsenet/pathogens/pfge.html.
196. Ninewells D of MM. NB027v12 Sensitivity testing.

197. Turton JF, Turton SE, Yearwood L, Yarde S, Kaufmann ME, Pitt TL. Evaluation of a nine-locus variable-number tandem-repeat scheme for typing of Pseudomonas aeruginosa. Clin Microbiol Infect. 2010;16(8):1111–6.
198. Woodford N, Ward ME, Kaufmann ME, Turton J, Fagan EJ, James D, et al. Community and hospital spread of Escherichia coli producing CTX-M extended-spectrum β-lactamases in the UK. J Antimicrob Chemother. 2004;54(4):735–43.
199. Lindsay DSJ, Brown AW, Scott KJ, Denham B, Thom L, Rundell G, et al. Circulating emm types of Streptococcus pyogenes in Scotland: 2011-2015. J Med Microbiol. 2016;65(10):1229–31.
200. DNA EXTRACTION. QIAGEN DNA Blood Mini Handbook [Internet]. Qiagen. 2016. Available from: http://www.qiagen.com/knowledge-and-support/resource-center/resource-download.aspx?id=67893a91-946f-49b5-8033-394fa5d752ea&lang=en
201. Lucigen E. MasterPure™ Gram Positive DNA Purification Kit Cat. No. MGP04100. Madison, WI, USA; 2018.
202. Pettigrew K. NEXTERA XT MiSeq protocol - 2017-05-16. 2019.
203. US B. NanoVue Plus ™ Product User Manual.
204. Illumina. MiSeq System Guide Document # 15027617 v05 M. 2019.
205. Illumina. Illumina Experiment Manager Software Guide Document # 15031335 v08. 2018;(December):1–18.
206. Illumina. MiSeq™ system Denature and Dilute Libraries Guide Document # 15039740 v10. 2019;1–14.
207. Inouye M, Dashnow H, Raven L, Schultz MB, Pope BJ, Tomita T, et al. SRST2 : Rapid genomic surveillance for public health and hospital microbiology labs. 2014;1–16.
208. Page AJ, Parkhill J, Keane JA, Hunt M, Mather AE, Leonor S, et al. ARIBA : rapid antimicrobial resistance genotyping directly from sequencing reads. 2017;
209. Nelson KE, Fouts DE, Mongodin EF, Ravel J, DeBoy RT, Kolonay JF, et al. Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen Listeria monocytogenes reveal new insights into the core genome components of this species. Nucleic Acids Res. 2004;32(8):2386–95.
210. Dickson E, Phillips G, Lee J, Imrie L, Alexander C, Coia J. First description of t10939 BORSA isolates in Scotland. 2016;18(5):10939.
211. Nadasy KA, Domiati-Saad R, Tribble MA. Invasive Klebsiella pneumoniae Syndrome in North America. Clin Infect Dis. 2007;45(3):e25–8.
212. Holt KE, Wertheim H, Zadoks RN, Baker S, Whitehouse CA, Dance D, et al. Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae* , an urgent threat to public health. Proc Natl Acad Sci. 2015;112(27):E3574–81.
213. Prokesch BC, TeKippe M, Kim J, Raj P, TeKippe EME, Greenberg DE. Primary osteomyelitis caused by hypervirulent Klebsiella pneumoniae. Lancet Infect Dis. 2016;16(9):e190–5.
214. Forde BM, Ben Zakour NL, Stanton-Cook M, Phan MD, Totsika M, Peters KM, et al. The complete genome sequence of escherichia coli EC958: A high quality reference sequence for the globally disseminated multidrug resistant E. coli O25b:H4-ST131 clone. PLoS One. 2014;9(8).
215. (BAPM) BA of PM. BAPM Guidance on Cot Capacity and the use of Nurse Staffing standards This document has been written by BAPM EC and will be further developed in consultation with the BAPM membership and with neonatal networks. 2018;(June):1–3.

216. Kwong JC, Lane CR, Romanes F, Gonçalves da Silva A, Easton M, Cronin K, et al. Translating genomics into practice for real-time surveillance and response to carbapenemase-producing Enterobacteriaceae: evidence from a complex multi-institutional KPC outbreak. PeerJ. 2018;6:e4210.
217. Wellcome Sanger Institute. pathogenwatch [Internet]. Pathogenwatch. Available from: https://www.sanger.ac.uk/tool/pathogenwatch/
218. Gargis A, Kalman L, Lubin IM. Assuring the quality of next-generation sequencing in clinical laboratory practice. J Clin Microbiol. 2016;54(12):2857–65.
219. Kozyreva VK, Truong CH, Greninger A, Crandall J, Mukhopadhyay R C V. Validation and Implementation of Clinical Laboratory Improvements Act- Compliant Whole-Genome Sequencing in the Public Health Microbiology Laboratory. J Clin Microbiol. 2017;55(8):2502–20.
220. De Almeida OGG, De Martinis ECP. Relating next-generation sequencing and bioinformatics concepts to routine microbiological testing. Electron J Gen Med. 2019;16(3):1–15.
221. Caboche S, Even G, Loywick A, Audebert C, Hot D. MICRA: An automatic pipeline for fast characterization of microbial genomes from high-throughput sequencing data. Genome Biol. 2017;18(1):1–14.
222. Quijada NM, Rodríguez-Lázaro D, Eiros JM, Hernández M, Valencia A. TORMES: An automated pipeline for whole bacterial genome analysis. Bioinformatics. 2019;35(21):4207–12.
223. Hong C, Manimaran S, Shen Y, Perez-Rogers JF, Byrd AL, Castro-Nallar E, et al. PathoScope 2.0: A complete computational framework for strain identification in environmental or clinical sequencing samples. Microbiome. 2014;2(1).
224. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, et al. VFDB: A reference database for bacterial virulence factors. Nucleic Acids Res. 2005;33(DATABASE ISS.):325–8.
225. Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, et al. ARG-annot, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Chemother. 2014;58(1):212–20.
226. U.S Food & Drug Administration (FDA). Database for Reference Grade Microbial Sequences (FDA-ARGOS) [Internet]. Database for Reference Grade Microbial Sequences (FDA-ARGOS). 2018 [cited 2018 Jul 16]. Available from: https://www.fda.gov/medicaldevices/scienceandresearch/databaseforreferencegrademicrobialsequences/default.htm
227. Kumar S, Shankar B, Arya S, Deb M, Chellani H. Healthcare associated infections in neonatal intensive care unit and its correlation with environmental surveillance. J Infect Public Health [Internet]. 2018;11(2):275–9. Available from: http://dx.doi.org/10.1016/j.jiph.2017.08.005
228. Dworkin MS. Cases in Field Epidemiology: A Global Perspective. 2011;477.
229. Wellcome Genome Campus Course: Genomics and Clinical Microbiology [Internet]. Available from: https://coursesandconferences.wellcomegenomecampus.org/our-events/genomics-clinical-microbiology-2021/
230. Futurelearn Introduction to Bacterial Genomics [Internet]. 2020. Available from: https://www.futurelearn.com/courses/introduction-to-bacterial-genomics
231. Fox EM, Delappe N, Garvey P, McKeown P, Cormican M, Leonard N, et al. PFGE analysis of Listeria monocytogenes isolates of clinical, animal, food and environmental origin from Ireland. J Med Microbiol. 2012;61(4):540–7.
232. O'Reilly LC. A method for overcoming DNA degradation during PFGE for Serratia marcescens. J Microbiol Methods [Internet]. 2011;85(2):173–4. Available from: http://dx.doi.org/10.1016/j.mimet.2011.02.014

233. Römling U, Tümmler B. Achieving 100% typeability of Pseudomonas aeruginosa by pulsed-field gel electrophoresis. J Clin Microbiol. 2000;38(1):464–5.
234. Marshall S, Clark CG, Wang G, Mulvey M, Kelly MT, Johnson WM. Comparison of molecular methods for typing Vibrio parahaemolyticus. J Clin Microbiol. 1999;37(8):2473–8.
235. Corkill JE, Graham R, Hart CA, Stubbs S. Pulsed-field gel electrophoresis of degradation-sensitive DNAs from Clostridium difficile PCR ribotype 1 strains. J Clin Microbiol. 2000;38(7):2791–2.
236. Moissenet D, Goujon CP, Garbarg-Chenon A, Vu-Thien H. CDC group IV c-2: A new Ralstonia species close to Ralstonia eutropna. J Clin Microbiol. 1999;37(6):1777–81.
237. Kinscherf TG, Yap MN, Charkowski AO, Willis DK. Chef procedures: A rapid high-temperature method for sample preparation, a high voltage hepes buffer system and the use of nusieve® agarose. J Rapid Methods Autom Microbiol. 2009;17(1):9–16.
238. Ishii Y, Alba J, Kimura S, Nakashima K, Abe Y YK. Rapid pulsed-field gel electrophoresis technique for determination of genetic diversity of Serratia marcescens. J Infect Chemother. 2002;8:368–370.
239. Cheah ALY, Spelman T, Liew D, Peel T, Howden BP, Spelman D, et al. Enterococcal bacteraemia: Factors influencing mortality, length of stay and costs of hospitalization. Clin Microbiol Infect. 2013;19(4):E181–9.
240. Neely AN, Maley MP. Survival of Enterococci and Sthaphylococci on Hospital Fabrics and Plastics. J Clin Microbiol. 2000;38(2):724–6.
241. Raven KE, Gouliouris T, Brodrick H, Coll F, Brown NM, Reynolds R, et al. Complex Routes of Nosocomial Vancomycin-Resistant Enterococcus faecium Transmission Revealed by Genome Sequencing. Clin Infect Dis. 2018;64(July):886–93.
242. Howden BP, Holt KE, Lam MM, Seemann T, Ballard S, Coombs GW, Tong SY, Grayson ML, Johnson PD, Stinear TP. Genomic Insights to Control the Emergence of Vancomycin-Resistant. MBio. 2013;4(4):1–9.
243. Mellmann A, Bletz S, Böking T, Kipp F, Becker K. Real-Time Genome Sequencing of Resistant Bacteria Provides Precision Infection Control in an Institutional Setting. Am Soc Microbiol. 2016;54(August):2874–81.
244. Otter JA, Burgess P, Davies F, Mookerjee S, Singleton J, Gilchrist M, et al. Counting the cost of an outbreak of carbapenemase-producing Enterobacteriaceae: an economic evaluation from a hospital perspective. Clin Microbiol Infect. 2017;23(3):188–96.
245. Jiang Y, Wei Z, Wang Y, Hua X, Feng Y, Yu Y. Tracking a hospital outbreak of KPC-producing ST11 Klebsiella pneumoniae with whole genome sequencing. Clin Microbiol Infect. 2015;21(11):1001–7.
246. Snitkin ES, Zelazny AM, Thomas PJ, Stock, Frida DK, Palmore TN, Segre JA. Tracking a hospital outbreak of carbapenem-resistant Klebsiella pneumoniae with whole-genome sequencing. Sci Transl Med. 2012;4(148):1–18.
247. Rimoldi SG, Gentile B, Pagani C, Di Gregorio A, Anselmo A, Palozzi AM, et al. Whole genome sequencing for the molecular characterization of carbapenem-resistant Klebsiella pneumoniae strains isolated at the Italian ASST Fatebenefratelli Sacco Hospital, 2012-2014. BMC Infect Dis. 2017;17(1):1–11.
248. Onori R, Gaiarsa S, Comandatore F, Pongolini S, Brisse S, Colombo A, et al. Tracking nosocomial Klebsiella pneumoniae infections and outbreaks by whole-genome analysis: Small-scale Italian scenario within a single hospital. J Clin Microbiol. 2015;53(9):2861–8.
249. Ellington MJ, Ekelund O, Aarestrup FM, Canton R, Doumith M, Giske C, et al. The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria : report from the EUCAST Subcommittee. 2017;23:2–22.

250. Carattoli A, Zankari E, Garciá-Fernández A, Larsen MV, Lund O, Villa L, et al. In Silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother. 2014;58(7):3895–903.
251. Choi Y, Chan AP. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics. 2015;31(16):2745–7.
252. Martin J, Phan HTT, Findlay J, Stoesser N, Pankhurst L, Navickaite I, et al. Covert dissemination of carbapenemase-producing Klebsiella pneumoniae (KPC) in a successfully controlled outbreak: Long- and short-read whole-genome sequencing demonstrate multiple genetic modes of transmission. J Antimicrob Chemother. 2017;72(11):3025–34.
253. Munoz-Price LS, Poirel L, Bonomo RA, Schwaber MJ, Daikos GL, Cormican M, et al. Clinical epidemiology of the global expansion of Klebsiella pneumoniae carbapenemases. Lancet Infect Dis. 2013;13(9):785–96.
254. Doi Y, Paterson DL. Carbapenemase-Producing Enterobacteriaceae. Semin Respir Crit Care Med. 2016;36(1):74–84.
255. Van Duin D, Perez F, Rudin SD, Cober E, Hanrahan J, Ziegler J, et al. Surveillance of carbapenem-resistant Klebsiella pneumoniae: Tracking molecular epidemiology and outcomes through a regional network. Antimicrob Agents Chemother. 2014;58(7):4035–41.
256. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. 2018;(July 2012):2640–4.
257. Stoesser N, Batty EM, Eyre DW, Morgan M, Wyllie DH, Elias CDO, et al. Predicting antimicrobial susceptibilities for Escherichia coli and Klebsiella pneumoniae isolates using whole genomic sequence data. 2018;(June):2234–44.
258. Wang Y, Lv Y, Cai J, Schwarz S, Cui L, Hu Z, et al. A novel gene, optrA, that confers transferable resistance to oxazolidinones and phenicols and its presence in Enterococcus faecalis and Enterococcus faecium of human and animal origin. J Antimicrob Chemother. 2015;70(8):2182–90.
259. Woodford N. AMRHAI News. Public Heal Engl. 12000BC;
260. Agency HP. Potentially transferable linezolid resistance in Enterococcus faecium in the UK. 2012;(July):2012.
261. Health Protection Scotland. Oxazolidinone-resistance due to optrA in Enterococcus faecalis. HPS Wkly Rep. 2016;50(2016/29):230–1.
262. Gawryszewska I, Żabicka D, Hryniewicz W, Sadowy E. Linezolid-resistant enterococci in Polish hospitals: species, clonality and determinants of linezolid resistance. Eur J Clin Microbiol Infect Dis. 2017;36(7):1279–86.
263. Biek R, Pybus OG, Lloyd-Smith JO, Didelot X. Measurably evolving pathogens in the genomic era. Trends Ecol Evol. 2015;30(6):306–13.
264. Golubchik T, Batty EM, Miller RR, Farr H, Young BC, Larner-Svensson H, et al. Within-Host Evolution of Staphylococcus aureus during Asymptomatic Carriage. PLoS One. 2013;8(5):1–14.
265. Chen Y, Burall LS, Luo Y, Timme R, Melka D, Muruvanda T, et al. Listeria monocytogenes in Stone Fruits Linked to a Multistate Outbreak : Enumeration of Cells and Whole-Genome Sequencing. 2016;82(24):7030–40.
266. Moura A, Tourdjman M, Leclercq A, Hamelin E, Laurent E, Fredriksen N, et al. Real-time whole-genome sequencing for surveillance of Listeria monocytogenes, France. Emerg Infect Dis. 2017;23(9):1462–70.
267. Darch SE, McNally A, Harrison F, Corander J, Barr HL, Paszkiewicz K, et al. Recombination is a key driver of genomic and phenotypic diversity in a Pseudomonas aeruginosa population during cystic fibrosis infection. Sci Rep. 2015;5:1–12.

268. Zhou K, Lokate M, Deurenberg RH, Tepper M, Arends JP, Raangs EGC, et al. Use of whole-genome sequencing to trace, control and characterize the regional expansion of extended-spectrum β-lactamase producing ST15 Klebsiella pneumoniae. Sci Rep [Internet]. 2016;6(January):1–10. Available from: http://dx.doi.org/10.1038/srep20840
269. Andrew S. Scottish Microbiology Reference Laboratories. Heal Prot Scotl [Internet]. 2018; Available from: http://kiss.kstudy.com/search/detail_page.asp?key=3424155%0Ahttps://doi.org/10.1016/j.rcim.2018.05.010%0Ahttps://doi.org/10.1016/j.cie.2018.07.041
270. Scheckler WE, Buck AS, Farr BM, Friedman C, Gross PA, Harris J. SPECIAL COMMUNICATIONRequirements for infrastructure and essential activities of infection control and epidemiology in hospitals: A Consensus Panel report. 2013;1–14. Available from: papers2://publication/uuid/D763E50B-F17E-4856-B7ED-488D41900E3D
271. Peacock SJ, Parkhill J, Brown NM. Changing the paradigm for hospital outbreak detection by leading with genomic surveillance of nosocomial pathogens. Microbiol (United Kingdom). 2018;164(10):1213–9.
272. Blane B, Raven KE, Leek D, Brown N, Parkhill J, Peacock SJ. Rapid sequencing of MRSA direct from clinical plates in a routine microbiology laboratory. J Antimicrob Chemother. 2019;74(8):2153–6.
273. COG-UK. COVID-19 Genomics UK (COG-UK) Consortium Report. Vol. Report #10. 2020.
274. O'Toole A, Hill V, Rambaut A. CIVET Cluster Investigation & Virus Epidemiology Tool [Internet]. 2020 [cited 2021 Feb 14]. Available from: https://artic-network.github.io/civet/
275. O'Toole A, Hill V, McCrone J, Scher E, Rambaut A. Pangolin COVID-19 Lineage Assigner Phylogenetic Assignment of Named Global Outbreak LINeages [Internet]. v2.2.2, lineages version 2021-02-12. 2021 [cited 2021 Feb 14]. Available from: https://pangolin.cog-uk.io

8 Appendices

8.1 Appendix 1 Caldicott approval NHS Grampian

NHS
Grampian

**APPLICATION FORM FOR
CALDICOTT APPROVAL
FOR USE OF PATIENT IDENTIFIABLE DATA**

After completion please return this form to

Caldicott, Information Governance, NHS Grampian, Rosehill House, Foresterhill Site, Cornhill Road, Aberdeen AB25 2ZG

Email: nhsg.caldicott@nhs.net

Project Title
Whole Genome Sequencing Service for the Investigation of Healthcare Associated Infection

Description:

Whole Genome Sequencing (WGS) can be used so that outbreaks can be detected in real time, enabling rapid implementation of targeted infection prevention and control measures. It may also rule out outbreaks and as a result outbreak meetings/closing of wards may not be necessary. Project aims to determine how to establish a WGS service for the investigation of HAI outbreaks in real time, examine the health economics (cost effectiveness and impact on services), provide further evidence that WGS is a superior typing method able to unequivocally define strains compared to current typing techniques for different HAI organisms and add to published literature.

Name of Applicant: **Dr Ben Parcell**
Consultant in Medical Microbiology
Community and Dr Gray's Infection Prevention and Control Doctor

Address: **Medical Microbiology**
Aberdeen Royal Infirmary
Foresterhill
Aberdeen AB25 2ZN

Tel No

Email address:

Application Number .....(for office use only) 1

NHS
Grampian

Name of organisation receiving data: NHS Grampian

and their Data Protection Registration Number: _____

What patient identifiable information are you looking to use?

CHI Number	NO
Forename	NO
Surname	NO
Initials	NO
Date of Birth	NO
Address	NO
Postcode	NO
Other, please specify	Diagnosis
Age	Yes
Gender	Yes

Application Number .....(for office use only)

NHS
Grampian

How will the data be transferred?
Paper records □
Computer records □ X

(Note – patient/user identifiable data must not be transferred via e-mail unless anonymised, encrypted or using secure NHS network i.e. nhs.net)

Who else will have access to the data?
(If data recipients are not employed by the NHS please state whether NHS honorary contracts are in place. If not – detail confidentiality agreements)

Only data such as organism result (identification of bacteria, antibiotic susceptibility) will be shared with Professor Stephen Henry Gillespie and four of his team members at School of Medicine, University of St Andrews. He is also employed by NHS Fife as a Consultant Microbiologist.

How will the service users be contacted?

NA

How will service users consent be obtained?

NA

If no consent being obtained, please detail the reason why not?

We will only use residual bacterial isolates which are surplus to diagnostic requirements and would otherwise be discarded as clinical waste. No extra samples taken.

Bacterial isolates identified as part of routine care retrospectively and prospectively collected. All patient details de-identified. Patients will remain anonymous.

Where will the data be stored?

The information will be stored on an password locked NHS computer in the NHS Microbiology department in a room with a locked door.

How will the data be protected? (Please detail security measures to be taken)

The NHS computer is password protected.

If the data is on a computer is there access via a network?

It will be stored in Dr Ben Parcells own U drive which cannot be accessed by anyone else.

NHS
Grampian

How long will the data be stored? 5 years
At the end of this period, how will the data be disposed of? It will be deleted
Who will be responsible for ensuring that the data is disposed of in a confidential manner? Dr Ben Parcell

Please refer to the last page for the six Caldicott Principles before answering the questions below.

Q.1 What is the purpose for which data are to be used? (Principle 1)

To assess the impact of using sequencing on:

- **Management of HAI's**

Q.2 Why is it necessary to use identifiable data? (Principle 2)

In order to show that this technology will benefit patients and NHS Grampian.

Q.3 Justify the use of each patient-identifiable data field (Principle 3)

Diagnosis – will provide more information of when it is of benefit to use this test.
Age and gender - will give us an idea of which groups of people this test can be used on.

This information will be collected from the Microbiology departments IT system apex.

Application Number .....(for office use only)

NHS
Grampian

Q.4 Who will have access to patient-identifiable information and what control will there be? (Principle 4)

Data will be anonmyised and held on computer in a locked room in the Microbiology department

Q.5 Outline actions taken to ensure individuals with access to patient-identifiable information are aware of their responsibilities and obligations to respect patient confidentiality (Principle 5)

I am aware of the responsibilities and obligations to respect patient confidentiality

Q.6 Outline the organisational arrangements for complying with legal requirements (Principle 6)

We will all ensure that we comply with legal requirements. As every use of patient-identifiable information must be lawful we have asked for Caldicott guardian so that we can extract this information from 'apex' our microbiology testing.

NHS
Grampian

I confirm that the data will be held and used according to the condition and information given as described with this approval form.

Applicant:

Job Title:

Signature:

FOR OFFICE USE ONLY

Data Protection Act compliant **Yes** ☐ **No** ☐

Comments:

Information Governance Manager: Mr Chris Morrice

Signature ..... Date .....

Authorisation Granted **Yes** ☑ **No** ☐

Comments:

Caldicott Guardian (NHS Grampian): Dr Nick Fluck, Medical Director, NHS Grampian

Signature ..... Date 07.06.2017

Application Number .....(for office use only)

6

NHS
Grampian

Applicant Notified YES ✓ NO ☐ DATE 07.06.17

1. The data received from NHS Grampian will be treated as confidential
2. The data received from NHS Grampian will be used only for the purpose(s) described
3. In the case of anonymised or confidential aggregated data, no attempt will be made to identify or contact individuals or organisations identified through this data.
4. The data received from NHS Grampian may be disclosed to staff of the above organisation but only for the described purpose(s)
5. The data received from NHS Grampian may not be disclosed to any third party
6. The data received from NHS Grampian will be stored in secure conditions at all times whether held in electronic medium or as printed hard copies
7. The organisation to which the data is released will maintain and comply with a Data Protection Registration which encompasses the data an data storage usage
8. The data will be destroyed when the work is completed: any printed copies will be destroyed, and files deleted from computer systems (including any copies held on backup or archive media)

All staff given access to data will be made aware of these conditions (Principle 5).

Caldicott Guardian Principles

1. **Justify the purpose(s)**

   Every proposed use or transfer of patient-identifiable information within or from an organisation should be clearly defined and scrutinised, with continuing uses regularly reviewed by an appropriate guardian.

2. **Don't use patient-identifiable information unless it is absolutely necessary.**

   Patient-identifiable information items should not be used unless there is no alternative.

3. **Use the minimum necessary patient-identifiable information.**

   Where use of patient-identifiable information is considered to be essential, each individual item of information should be justified with the aim of reducing identifiability.

4. **Access to patient-identifiable information should be on a strict need to know basis.**

   Only those individuals who need access to patient-identifiable information should have access to it, and they should only have access to the information items that they need to see.

Application Number .....(for office use only) 7

NHS
Grampian

Everyone should be aware of their responsibilities.

Action should be taken to ensure that those handling patient-identifiable information – both clinical and non-clinical staff – are aware of their responsibilities and obligations to respect patient confidentiality.

Understand and comply with the law

Every use of patient-identifiable information must be lawful. Someone in each Organisation should be responsible for ensuring that the organisation complies with legal requirements.

Application Number .....(for office use only)

8.2 Appendix 2 Caldicott approval NHS Tayside

Information Governance
Maryfield House South
Mains Loan
Dundee
DD4 7BT
Tel. 01382 740074
Ext. 70249
www.nhstayside.scot.nhs.uk

Dr Ben Parcell
Consultant Medical Microbiologist
Medical Microbiology
Ninewells Hospital
Dundee
DD1 9SY

Date 26 September 2018
Your Ref
Our Ref IGTCAL5071
Enquiries to Mr J. Donnelly
Extension 70249
Direct Line N/A
Email

Dear Dr Parcell

CALDICOTT APPROVAL – Whole-Genome Sequencing Service for the Investigation of Healthcare Associated Infection

Proposal Sponsor: Dr Ben Parcell, Consultant Medical Microbiologist, NHS Tayside

Data User(s): Dr Ben Parcell, Consultant Medical Microbiologist, NHS Tayside
Dr Catriona Harkins, Specialist Registrar Dermatology, NHS Tayside
Professor Stephen H Gillespie, Sir James Black Professor of Medicine, University of St Andrews
Professor Matt Holden, School of Medicine, University of St Andrews

Caldicott approval is given for you to identify patients from the Microbiology Department computer system (Lab Centre), and to access IPC notes to collect relevant and proportionate data on routine typing results and antibiograms, then to compare the results with those for surplus bacterial isolates that have undergone Whole Genome Sequencing at St Andrews University, in order to assess the impact of using sequencing on the management of Hospital Acquired Infections, as described in your application and supporting information.

It is noted that all data will be anonymised.

Thank you for your co-operation in providing us with the information requested by us in this process.
Please contact me should any queries arise from the application of this approval.

Yours sincerely

Joseph Donnelly
Data Protection Officer

Copies to: Dr Catriona Harkins, Specialist Registrar Dermatology, NHS Tayside
Professor Stephen H Gillespie, Sir James Black Professor of Medicine, University of St Andrews
Professor Matt Holden, School of Medicine, University of St Andrews

8.3 Appendix 3 Biorepository Approval

Grampian Biorepository
Aberdeen Royal Infirmary
Link Building
Aberdeen AB25 2ZD
Tel: 01224 550928

Dr Ben Parcell
NHS Grampian

Wed, 15 Feb 2017

Dear Dr Ben Parcell,

RE: Whole Genome Sequencing Service for the Investigation of Healthcare Associated Infection TR000126

We are pleased to inform you that the Tissue Bank Committee has approved your recent application for the above project.

If the samples you require are within the custody of the Grampian Biorepository or the NHS Pathology Archive please contact William Mathieson (william.mathieson@nhs.net or telephone 01224551924) to arrange release of the requested material for your use. Alternatively please phone the Grampian Biorepository on 01224 550928.

Please note that, unless otherwise indicated approval of projects and the subsequent release and use of tissue samples is given on the following conditions:

(a) Samples should only be used for the purposes outlined in the application
(b) Tissue should not be transferred to a third party (other than collaborators included in the Tissue Request) without the prior agreement of the Grampian Biorepository
(c) No attempt to identify any individual from the materials should be made and the anonymity of donors must be protected and respected at all times
(d) An annual update on the use of the material should be provided to the Grampian Biorepository. In addition, their ultimate fate should be confirmed indicating that they have either been entirely depleted or appropriately discarded.

The continuance of the Tissue Bank facility is dependent on funding from the Chief Scientist Office. The Grampian Biorepository must be acknowledged in any presentation, abstract, publication or product resulting from the use of the tissue. The Grampian Biorepository must also be informed when any paper or abstract resulting from a study is published. Full Terms and Conditions are available on our website: biorepository.nhsgrampian.org

We wish you every success with your investigations.

Kind regards,

William Mathieson
Grampian Biorepository Manager

8.4 Appendix 4 University of St Andrews Ethics approval

University of St Andrews | FOUNDED 1413

University Teaching and Research Ethics Committee

20 November 2017

Dr Benjamin Parcell
School of Medicine

Dear Ben

Thank you for submitting your ethical application which was considered by the School of Medicine Ethics Committee meeting on 19th October 2017 when the following documents were reviewed:

1. Ethical Application Form
2. External Permissions

The School of Medicine Ethics Committee has been delegated to act on behalf of the University Teaching and Research Ethics Committee (UTREC) and has granted this application ethical approval. The particulars relating to the approved project are as follows -

Approval Code:	MD13171	**Approved on:**	19/10/17	**Approval Expiry:**	19/10/22
Term of Approval	5 YEARS				
Project Title:	Whole Genome Sequencing for the Investigation of Healthcare Associated Infection				
Researcher(s):	Dr Benjamin Parcell				
Supervisor(s):	Professor Stephen Gillespie				

Approval is given for the term granted above. Projects which have not commenced within two years of approval must be re-submitted for review by your School Ethics Committee. If you are unable to complete your research within the approval period, you are required to write to your School Ethics Committee Convener to request a discretionary extension of no greater than 6 months or to re-apply if directed to do so, and you should inform your School Ethics Committee when your project reaches completion.

If you make any changes to the project outlined in your approved ethical application form, you should inform your supervisor and seek advice on the ethical implications of those changes from the School Ethics Convener who may advise you to complete and submit an ethical amendment form for review.

Any adverse incident which occurs during the course of conducting your research must be reported immediately to the School Ethics Committee who will advise you on the appropriate action to be taken.

Approval is given on the understanding that you conduct your research as outlined in your application and in compliance with UTREC Guidelines and Policies (http://www.st-andrews.ac.uk/utrec/guidelinespolicies/). You are also advised to ensure that you procure and handle your research data within the provisions of the Data Provision Act 1998 and in accordance with any conditions of funding incumbent upon you.

Yours sincerely

Dr Morven Shearer
Convener of the School of Medicine Ethics Committee

School of Medicine Ethics Committee
Medical and Biological Sciences Building, North Haugh, St Andrews, Fife, KY16 9TF, Scotland, UK
Email: medethic@st-andrews.ac.uk Tel No: 01334 463585
The University of St Andrews is a charity registered in Scotland: No SC013532

8.5 Appendix 5 Sample and study accession numbers for bacterial isolates submitted to European Nucleotide Archive

Title	Study Accession	Sample Accession	Experiment Accession	Run Accession	Tax Id	Scientific name
Patient D ETA	PRJEB21208	SAMEA104148503	ERX2081507	ERR2022352	208964	*P. aeruginosa* PAO1
Patient B ETA	PRJEB21208	SAMEA104148505	ERX2081508	ERR2022353	208964	*P. aeruginosa* PAO1
Bed 4 WHB Drain	PRJEB21208	SAMEA104148502	ERX2081509	ERR2022354	208964	*P. aeruginosa* PAO1
Bed 7 WHB Drain	PRJEB21208	SAMEA104148501	ERX2081510	ERR2022355	208964	*P. aeruginosa* PAO1
Kitchen Ice Machine	PRJEB21208	SAMEA104148506	ERX2081511	ERR2022356	208964	*P. aeruginosa* PAO1
Kitchen	PRJEB21208	SAMEA104148506	ERX2081512	ERR2022357	208964	*P. aeruginosa* PAO1
Kitchen	PRJEB21208	SAMEA104148506	ERX2081513	ERR2022358	208964	*P. aeruginosa* PAO1
Bed 8 WHB	PRJEB21208	SAMEA104148504	ERX2081514	ERR2022359	208964	*P. aeruginosa* PAO1
Patient F ETA	PRJEB21208	SAMEA104148500	ERX2081515	ERR2022360	208964	*P. aeruginosa* PAO1
Patient C ETA	PRJEB21208	SAMEA104148583	ERX2081516	ERR2022361	208964	*P. aeruginosa* PAO1
Patient C ETA	PRJEB21208	SAMEA104148583	ERX2081517	ERR2022362	208964	*P. aeruginosa* PAO1
Patient E ETA	PRJEB21208	SAMEA104148496	ERX2081518	ERR2022363	208964	*P. aeruginosa* PAO1
Patient A Abdominal Drain Fluid	PRJEB21208	SAMEA104148497	ERX2081519	ERR2022364	208964	*P. aeruginosa* PAO1
ICU Entrance WHB drain	PRJEB21208	SAMEA104148498	ERX2081520	ERR2022365	208964	*P. aeruginosa* PAO1

Domestic Service Room WHB	PRJEB21208	SAMEA104148499	ERX2081521	ERR2022366	208964	*P. aeruginosa* PAO1
Domestic Service Room	PRJEB21208	SAMEA104148499	ERX2081522	ERR2022367	208964	*P. aeruginosa* PAO1

WHB, Wash hand basins; ETA, endotracheal aspirate

8.6 Appendix 6 Assessment of the utility of WGS for Gram positive suspected outbreak investigations analysing WGS output versus standard investigations and impact on IPC

Incident	Comparison of WGS preparation/results generation time (days) and routine typing	Additional information provided by WGS	Direct impact on IPC management
MRSA	Routine typing (TAT 27) WGS (TAT 49)	Greater discrimination recent transmission	No
VRE	WGS results not produced in real-time due to batching.	Greater discrimination only one ST80 cluster. Four outbreak isolates differentiated by only 21 SNP sites suggesting a common source. Enhanced 'alert organism' detection	No
optrA gene positive *E. faecalis*	WGS results were not produced in real-time due to batching.	Greater discrimination- gave ST and investigated new/unusual resistance mechanism	No
L. monocytogenes	Routine typing (TAT 8) WGS (TAT 12)	Greater discrimination and enhanced 'alert organism' detection greater clarity that transmission had occurred in 2 patient	Yes- repeated kitchen inspections
GAS general ward	WGS results not produced in real-time due to batching.	Greater discrimination confirmed isolates were ST28 and indistinguishable.	No
GAS maternity unit	WGS results not produced in real-time due to batching.	Greater discrimination- MLST supported *emm* typing	No
GAS Midwife unit	WGS results were not produced in real-time due to batching.	Greater discrimination MLST confirmed transmission had not taken place.	No
GAS Care home	WGS results not produced in real-time due to batching.	Greater discrimination- first two isolates indistinguishable	No
BORSA	WGS results produced in real-time throughout outbreak.	Greater discrimination- isolates closely related. Investigate a new/unusual resistance mechanism Streamlining testing	Yes - range of layered mitigations used

TAT, turnaround time; MRSA, Meticillin-resistant *Staphylococcus aureus*; VRE, Vancomycin-resistant enterococci; GAS, Group A *Streptococcus*; BORSA, Borderline oxacillin resistant *Staphylococcus aureus*

8.7 Appendix 7 Assessment of the utility of WGS for Gram negative suspected outbreak investigations analysing WGS output versus standard investigations and impact on IPC

Incident	**Comparison of WGS preparation/results generation time (days) and routine typing**	**Additional information provided by WGS**	**Direct impact on IPC management**
CPE	Routine typing (TAT 8) WGS (TAT 17)	Greater discrimination- SNP distances. Found mixed sample (ST3 *E. coli),* streamlined testing	Yes -IPC measures and screening carried out
P. aeruginosa ICU	WGS carried out in retrospect	Greater discrimination and streamlined tests, replaced 2 techniques.	No
P. aeruginosa CF clinical tertiary hospital	WGS results were not produced in real-time due to batching	Greater discrimination- confirmed ST/SNP distances	No
ESBL *E.coli* Community hospital	Routine typing (TAT 12) WGS (TAT 11)	Greater discrimination SNP distances showed 7 ST131 isolates were closely related	Yes- range of outbreak measures instigated
ESBL *E. coli* Residential care home	Routine typing (TAT 17) WGS (TAT 23)	Greater discrimination isolates not closely related	Yes- assurance outbreak meeting not required.
ESBL *E. coli* Household transmission	WGS results were not available in real-time due to batching	Greater discrimination gave ST and SNP distance suggested highly similar	No
ESBL *E. coli* Maternity unit	Routine typing (TAT 10) WGS (TAT 8)	Greater discrimination ST and SNP distance given	Yes outbreak ruled out, meeting was not needed
P.aeruginosa NICU	Routine typing (TAT 12) WGS (TAT 10)	Greater discrimination detail of diversity amongst isolates	Yes IPC team opened ward earlier.
K. pneumoniae NICU	Routine Typing (TAT 7) WGS (TAT 50)	Greater discrimination greater detail of diversity	No
K. pneumoniae NICU	Routine Typing (TAT 8) WGS (TAT 11).	Greater discrimination- WGS supported VNTR results and gave ST	No
K. oxytoca NICU	Routine typing (TAT	Greater	Yes – assured

	20) WGS (TAT 11)	discrimination- PFGE DNA degraded. WGS typed (patient 7 had *K. michiganensis*)	this was an outbreak and IPC measures should remain
E.asburiae NICU	Routine typing (TAT 16) WGS (TAT 24).	Greater discrimination isolates >50,000 SNPs.	No

TAT, turnaround time; CPE, Carbapenamase-producing *Enterobacterales*; CF, cystic fibrosis; ESBL, Extended-spectrum β-lactamase (ESBL)

8.8 Appendix 8 Publication – "*Pseudomonas aeruginosa* intensive care unit outbreak: winnowing of transmissions with molecular and genomic typing"

Journal of Hospital Infection 98 (2018) 282–288

Available online at www.sciencedirect.com

Journal of Hospital Infection

journal homepage: www.elsevier.com/locate/jhin

Pseudomonas aeruginosa intensive care unit outbreak: winnowing of transmissions with molecular and genomic typing

B.J. Parcell [a,*], K. Oravcova [b,†], M. Pinheiro [b], M.T.G. Holden [b], G. Phillips [a], J.F. Turton [c], S.H. Gillespie [b]

[a] *Ninewells Hospital & Medical School, Dundee, UK*
[b] *School of Medicine, University of St Andrews, St Andrews, UK*
[c] *Antimicrobial Resistance and Healthcare Associated Infections Reference Unit, Public Health England, Colindale, UK*

ARTICLE INFO

Article history:
Received 6 November 2017
Accepted 3 December 2017
Available online 8 December 2017

Keywords:
Sequencing
Pseudomonas spp.
Resistance
Outbreak
Water

SUMMARY

Background: *Pseudomonas aeruginosa* healthcare outbreaks can be time consuming and difficult to investigate. Guidance does not specify which typing technique is most practical for decision-making.
Aim: To explore the usefulness of whole-genome sequencing (WGS) in the investigation of a *P. aeruginosa* outbreak, describing how it compares with pulsed-field gel electrophoresis (PFGE) and variable number tandem repeat (VNTR) analysis.
Methods: Six patient isolates and six environmental samples from an intensive care unit (ICU) positive for *P. aeruginosa* over two years underwent VNTR, PFGE and WGS.
Findings: VNTR and PFGE were required to fully determine the potential source of infection and rule out others. WGS results unambiguously distinguished linked isolates, giving greater assurance of the transmission route between wash-hand basin water and two patients, supporting the control measures employed.
Conclusion: WGS provided detailed information without the need for further typing. When allied to epidemiological information, WGS can be used to understand outbreak situations rapidly and with certainty. Implementation of WGS in real-time would be a major advance in day-to-day practice. It could become a standard of care as it becomes more widespread due to its reproducibility and lower costs.

Introduction

Pseudomonas aeruginosa is a Gram-negative bacterium that is ubiquitous in moist hospital environments [1,2]. It is an opportunistic pathogen in immunocompromised patients that causes a wide range of infections [2–6]. Hospital water can be a source of outbreaks in neonatal units and both adult and paediatric intensive care units (ICUs), colonizing and forming

* Corresponding author. Address: Ninewells Hospital & Medical School, Dundee DD1 9SY, UK. Tel.: +44(0)1382 632855; fax: +44(0) 1382 496462.
E-mail address: b.parcell@nhs.net (B.J. Parcell).
† Present address: Glasgow Institute of Biodiversity, Animal Health and Comparative Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK.

https://doi.org/10.1016/j.jhin.2017.12.005
0195-6701/

biofilms in water, taps, sinks, toilets, showers and drains [2,7–13]. Routes of transmission include environment to patient – either directly from contaminated water or splashes from water outlets, or indirectly from contaminated hands or equipment. Transmission from colonized patients to the environment and between patients can occur during clinical procedures that create aerosols. Infection can be acquired and arise from the patient's own gut microbiota after pseudomonads have been selected out by antibiotics [14]. Multi-drug resistance in *P. aeruginosa* is common, and the mortality rate in invasive infections is up to 29%; therefore, controlling the spread of this organism is important [15,16].

Differentiating strains is essential to identify routes of transmission of organisms, identify reservoirs and plot potential chains of transmission. Variable number tandem repeat (VNTR) typing, a polymerase-chain-reaction-based method, represents an improvement in speed and reproducibility over pulsed-field gel electrophoresis (PFGE) whilst providing a similar level of discrimination [1,16]. Turton *et al*. suggested that isolates similar by VNTR with no strong epidemiological links between them should be confirmed by PFGE [1]. Newer methods of whole-genome sequencing (WGS) offer the potential for greater resolution and reproducibility, and may be faster at identifying strains in an outbreak and deducing the lines of transmission. WGS has been used in the investigation of a variety of bacterial outbreaks and, in some instances, has been used for the investigation of pseudomonas outbreaks. To the authors' knowledge, this is the first study to report data comparing the utility of rapid WGS with the current typing methods (VNTR and PFGE) [17–27].

Materials and methods

Four patients were identified as colonized or infected with a strain of *P. aeruginosa* with an unusual resistance profile in an ICU at Ninewells Hospital, Dundee between January 2013 and May 2013. A case finding exercise was undertaken using the definition: 'a sample positive with *P. aeruginosa* resistant to imipenem isolated from a patient admitted to ICU since 2012'. The case finding exercise yielded a further five patients; however, only two patients had isolates that had been stored by the hospital laboratory. As such, six patient isolates were available for further testing. Water (pre- and post-flush samples) was sampled from 14 water outlets in the ICU for *P. aeruginosa* on 16th May 2013 (Table I). Monitoring swabs were also taken from 11 water outlet drains on the same day [domestic service room wash-hand basin (WHB), Bed 7 WHB, Bed 8 WHB, kitchen sink, kitchen drinking water tap, domestic service room sink, ventilator room sink, ICU entrance WHB, ward area WHB 1, ward area WHB 2, Bed 4 WHB]. Caldicott guardian approval was gained in order to protect patient confidentiality and enable appropriate information sharing.

Samples and susceptibility testing

All clinical specimens had been collected during routine care and processed at the Department of Medical Microbiology Laboratory, Ninewells Hospital. Environmental samples from each water outlet drain were incubated aerobically on MacConkey agar at 37°C and *Pseudomonas* CN Selective Agar (Oxoid Ltd, Basingstoke, UK) at 35°C, and examined after 24 and 48 h.

Table I
Results of microbiological detection for *Pseudomonas aeruginosa* in water samples (pre- and post-flush) from water outlets in the intensive care unit (ICU)

Source	Pseudomonas count (cfu/mL) pre-flush samples	Pseudomonas count (cfu/mL) post-flush samples
Ice machine	>100	>100
Domestic service room WHB	37	1
Bed 7 WHB	>100	28
Bed 8 WHB	41	0
Kitchen sink	0	0
Kitchen drinking water tap	0	0
Kitchen hydroboil	0	0
Domestic service room sink	0	0
Ventilator room sink	0	0
ICU entrance WHB	0	0
Chilled drinking water dispenser	0	0
Ward area WHB 1	0	0
Ward area WHB 2	0	0
Bed 4 WHB	0	0

WHB, wash-hand basin; cfu, colony-forming units.

VITEK 2 (bioMérieux, Marcy l'Etoile, France) was used for organism identification and antibiotic susceptibility testing using minimum inhibitory concentrations according to the European Committee on Antimicrobial Susceptibility Testing. An external contractor sampled all water outlets for *P. aeruginosa* using pre- and post-flush samples. Water samples were processed by a laboratory approved by the UK Accreditation Service within 4 h of collection.

VNTR and PFGE

Environmental and patient isolates were sent to the Antimicrobial Resistance and Healthcare Associated Infections Reference Unit, Public Health England, Colindale for typing (VNTR typing at nine loci and PFGE), as described previously [1].

WGS and phylogenetic analysis

Isolates were stored on beads at −80°C until processed. The cultures were recultured by the Infection Group, School of Medicine, University of St Andrews. DNA was extracted using QIAamp DNA Mini Kit (Qiagen, Hilden, Germany). The quality of the DNA was measured as A280 nm/A260 nm ratio on NanoVue (GE Healthcare, Little Chalfont, UK), and the concentration of double-stranded DNA was assessed using dsDNA BR Kit on a Qubit 2.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). One-nanogram samples of DNA were used to construct the libraries with Nextera XT kit (Illumina Inc, San Diego, CA, USA). The normalized libraries were sequenced using a 2 × 250 pair-end read of a 500-cycle v2 kit on a MiSeq platform (Illumina Inc, San Diego, CA, USA) using a resequencing workflow. The Illumina sequences generated were deposited in the European Nucleotide Archive under the study accession number ERP023446. Using SMALT (Wellcome Trust Sanger Institute;

www.sanger.ac.uk/resources/software/smalt/), reads were initially mapped to the chromosome of *P. aeruginosa* PAO1 (accession number AE004091), and single nucleotide polymorphisms (SNPs) were identified as described previously [28]. In addition, the chromosomes of a representative selection of *P. aeruginosa* strains – B136-33 (accession number CP004061), DK2 (CP003149), LES431 (CP006937), LESB58 (FM209186), M18 (CP002496), MTB-1 (CP006853), NCGM2.S1 (AP012280), PA1 (CP004054), PA38182 (HG530068), RP73 (CP006245), SCV20265 (CP006931), UCBPP-PA14 (CP000438), VRFPA04 (CP008739) and YL84 (CP007147) – were used to provide a wider context for the hospital isolates. For each of these additional *P. aeruginosa* strains, artificial 250bp pair-end reads fastq files were generated using a python script. The generated fastq files were mapped along with the outbreak isolates to the chromosome of *P. aeruginosa* PAO1 and SNPs. Recombination was detected in the genomes using Gubbins (http://sanger-pathogens.github.io/gubbins/) [29].

The core genome regions of the PAO1 and UCBPP-PA14 chromosome were defined by human curation using pairwise Blast comparisons with each other and other *P. aeruginosa* strains [30]. The Artemis Comparison Tool was used to visualize the comparisons [31]. SNPs falling inside mobile genetic elements were excluded from the core genome, as well as those falling in regions predicted by Gubbins to have occurred by recombination. Phylogenetic trees were constructed separately using RAxML v7.0.4 for all sites in the core genomes containing SNPs, using a general time reversible model with a gamma correction for among-site rate variation [32,33]. For a higher resolution phylogeny, ICU isolates that clustered on a branch with UCBPP-PA14 (CP000438) were mapped to this sequence as described above.

Results

Epidemiology

The ICU ward had been free from outbreaks in 2013. It was an eight-bed unit of a teaching hospital with approximately 950 acute beds. The infection prevention and control team (IPCT) were initially concerned that four patients were colonized/infected with a strain of *P. aeruginosa* with the same resistance profile between January 2013 and May 2013. All patient isolates were sensitive to gentamicin, ciprofloxacin, piperacillin-tazobactam and ceftazidime, and resistant to imipenem; most patients had received carbapenem treatment at some point during their admission. Patients had a mixture of diagnoses on admission.

Fluids such as bed bath water and endotracheal aspirate (ETA) were disposed of in the wash-hand basins. The IPCT visited the ward and gave advice in line with national guidance for the appropriate disposal of these potentially contaminated fluids. Procedures for the decontamination of two small pieces of equipment – the ventilator flow sensor and temperature probes – were also reviewed. These items were decontaminated by immersion in a sink filled with hot soapy water. This method was discontinued and sporicidal wipes were used after ascertaining their suitability with the manufacturer. The decontamination sink in the back room was found to have crusting on taps; these taps were replaced. This was the sink in which probes were decontaminated. WHBs were supplied by Pillar mixer taps with integral thermostats (Figure A, see online supplementary material). The flow from the tap ran close to the sink drains, causing splashing. Taps with flow straighteners were at risk of contamination by biofilm. To become more compliant with National Health Service building regulations, these were removed and sink basins were replaced to remove overflow drains. The ice machine was supplied by cold water via a flexible hose. This was identified as an area where bacteria could proliferate; the hose was therefore replaced by a Water Regulations Advisory Scheme (WRAS)-approved hose. Localized cleaning of all affected outlets was performed. An increased flushing regimen was introduced to remove any biofilm that was present within the affected outlets. The recommended flushing regimen was twice per day for 2 min at a time. Remediation works were successful as there was no growth of *Pseudomonas* spp. on repeat testing of outlets and water following these changes.

Antibiotic susceptibility testing

All isolates were confirmed to be *P. aeruginosa*, and five isolates had an indistinguishable antibiotic susecptibility pattern (Patient A, abdominal drain fluid; Patient B, ETA; Patient D, ETA; Patient E, ETA; Bed 8, WHB).

Environmental investigation

Water (pre- and post-flush samples) was sampled from 14 water outlets in the ICU for *P. aeruginosa* (Table I). Four areas were found to be positive: ice machine (pre- and post-flush), domestic service room WHB (pre- and post-flush), Bed 7 WHB (pre- and post-flush), and Bed 8 WHB (pre-flush). The initial results suggested that the pseudomonas contamination was most likely local to the outlets as the post-flush samples yielded much lower growth results and negative results compared with the pre-flush samples. Monitoring swabs were also taken from 11 water outlet drains, and three were positive (Bed 7 WHB water outlet drain, Bed 4 WHB water outlet drain and ICU entrance WHB water outlet drain).

VNTR and PFGE analysis

VNTR analysis of the isolates from the ICU identified that six isolates belonged to a cluster of related profiles, which included Patients B and D and the four environmental isolates from Bed 8 WHB water, Bed 4 WHB water outlet drain, Bed 7 WHB water outlet drain and kitchen ice machine water (Table II). All of these isolates had VNTR profiles that were similar to the PA14 strain, one of the most abundant clonal complexes in the *P. aeruginosa* population, which can be readily isolated from aquatic sources causing infections in humans [34]. The close relationship of these isolates in the PA14 cluster suggested that these isolates may be part of an outbreak. In contrast, the isolates from Patients A, C, E and F had distinct VNTR profiles, both from one another and the PA14 cluster, and also from the remaining environmental samples, suggesting that these were unlinked and therefore could be ruled out of the outbreak.

PFGE was used to distinguish the PA14 cluster isolates. Analysis of the banding pattern divided the isolates into three distinct subtypes designated NINE04PA-1 (Bed 8 WHB water,

Table II
Variable number tandem repeat (VNTR) profiles of *Pseudomonas aeruginosa* isolates from the intensive care unit (ICU)

Source	Date of sampling	VNTR
Bed 4 WHB water outlet drain	16/05/2013	12,2,1,5,5,2,4,5,11
Bed 7 WHB water outlet drain	16/05/2013	12,2,1,5,5,2,4,5,11
Bed 8 WHB water	16/05/2013	12,2,1,5,5,2,4,5,12
Domestic service room WHB water	16/05/2013	12,3,6,3,1,4,14,5,10
ICU entrance WHB water outlet drain	16/05/2013	12,3,-,3,1,4,14,5,10
Ice machine water	16/05/2013	12,2,1,5,5,2,4,5,14
Patient A abdominal drain fluid	11/03/2012	12,6,7,5,3,4,8,1,11
Patient B ETA	21/09/2012	12,2,1,5,5,2,4,5,12
Patient C ETA	04/01/2013	11,2,6,-,3,6,6,6,12
Patient D ETA	15/04/2013	12,2,1,5,5,2,4,5,12
Patient E ETA	11/05/2013	12,4,-,-,3,1,6,4,13
Patient F ETA	05/05/2013	12,2,-,3,2,2,-,5,6

WHB, wash-hand basin; ETA, endotracheal aspirate.

Patient D ETA, Patient B ETA), NINE04PA-1' (Bed 4 and 7 WHB water outlet drain) and NINE04PA-1" (kitchen ice machine water). There were clear and definite band differences between the ice machine isolate and the patient isolates.

Genome sequencing and phylogenetic reconstruction

WGS and phylogenetic analysis were performed in order to resolve the fine-scale relationship between outbreak isolates and explore epidemiological links between the isolates (Figure 1). In order to provide a wider genetic context and a snapshot of diversity within the species, 15 additional *P. aeruginosa* genome sequences from EMBL nucleotide database were included in the analysis. For this overview of the *P. aeruginosa* population, the WGS reads of the isolates were mapped to the chromosome of the reference strain PA01 [25]. In total, 182,476 SNP sites were identified amongst all analysed isolates and revealed a diverse population structure, throughout which the ICU isolates were distributed. The cluster of isolates identified by VNTR as belonging to the PA14 clone formed a distinct clade in the phylogenetic tree that included the reference isolate UCBPP-PA14, which belonged to the PA14 clone. The next closest isolates to the PA14 cluster were those

Figure 1. Phylogenomic analysis of *Pseudomonas aeruginosa* isolates. Maximum likelihood phylogenetic tree built with core single nucleotide polymorphisms (SNPs) identified by mapping to the PAO1 reference genome is presented on the left of the figure. The box on the right contains a maximum likelihood phylogeny of the intensive care unit isolates belonging to the PA14 clone, where reads were mapped to the PA14 reference genome of UCBPP-PA14R. The tree was built with core SNPs, excluding SNPs identified in regions that had arisen by recombination (the number of SNPs associated with recombination is given in red text above the branches on which they were identified). Scale bars illustrating the relative SNPs' distances of the phylogenetic trees are displayed. ETA, endotracheal aspirate; ICU, intensive care unit; WHB, wash-hand basin.

belonging to Patient C ETA and Patient F ETA, and differed from the cluster by approximately 49,000 SNPs and approximately 50,000 SNPs, respectively.

In order to provide greater resolution for the relationship of the PA14 cluster isolates, the WGS reads were remapped to the reference chromosome of UCBPP-PA14 [26]. This isolate was genetically closer to the outbreak isolates than PAO1, and therefore remapping to this isolate's chromosome would provide increased genomic coverage and consequently greater resolution. Initial phylogenetic analysis of the SNP data mapped to UCBPP-PA14 differentiated the isolates into two separate clusters and a further outlier that were each distinguished by over 1000 SNPs: one cluster containing Patient D ETA, Patient B ETA and Bed 8 WHB water, which was distinguished from the second cluster containing Bed 4 WHB water outlet drain isolate and Bed 7 WHB water outlet drain isolate by 4515 SNPs, which in turn was distinguished from the kitchen ice machine water isolate by 1852 SNPs. Analysis of the distribution of SNPs in the chromosome identified regions of high SNP density, indicative of this variation arising by homologous recombination. Utilizing Gubbins to detect potential regions of recombination identified five regions that distinguished the PA14 clone ICU population. Excluding the SNPs in these regions from the phylogenetic reconstruction reduced the apparent genetic diversity of the PA14 group, but still maintained the distinction of the two clusters and the outlier. In the Patient D ETA, Patient B ETA and Bed 8 WHB water cluster, the Patient B ETA and Bed 8 WHB water isolates were indistinguishable, and differed from the Patient D ETA isolate by four SNPs. The minimal genetic distance between these isolates strongly supports transmission between the Bed 8 WHB water and Patients D and B, and is within the range of SNP distances observed in a study that investigated *P. aeruginosa* transmission in a hospital setting [26]. These patients were not in the ICU department at the same time. An overview of conventional typing and genomic analysis and the timeline for the delivery of these results is illustrated in Figure 2.

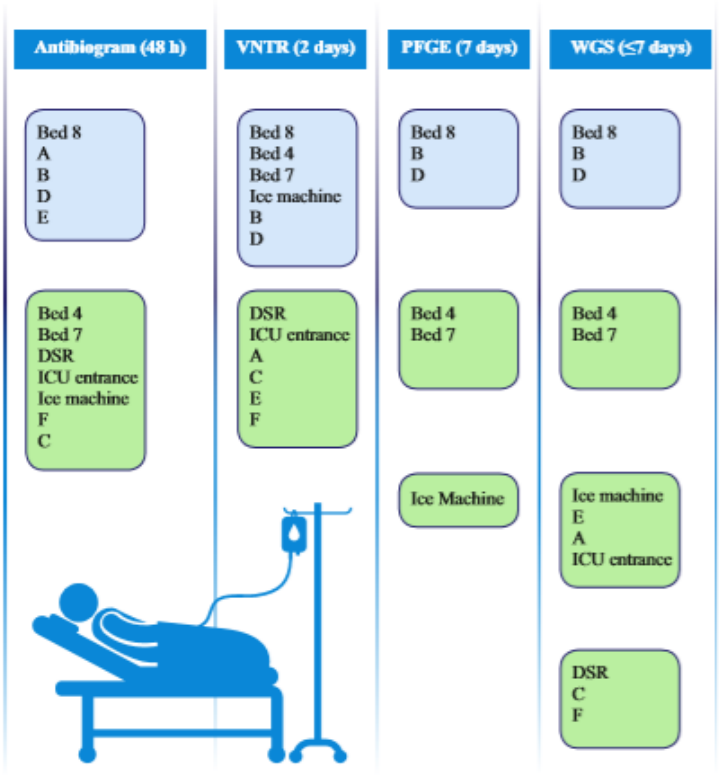


Figure 2. Overview of conventional typing and genomic analysis. Environmental samples: Bed 4, Bed 4 wash-hand basin (WHB) water outlet drain; Bed 7, Bed 7 WHB water outlet drain; Bed 8, Bed 8 WHB water; DSR, domestic service room WHB water; ICU entrance, ICU entrance WHB water outlet drain; ice machine, water. Patient samples: A, Patient A abdominal drain fluid; B, Patient B endotracheal aspirate (ETA); C, Patient C ETA; D, Patient D ETA; E, Patient E ETA; F, Patient F ETA. Light blue, strains deemed 'in' by typing; light green, strains deemed 'out' by typing. VNTR, variable number tandem repeat; PFGE, pulsed-field gel electrophoresis; WGS, whole-genome sequencing; ICU, intensive care unit.

Discussion

The IPCT had urgent questions to answer: is there an outbreak, is there a common source, who is involved, how did the outbreak arise? This study evaluated the relative utility of typing methods to answer these questions. As one moves from routine methods of antibiogram through VNTR and PFGE to WGS, the understanding of the nature of this outbreak becomes apparent. There is a progressive winnowing of possibly involved patients and infection sources. The antibiogram showed a linkage between the Bed 8 WHB water and Patients B and D, but also included Patients A and E. VNTR correctly identified the two patients who were part of the outbreak, but also identified several false-positive environmental links. This left the IPCT considering various routes of transmission; for instance, patient care using ice machine water. PFGE was required for complete clarification. The clinical benefit of using WGS in this situation is that it rapidly provides absolute clarity in distinguishing the isolates in one step, negating the need for the IPCT to spend unnecessary time contemplating other scenarios of how the transmission came about without certainty. In this situation, this information was combined with epidemiology. These patients were not in the ICU at the same time, suggesting that the water supply had acted as a reservoir and source of ongoing infection.

Recognized interventions to prevent transmission from water to patients were effective in preventing further transmission. These included removal of taps with flow straighteners, replacement of sink basins with overflow drains, introduction of increased flushing regimen, and monitoring of water temperature to become fully compliant with national guidance. The IPCT also reviewed procedures for the decontamination of the ventilator flow sensor and temperature probes, in addition to making recommendations for the disposal of potentially contaminated fluids to prevent transmission of organisms from patients to sink drains and distal ends of taps. ICU staff supported changes to their decontamination practices, and training on the new cleaning protocol for sinks was also given. Following this, areas were resampled and it was confirmed that remediation works were successful.

This study has shown that WGS is a potent tool to direct effective intervention in outbreak situations in comparison to, and providing additional information to, molecular typing methods. WGS can provide results within seven days; however, it should be noted that current start-up costs for this technology remain high. For WGS to be introduced into routine clinical microbiology laboratories, investment in infrastructure including bioinformatics and expertise for the interpretation, management and storage of data is required. Standard operating procedures, validation of methods and quality control measures are in place for VNTR and PFGE testing, and will be required for WGS to take place in clinical laboratories. The results are limited by the fact that some isolates were not stored and not all were recultured successfully. Only one colony was processed from each sample, and this may have limited assessment of the diversity of *Pseudomonas* spp. in patient and environment samples.

WGS would be of particular use when there are no obvious epidemiological links between the patients, enabling IPCTs to have timely results using one method. WGS alone provided the necessary resolution to identify the transmission pathway, demonstrating unequivocally the spread between single water supply to patients, and eliminating other potential transmission events and sources. Thought should be given as to how to make these powerful data available routinely in a timely manner, and in a format that is easily interpretable and clinically relevant. Establishing tools such as sequencing machines locally can reduce the turnaround time. It is essential that clinicians develop a new approach to investigate hospital outbreaks, and escalate to WGS at an early stage to allow accurate and rapid description of the causes. It is only by using WGS in real-time that it will be used as a powerful tool to improve patient outcomes.

Acknowledgements

The authors wish to thank Patricia Kydd, Pamela Davidson, Harry McEwan, Derek Foote and Wai-Lum Sung for their assistance in data collection, sampling and development of graphics. In addition, the authors wish to thank the Grampian Biorepository and Tayside Biorepository for their approval for use of isolates.

Conflict of interest statement
None declared.

Funding source
Bioinformatics and computational biology analyses were supported by the University of St Andrews Bioinformatics Unit which is funded by a Wellcome Trust ISSF award (Grant 097831/Z/11/Z). BJP, KO, MP, MTGH, GP and SHG are funded by the Chief Scientist Office through the Scottish Infection Research Network, a part of the SHAIPI consortium (Grant Reference Number SIRN/10).

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jhin.2017.12.005.

References

[1] Turton JF, Turton SE, Yearwood L, Yarde S, Kaufmann ME, Pitt TL. Evaluation of a nine-locus variable-number tandem-repeat scheme for typing of *Pseudomonas aeruginosa*. Clin Microbiol Infect 2010;16:1111–6.
[2] Martin K, Baddal B, Mustafa N, Perry C, Underwood A, Constantidou C, et al. Clusters of genetically similar isolates of *Pseudomonas aeruginosa* from multiple hospitals in the UK. J Med Microbiol 2013;62:988–1000.
[3] Walker JT, Jhutty A, Parks S, Willis C, Copley V, Turton JF, et al. Investigation of healthcare-acquired infections associated with *Pseudomonas aeruginosa* biofilms in taps in neonatal units in Northern Ireland. J Hosp Infect 2014;86:16–23.
[4] Kerr KG, Snelling AM. *Pseudomonas aeruginosa*: a formidable and ever-present adversary. J Hosp Infect 2009;73:338–44.
[5] Pier GB, Ramphal R. *Pseudomonas aeruginosa*. In: Mandell GL, Bennett JE, Dolin R, editors. Principles and practice of infectious diseases. 7th ed. Philadelphia: Churchill Livingstone; 2010. p. 2835–60.
[6] Giamarellou H. Prescribing guidelines for severe pseudomonas infections. JAC 2002;49:229–33.
[7] Durojaiye OC, Carbarns N, Murray S, Majumdar S. Outbreak of multidrug-resistant *Pseudomonas aeruginosa* in an intensive care unit. J Hosp Infect 2011;78:154–5.

[8] Cholley P, Thouverez M, Floret N, Bertrand X, Talon D. The role of water fittings in intensive care rooms as reservoirs for the colonization of patients with *Pseudomonas aeruginosa*. Intensive Care Med 2008;34:1428–33.
[9] Anaissie EJ, Penzak SR, Dignani MC. The hospital water supply as a source of nosocomial infections: a plea for action. Arch Intern Med 2002;162:1483–92.
[10] Aumeran C, Paillard C, Robin F, Kanold J, Baud O, Bonnet R, et al. *Pseudomonas aeruginosa* and *Pseudomonas putida* outbreak associated with contaminated water outlets in an oncohaematology paediatric unit. J Hosp Infect 2007;65:47–53.
[11] Gershman MD, Kennedy DJ, Noble-Wang J, Kim C, Gullion J, Kacica M, et al.; *Pseudomonas fluorescens* Investigation Team. Multistate outbreak of *Pseudomonas fluorescens* bloodstream infection after exposure to contaminated heparinized saline flush prepared by a compounding pharmacy. Clin Infect Dis 2008;47:1372–9.
[12] Hota S, Hirji Z, Stockton K, Lemieux C, Dedier H, Wolfaardt G, et al. Outbreak of multidrug-resistant *Pseudomonas aeruginosa* colonization and infection secondary to imperfect intensive care unit room design. Infect Control Hosp Epidemiol 2009;30:25–33.
[13] Snyder LA, Loman NJ, Faraj LA, Levi K, Weinstock G, Boswell TC, et al. Epidemiological investigation of *Pseudomonas aeruginosa* isolates from a six-year-long hospital outbreak using high-throughput whole genome sequencing. Euro Surveill 2013;18:1–9.
[14] HFS. HIPS and *Pseudomonas aeruginosa* and Water (Scotland) Group Guidance for neonatal units (NNUs) (levels 1, 2 & 3), adult and paediatric intensive care units (ICUs) in Scotland to minimise the risk of *Pseudomonas aeruginosa* infection from water. 2014. Version 2.0.
[15] Castanheira M, Deshpande LM, Costello A, Davies TA, Jones RN. Epidemiology and carbapenem resistance mechanisms of carbapenem-non-susceptible *Pseudomonas aeruginosa* collected during 2009–11 in 14 European and Mediterranean countries. J Antimicrob Chemother 2014;69:1804–14.
[16] Kelsey M. Pseudomonas in augmented care: should we worry? J Antimicrob Chemother 2013;68:2697–700.
[17] Robinson ER, Walker TM, Pallen MJ. Genomics and outbreak investigation: from sequence to consequence. Genome Med 2013;5:36.
[18] Mellmann A, Harmsen D, Cummings CA, Zentz EB, Leopold SR, Rico A, et al. Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. PLoS One 2011;6:e22751.
[19] Köser CU, Holden MT, Ellington MJ, Cartwright EJ, Brown NM, Ogilvy-Stuart AL, et al. A neonatal MRSA outbreak investigation using rapid whole genome sequencing. N Engl J Med 2012;366:2267–75.
[20] Harris SR, Cartwright EJ, Török ME, Holden MT, Brown NM, Ogilvy-Stuart AL, et al. Whole-genome sequencing for analysis of an outbreak of meticillin-resistant *Staphylococcus aureus*: a descriptive study. Lancet Infect Dis 2013;13:130–6.
[21] Vogel U, Szczepanowski R, Claus H, Jünemann S, Prior K, Harmsen D. Ion torrent personal genome machine sequencing for genomic typing of *Neisseria meningitidis* for rapid determination of multiple layers of typing information. J Clin Microbiol 2012;50:1889–94.
[22] Reuter S, Harrison TG, Köser CU, Ellington MJ, Smith GP, Parkhill J, et al. A pilot study of rapid whole-genome sequencing for the investigation of a legionella outbreak. BMJ Open 2013;3: pii:e002175.
[23] Walker TM, Ip CL, Harrell RH, Evans JT, Kapatai G, Dedicoat MJ, et al. Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study. Lancet Infect Dis 2013;13:137–46.
[24] Chin CS, Sorenson J, Harris JB, Robins WP, Charles RC, Jean-Charles RR, et al. The origin of the Haitian cholera outbreak strain. N Engl J Med 2011;364:33–42.
[25] Davis RJ, Jensen SO, Van Hal S, Espedido B, Gordon A, Farhat R, et al. Whole genome sequencing in real-time investigation and management of a *Pseudomonas aeruginosa* outbreak on a neonatal intensive care unit. Infect Control Hosp Epidemiol 2015;36:1058–64.
[26] Quick J, Cumley N, Wearn CM, Niebel M, Constantinidou C, Thomas CM, et al. Seeking the source of *Pseudomonas aeruginosa* infections in a recently opened hospital: an observational study using whole-genome sequencing. BMJ Open 2014;4:1–10.
[27] Mellmann A, Bletz S, Böking T, Kipp F, Becker K, Schultes A, et al. Real-time genome sequencing of resistant bacteria provides real-time genome sequencing of resistant bacteria provides precision infection control in an institutional setting. J Clin Microbiol 2016;54:2874–81.
[28] Hsu LY, Harris SR, Chlebowicz MA, Lindsay JA, Koh TH, Krishnan P, et al. Evolutionary dynamics of methicillin-resistant *Staphylococcus aureus* within a healthcare system. Genome Biol 2015;16:81.
[29] Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucl Acids Res 2014;43:e15.
[30] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215:403–10.
[31] Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J. ACT: the Artemis Comparison Tool. Bioinformatics 2005;21:3422–3.
[32] Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 2006;22:2688–90.
[33] Harris SR, Feil EJ, Holden MT, Quail MA, Nickerson EK, Chantratita N, et al. Evolution of MRSA during hospital transmission and intercontinental spread. Science 2010;327:469–74.
[34] Mikkelsen H, McMullan R, Filloux A. The *Pseudomonas aeruginosa* reference strain PA14 displays increased virulence due to a mutation in ladS. PLoS One 2011;6:e29113.

8.9 Appendix 9- Publication "Clinical Perspectives in Integrating Whole Genome Sequencing into the Investigation of Healthcare and Public Health Outbreaks - Hype or Help?"

Journal of Hospital Infection 109 (2021) 1–9

Available online at www.sciencedirect.com

Journal of Hospital Infection

journal homepage: www.elsevier.com/locate/jhin

Clinical perspectives in integrating whole-genome sequencing into the investigation of healthcare and public health outbreaks — hype or help?

B.J. Parcell [a,*], S.H. Gillespie [b], K.A. Pettigrew [b,†], M.T.G. Holden [b]

[a] Ninewells Hospital and Medical School, Dundee, UK
[b] School of Medicine, University of St Andrews, St Andrews, UK

ARTICLE INFO

Article history:
Received 2 November 2020
Accepted 2 November 2020
Available online 9 November 2020

Keywords:
Whole-genome sequencing
Healthcare-associated infections
Typing
Pulsed-field gel electrophoresis
Variable number of tandem repeats
Multi-locus sequence typing

SUMMARY

Outbreaks pose a significant risk to patient safety as well as being costly and time consuming to investigate. The implementation of targeted infection prevention and control measures relies on infection prevention and control teams having access to rapid results that detect resistance accurately, and typing results that give clinically useful information on the relatedness of isolates. At present, determining whether transmission has occurred can be a major challenge. Conventional typing results do not always have sufficient granularity or robustness to define strains unequivocally, and sufficient epidemiological data are not always available to establish links between patients and the environment. Whole-genome sequencing (WGS) has emerged as the ultimate genotyping tool, but has not yet fully crossed the divide between research method and routine clinical diagnostic microbiological technique. A clinical WGS service was officially established in 2014 as part of the Scottish Healthcare Associated Infection Prevention Institute to confirm or refute outbreaks in hospital settings from across Scotland. This article describes the authors' experiences with the aim of providing new insights into practical application of the use of WGS to investigate healthcare and public health outbreaks. Solutions to overcome barriers to implementation of this technology in a clinical environment are proposed.

Introduction

Whole-genome sequencing (WGS) has several advantages over conventional microbiological typing techniques. It can be applied to all micro-organisms (bacteria, fungi, viruses and parasites) and used to analyse their entire genomes [1]. Since Sanger *et al.* sequenced the first complete DNA genome (bacteriophage ϕX174) using the 'plus and minus' method in 1977, advances in this field have resulted in increased capacity, reduced costs, and improved speed and reproducibility of results, all of which present an opportunity for WGS to be further incorporated into routine microbiological workflows [2–5]. WGS can be used to pinpoint and track bacteria to a greater degree than traditional typing methods, and has been applied to the investigation of a wide variety of outbreaks [6–8]. WGS has also been used to investigate the

* Corresponding author. Address: Ninewells Hospital and Medical School, Dundee DD1 9SY, UK. Tel.: +44(0)1382 632855; fax: +44(0) 1382 496462.
E-mail address: benjamin.parcell@nhs.scot (B.J. Parcell).
† Present address: Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK.

https://doi.org/10.1016/j.jhin.2020.11.001

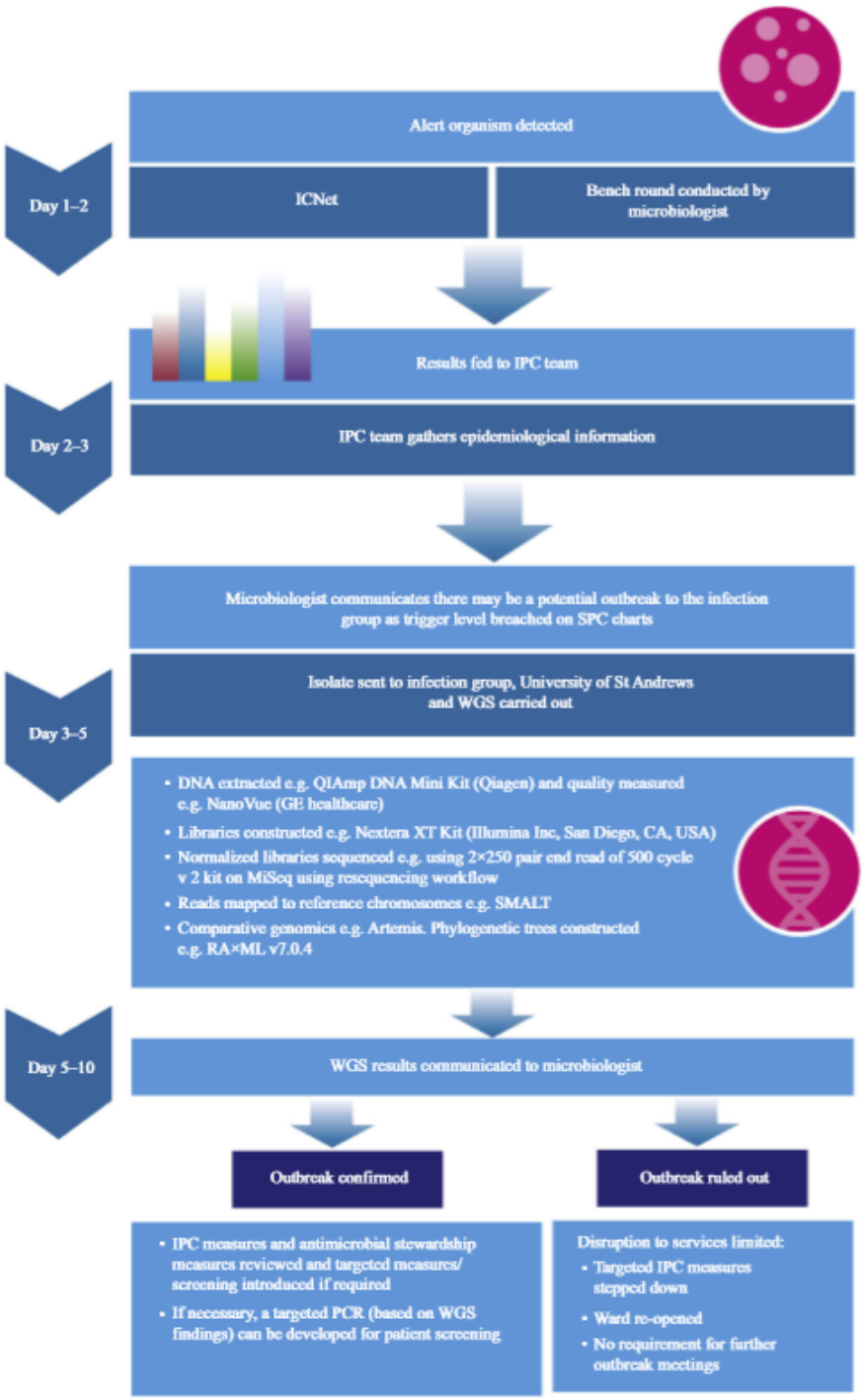


Figure 1. Incorporation of whole-genome sequencing (WGS) into conventional outbreak analysis workflows for in-silico outbreak investigation. IPC, infection prevention and control; SPC, statistical process control; PCR, polymerase chain reaction.

emergence and spread of viruses. Due to its portability, nanopore DNA sequencing technology in the form of the MinION (Oxford Nanopore Technologies, Oxford, UK) was utilized for real-time genome sequencing of the Ebola virus disease epidemic in West Africa and yellow fever virus in Brazil [9,10]. The Zika genome has also been sequenced directly from clinical samples using a protocol involving multiplex polymerase chain reaction (PCR) for MinION and Illumina sequencing [11]. More recently, a combined phylogenetic and epidemiological approach was undertaken using Oxford Nanopore and Illumina MiSeq technology to investigate the first 4 weeks of emergence of severe acute respiratory syndrome coronavirus-2 in Scotland [12].

In recent years, reference laboratories have adopted WGS as a standard typing technique for *Escherichia coli*, *Shigella* spp., *Listeria* spp., *Campylobacter* spp., *Staphylococcus aureus*, *Salmonella* spp. and *Mycobacterium* spp. [13]. In the case of *Salmonella* spp., WGS has emerged as an alternative to the previous gold standard of traditional serology and the Kauffmann–White scheme [14]. Single nucleotide polymorphism (SNP)-based genetic cluster analysis can also be performed to inform epidemiological investigations, such as the UK-wide *Salmonella enteritidis* 25-SNP cluster t25.12 outbreak and national surveillance [15]. Newer bioinformatic methods offer results with improved accuracy, reproducibility and greater resolution. With advances in sequencing technology, there is greater potential to utilize sequencing for real-time outbreak investigations to inform infection prevention and control (IPC) interventions. Greater knowledge and understanding of the transmission of microorganisms can be generated to inform the best ways to prevent the dissemination of antimicrobial resistance. Consideration must be given as to how best to refine this potentially disruptive technology, optimizing its use in clinical settings so that results are timely, clinically relevant and interpretable by the outbreak management team. This article considers how WGS could be used in the investigation of outbreaks in the future. By using examples from the authors' experience of applying WGS as a targeted IPC tool, this article will illustrate the benefits of harnessing the discriminatory power of WGS in outbreak investigations. In doing so, the aim is to provide pointers towards potential solutions when faced with the challenges brought by this new technology. Over a 5-year period, the authors established a WGS service, with no prior infrastructure, to confirm or refute nosocomial outbreaks in real-time in NHS clinical environments. Clinical specimens were first collected as part of routine care, and initially processed in National Health Service microbiology laboratories using standard methods prior to sequencing by the Infection and Global Health Division, University of St Andrews (Figure 1). Over 3 years, 761 isolates from more than 20 nosocomial outbreaks from health boards across Scotland were sequenced. The pathogens sequenced included: *Acinetobacter baumanii*, *Enterobacter cloacae*, *Enterococcus faecalis*, *Enterococcus faecium*, *E. coli*, *Klebsiella oxytoca*, *Klebsiella pneumoniae*, *Listeria monocytogenes*, *Morganella morganii*, *Mycobacterium abscessus*, *Pseudomonas aeruginosa*, *Enterobacter asburiae*, *Serratia marcescens*, *S. aureus*, *Staphylococcus epidermidis* and *Streptococcus pyogenes*.

The value of results with greater granularity due to WGS

It is recognized that routine typing results can obstruct outbreak investigations when results are not rapid, and also when they are unable to show unequivocally whether or not isolates are linked. Pulsed-field gel electrophoresis (PFGE) was previously considered by many to be the gold standard for microbiological typing; however, although it was standardized internationally, it could not decipher evolutionary relationships, or determine whether isolates were truly related in terms of transmission having taken place following identification of a prevalent profile [16,17] Multi-locus sequence typing, which interrogates regions of 400–500 bp from multiple (usually seven) housekeeping genes, is highly reproducible, standardized and easily comparable between laboratories, but lacks discriminatory power for deep epidemiological surveillance [16].

Previously, the authors illustrated the relative discriminatory power of WGS when used for the investigation of an outbreak of *P. aeruginosa* involving nine patients in an adult intensive care unit (ICU) [18]. Both variable number of tandem repeats (VNTR) and PFGE were required to fully determine that transmission had occurred between handwash basin water and two patients. The use of WGS would have provided the necessary information in one step, negating the need for further typing.

WGS can assist IPCTs when routine typing fails to determine whether or not bacterial isolates are linked due to the commonality of distribution of pathogen types

WGS was used to assist the outbreak team [infection prevention and control team (IPCT) and public and environmental health team] in a hospital outbreak investigation of two cases of invasive listeria infection. *L. monocytogenes* is a notifiable foodborne pathogen which can cause gastroenteritis. Early identification of the source is essential in order to prevent further transmission to patients, staff and the public as listeriosis has a crude mortality rate of 20% and is associated with sepsis, meningoencephalitis and, in pregnant patients, abortion or premature delivery [17]. In this investigation, both patients were immunocompromised and had been admitted to hospital at various points. Routine typing identified the listeria isolates as serotype 4 [clonal complex (CC) 1 and sequence type (ST) 1]. Hospital-wide and ward catering facilities had been inspected and no concerns regarding the practice of food hygiene had been noted. The outbreak team were unsure whether there had been transmission as this was one of the more common serotypes found in clinical isolates (one in six invasive listeria isolates are CC1), and consensus opinion amongst the outbreak management team had been that further catering facility inspections were not required. Both listeria isolates underwent WGS. Reads were initially mapped to the reference chromosome of strain F2365, and SNPs were identified against this. The two isolates were found to be indistinguishable, and therefore highly likely to be epidemiologically linked. In this example, WGS revealed additional information that prompted further action to look again for a common link. The outbreak team repeated hospital

kitchen inspections, and as a result of this identified that the handling of salads and meat did not meet national recommendations. Subsequently, hospital catering facilities were closed temporarily until remedial action was undertaken. A further case of invasive serotype 4 listeria infection was identified in a patient 5 months later. WGS ruled this isolate out of the outbreak as it was found to differ by approximately 10,000 SNPs, illustrating that it was too genetically divergent to share a recent common ancestry with the earlier cluster.

Enhancing the detection of 'alert organisms' using genomic analysis

Some countries have an agreed minimum list of micro-organisms that are deemed to pose a risk to patients due to antimicrobial resistance and/or virulence. Flagging up alert organisms such as these allows IPCTs to perform further investigations [19]. The detection of alert bacteria and fungi from patient and/or environmental samples in clinical microbiology laboratories traditionally relies upon the use of phenotypic methods. If the same alert organism is identified from the samples of multiple patients taken within the same timeframe and in a similar place, an outbreak investigation is usually initiated and bacterial isolates will be sent for typing. Statistical process control (SPC) charts can also be used to demonstrate how outbreak strains can accumulate. They display data chronologically and can reveal natural or unnatural variation [20]. However, using approaches which focus solely on alert micro-organisms to detect transmission does not account for the fact that there can be dissemination of genetic elements, such as transposons and plasmids carrying antibiotic resistance genes. This can be a dynamic process, and traditional typing may lack the discriminatory power to identify the genetic lineage of isolates beyond species. This could potentially result in the failure of IPC measures if an alert organism is misidentified or not detected. To explain this point, the benefits of using genomic analysis to detect alert organisms rather than solely accepting results from phenotypic tests was highlighted in the investigation of a seemingly small separate series of vancomycin-resistant *E. faecium* (VREfm) outbreaks. Initially, VREfm was identified in urine cultures from two patients admitted to the same ward of an orthopaedic rehabilitation hospital. Antibiograms were identical, and further screening samples were taken to investigate faecal carriage. Notably,

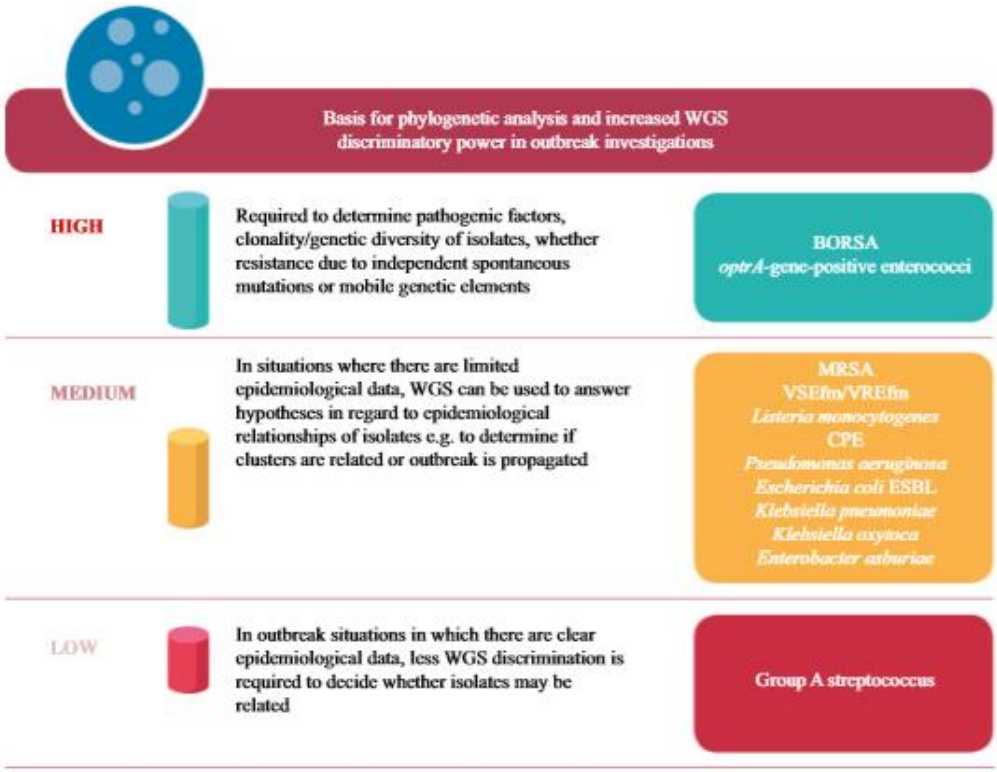


Figure 2. Basis for phylogenetic analysis and increased whole-genome sequencing (WGS) discriminatory power in outbreak investigations. BORSA, borderline oxacillin-resistant *Staphylococcus aureus*; MRSA, meticillin-resistant *S. aureus*; VSEfm, vancomycin-susceptible *Enterococcus faecium*; VREfm, vancomycin-resistant *E. faecium*; CPE, carbapenemase-producing Enterobacterales; ESBL, extended-spectrum beta-lactamase.

one patient was colonized with two different strains of VREfm – one identified from rectal swabs and another from urine – and each strain was related to an entirely different outbreak cluster in the main hospital. This demonstrated the importance of repeated and sequential patient sampling from different body sites, including multiple colonies for sequencing during an enterococcus outbreak investigation. Over a 2-year period, further positive patients were identified on the surgical high dependency unit and renal ward of the main hospital, totalling 11 cases. Routine PFGE typing of all the VREfm isolates from all patients revealed five separate clusters in total (three ST80 clusters, one ST64 culture and one ST203 cluster). When WGS was applied, it revealed that, in fact, there was only one ST80 cluster, rather than numerous discrete clusters as reported by PFGE. There was a difference of 65–77 SNPs between isolates taken from various patients, suggesting a recent common ancestor (i.e. within the last 5 years). The value of WGS to unravel the transmission pathways was demonstrated when two vancomycin-susceptible *E. faecium* (VSEfm) isolates from two separate patients, identified during a separate VSEfm outbreak in the ICU the preceding year, were linked to the ST64 cluster. This was an unexpected finding, uncovering a hidden transmission event of which the IPCT was not aware. The four isolates were differentiated by only 21 SNPs, suggesting a relatively recent common source. The investigation also revealed that interhospital transmission had occurred between local hospitals and a regional hospital performing renal transplants. These results were produced as part of a real-time VREfm outbreak investigation, and support findings from a retrospective study in the UK in which Raven *et al.* applied WGS to the genetic characterization of 293 bacteraemia isolates [21]. They found that the majority of bacteraemias were hospital-acquired or healthcare-associated, and over 50% of isolates were highly related. In 32% of cases, complex transmissions had occurred over a number of years and across various wards, there was a mixture of vancomycin-resistant and vancomycin-susceptible antibiotic profiles, and this was due to isolates having lost or gained transposons carrying the gene encoding vancomycin resistance (*vanA*) [21]. Taken together, these findings indicate that resistance in enterococci is not stable,

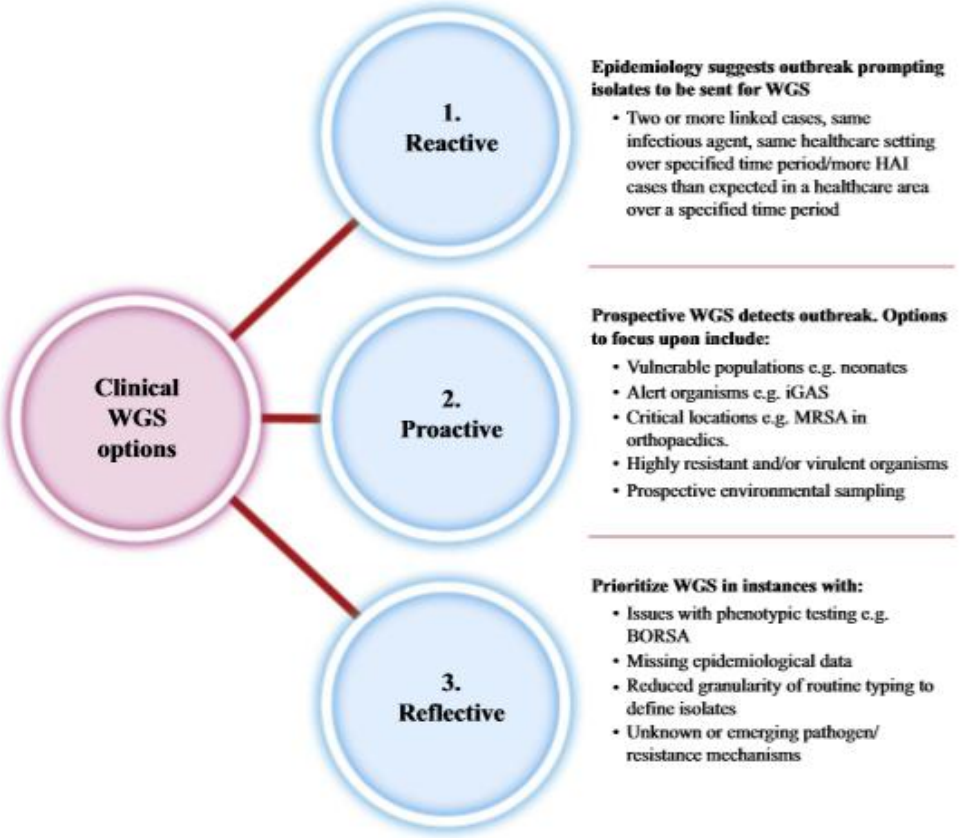


Figure 3. Recommendations for using whole-genome sequencing (WGS) for the detection of nosocomial outbreaks. HAI, hospital-acquired infection; MRSA, meticillin-resistant *Staphylococcus aureus*; BORSA, borderline oxacillin-resistant *S. aureus*.

and the use of enterococci resistance as a marker for transmission is not reliable. IPCTs therefore need to consider alert organisms in the context of background dissemination of genetic elements in their hospital. For instance, if an increase in the number of cases of enterococcal bacteraemia has been detected, it may be necessary to move quickly to deploy pathogen sequencing to type isolates, irrespective of whether isolates are VSEfm or VREfm.

Streamlining outbreak investigations in the laboratory

WGS can aid the in-silico investigation of outbreaks by improving turnaround times of results, thereby rapidly streamlining outbreak investigations. It can also be used to replace unnecessary routine laboratory testing, and reduce the transport of isolates to various reference laboratories for traditional typing. Carbapenemase-producing Enterobacterales are notoriously challenging to detect in clinical laboratories, and multiple methods are routinely used to detect them. In an outbreak involving three renal patients infected with bla_{KPC}-positive ST258 *K. pneumoniae*, WGS was found to negate the need for screening by various multiplex PCR assays in two different laboratories and VNTR analysis. Additionally, WGS revealed that one of the samples was mixed with an ST3 *E. coli* which had not been identified on routine testing.

Ruling out outbreaks with WGS

WGS can also be used by IPCTs to rule out outbreaks, avoiding disruption to services by removing the need for staff and time-consuming outbreak meetings. If WGS can be used to rule out an outbreak, staff can focus on preventive measures and tasks which could stop the occurrence of outbreaks. There would be additional benefits to the hospital and patients as outbreak control measures could be stopped, wards could open, and patients would not require screening. Extra cleaning regimes, such as twice-daily cleaning of commonly touched surfaces with chlorine disinfectant, with increased domestic staff input would not be necessary. In one example illustrating these points, a doctor highlighted that they had noticed three separate patients who had extended-spectrum beta-lactamase-positive *E. coli* urine cultures with identical antibiograms. The patients resided in a residential care home, and the doctor was concerned that this may represent a breakdown in IPC measures. An outbreak was suspected by the IPCT as an alert organism had been detected from the samples of three separate patients taken within the same timeframe. In this situation, the three isolates underwent WGS and were mapped to ST43 (ST131) reference chromosome sequence from an isolate that originated in the UK. Looking at the core genome, differences ranged from 99 SNPs to 162 SNPs, providing reassurance that this was not an outbreak.

Using WGS to uncover new resistance mechanisms

An essential part of IPC is horizon scanning: identifying new threats including new pathogens and resistance mechanisms. Aside from its utility for high-resolution SNP-based typing, WGS can capture the whole-genome inventory of an organism, and can therefore be used as a vital tool for the investigation of emerging resistance mechanisms. *optrA* is an ABC transporter gene, first reported in 2015, that encodes resistance to oxazolidinones such as linezolid via active efflux [22]. National resistance alerts have been issued highlighting the risk to public health as this new resistance mechanism is plasmid-mediated and could potentially transfer to other strains, species and genera present on the skin and gut of humans and animals [23]. In 2016, an IPCT investigation was undertaken as Public Health England's Antimicrobial Resistance and Healthcare Associated Infections Reference Unit confirmed a patient had a urine sample which was positive for *optrA* linezolid-resistant enterococcus. A retrospective search of the hospital's historical laboratory culture results identified two further patients with *E. faecalis* isolates that were linezolid resistant, identified in 2014 and 2015. WGS was applied to investigate these isolates and found that both were positive for the *optrA* gene. A review of the epidemiology revealed that all patients had urinary tract infections. WGS was valuable in this situation as it identified that the isolates were of three distinct sequence types (ST480, ST19 and ST330), confirming that resistance had emerged separately in the *E. faecalis* population.

Barriers and facilitators to translating the promise of WGS into clinical practice

This article has identified some of the barriers and facilitators to translating the promise of WGS into clinical practice in a system-based manner. The first barrier to integration of WGS into conventional outbreak analysis in any clinical microbiology laboratory is infrastructure. Options include placing equipment at national reference laboratories, at hub sites (e.g. large teaching hospitals) or outsourcing to private laboratories. However, the continued progress of sequencing technology has enabled clinical microbiology laboratories to come a step closer to performing low-cost WGS themselves, using simple bench-top technology and user-friendly library preparation protocols [24]. Benefits of placing facilities closer to patients include streamlining of work (replacement of multiple tests and reduction in sending samples away for confirmatory testing) which could result in reduced turnaround times. However, if WGS is incorporated into local teaching hospital facilities, there would need to be investment in appropriately trained staff, equipment and data analysis. It should also be considered that the future of clinical microbiology is changing and less conventional microbiology is being undertaken. There has been increased use of molecular methods such as PCR, point-of-care testing and automation. Staff skill mix is changing; for instance, consultant microbiologists are carrying out less authorization of routine results, and more staff are familiar with DNA extraction and PCR techniques. This could present an opportunity for WGS to be utilized within clinical microbiology services. To increase the success of a business case, the potential impact of WGS on patient management should be included. This could include its impact on antimicrobial stewardship, IPC measures, outbreak investigation and subsequent impact on services. To further strengthen a business case, WGS facilities could be shared with other departments, such as human genetics or a university department, provided that appropriate approvals are in place. Regardless of location, laboratories need to assure the quality of WGS and validate

Figure 4. Whole-genome sequencing (WGS) clinical decision aid. HAI, hospital-acquired infection; IPC, infection prevention and control; SPC, statistical process control; MRSA, meticillin-resistant *Staphylococcus aureus*; PCR, polymerase chain reaction; VRE, vancomycin-resistant enterococci; SCBU, special care baby unit; MLST, multi-locus sequence typing; iGAS, invasive group A streptococcal disease.

their methods [25,26]. It is also essential that laboratories are accredited by the United Kingdom Accreditation Service.

Additional guidance for sharing genomic data needs to be developed so that patient privacy is maintained yet genomic sequences can be shared and used in an early warning system for outbreaks. Thought must be given to producing an actionable result within a useful timeframe, and it is essential for results to be clear and meaningful so they can be interpreted by staff in terms of the clinical picture. Therefore, curriculums for microbiology and infection, public health and IPC training should also incorporate basic training on interpretation of WGS results.

Based on the authors' experience of using real-time WGS for outbreak investigations, instances in which increased WGS discriminatory power and phylogenetic analysis is required have been identified (Figure 2). There continues to be a significant cost to sequencing, and using WGS for the investigation of all outbreaks is not likely to be feasible for the foreseeable future. The discriminatory power of WGS was particularly valuable in outbreak situations with new or unusual resistance mechanisms. The majority of outbreaks required a medium level of WGS discriminatory power to determine if isolates were related. In instances where the epidemiology demonstrated that there was likely to be an outbreak, less WGS discrimination was required.

Considering the findings, it is suggested that there could be different approaches to the use of WGS for outbreak detection (Figure 3). The first approach involves a reactive automated response to epidemiology suggestive of an outbreak. In this instance, phenotypic methods may suggest that there is an outbreak and isolates could then undergo WGS. An alternative to this is a proactive approach in which WGS is used to detect an outbreak regardless of epidemiology. In these instances, users of WGS could prospectively sequence select populations of patients who may be vulnerable to infection who may be from a critical location, such as a neonatal ICU, or focus upon sequencing a highly resistant or virulent organism from patient samples or the environment. Sequencing a specific group of micro-organisms that cause nosocomial infections in a complete geographic area can give a high-resolution view of the pathogen population that can pinpoint the genetic basis of resistance and spread of the pathogen. This would represent a shift in the identification of outbreaks. A reflective approach to outbreak investigations could include the use of WGS in defined instances, when there is missing epidemiological data or in scenarios in which phenotypic testing lacks granularity.

A clinical decision aid (Figure 4) has been developed to assist clinicians in selecting isolates for sequencing, and actions following the production of results. The speed at which outbreaks emerge differs by organism and in accordance with the micro-organism's ecology or clinical setting. Clustered organisms can arise sufficiently far apart in time or location that they are not identified; therefore, local surveillance systems should be set to have a trigger/threshold that prompts IPC action. It is suggested that the process of outbreak identification should be automated by developing systems that function without human coordination with set 'action line' rules. This can be brought about by collecting epidemiological data, using SPC charts and implementing outbreak surveillance software such as ICNet. Care needs to be taken in relation to the interpretation of results, particularly with regard to the meaning of SNP differences. These need to be considered in the context of pathogen genome stability and the environment. For instance, the literature reports that various SNP differences have been found during the investigation of listeria outbreaks, with diversity ranging from zero to five SNPs and, in some outbreaks, up to 42 SNPs [27]. At the time when the authors investigated the listeria outbreak, Public Health England observed that listeria isolates in outbreaks linked to a single food premises can be as many as 20 SNPs apart, which is in contrast to findings in verocytotoxigenic *E. coli* and salmonella incidents, where only isolates within five SNPs of each other would be considered to be linked. This is because *Listeria* spp. can remain as environmental contaminants in premises over many years (G. Hawkins, Health Protection Scotland, Personal Communication, 1 September 2016).

WGS can be an asset at every stage of outbreak management, assisting IPCTs in the formulation of case definitions and supporting or refuting hypotheses in relation to lines of transmission. When coupled with epidemiological data, WGS can provide the ultimate discrimination of results, enabling IPCTs to carry out outbreak investigations efficiently and effectively. Its value also lies in testing the effectiveness of IPC measures and being able to rule out outbreaks, which negates the requirement for outbreak meetings and disruption to healthcare services. With the global threat of dissemination of antimicrobial resistance, WGS is a valuable tool that should be used to generate greater understanding of the development of new resistance mechanisms and dissemination of resistance elements. In the authors' opinion, the benefits of using WGS for outbreak investigation that have been encountered since establishment of a clinical WGS service in 2014 have far outweighed the efforts to confront the challenges of implementing this technology into routine care.

Acknowledgements

Caldicott Guardian approval was gained in order to undertake this study, which enabled protection of patient confidentiality and appropriate information sharing. NHS Biorepository approval was also gained so that surplus patient bacterial isolates could be used. The authors wish to thank NHS Tayside and NHS Grampian microbiology departments/infection prevention and control teams, national reference laboratories staff (Antimicrobial Resistance and Healthcare Associated Infections Reference Unit, Public Health England, Colindale; Scottish MRSA Reference Laboratory and the Scottish Antimicrobial Resistance Satellite Reference Laboratory, Glasgow), and Wai-Lum Sung, Graphic Designer at the University of Aberdeen, for graphics development.

Conflict of interest statement
None declared.

Funding
Bioinformatics and Computational Biology analyses were supported by the University of St Andrews Bioinformatics Unit, which is funded by a Wellcome Trust ISSF award (Grant 097831/Z/11/Z). The Scottish Healthcare Associated Infection Prevention Institute Consortium is funded by the Chief Scientist Office through the Scottish Infection Research Network (SIRN10).

References

[1] Didelot X, Bowden R, Wilson DJ, Peto TEA, Crook DW. Transforming clinical microbiology with bacterial genome sequencing. Nat Rev Genet 2012;13:601–12.

[2] Sanger F, Air GM, Barrell BG, Brown NL, Coulson R, Fiddes C, et al. Nucleotide sequence of bacteriophage phi X174 DNA. Nature 1977;265:687–95.

[3] Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 2016;17:333–51.

[4] Long SW, Williams D, Valson C, Cantu CC, Cernoch P, Musser JM, et al. A genomic day in the life of a clinical microbiology laboratory. J Clin Microbiol 2013;51:1272–7.

[5] Tagini F, Greub G. Bacterial genome sequencing in clinical microbiology: a pathogen-oriented review. Eur J Clin Microbiol Infect Dis 2017;36:2007–20.

[6] Mathers AJ, Stoesser N, Sheppard AE, Pankhurst L, Giess A, Yeh AJ, et al. *Klebsiella pneumoniae* carbapenemase (KPC)-producing *K. pneumoniae* at a single institution: insights into endemicity from whole-genome sequencing. Antimicrob Agents Chemother 2015;59:1656–63.

[7] Mellmann A, Harmsen D, Cummings CA, Zentz EB, Leopold SR, Rico A, et al. Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. PLoS One 2011;6:e22751.

[8] Köser CU, Holden MTG, Ellington MJ, Cartwright EJP, Brown NM, Ogilvy-Stuart AL, et al. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. N Engl J Med 2012;366:2267–75.

[9] Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, et al. Real-time, portable genome sequencing for Ebola surveillance. Nature 2016;530:228–32.

[10] Nikolay B, Thézé J, Azevedo V, Souza RP, Suchard MA, Santos DA, et al. Genomic and epidemiological monitoring of yellow fever virus transmission potential. Science 2018;361:894–9.

[11] Quick J, Grubaugh ND, Pullan ST, Claro IM, Smith AD, Gangavarapu K, et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. Nat Protoc 2017;12:1261–76.

[12] Filipe ADS, Shepherd J, Williams T, Hughes J, Aranday-Cortes E, Asamaphan P, et al. Genomic epidemiology of SARS-CoV-2 spread in Scotland highlights the role of European travel in COVID-19 emergence. medRxiv 2020. 2020.06.08.20124834.

[13] Grant K, Jenkins C, Arnold C, Green J, Zambon M. Implementing pathogen genomics. London: Public Health England; 2018. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/731057/implementing_pathogen_genomics_a_case_study.pdf [last accessed November 2020].

[14] Ibrahim GM, Morin PM. Salmonella serotyping using whole genome sequencing. Front Microbiol 2018;9:1–8.

[15] Health Protection Scotland. HPS Briefing Note 2020/15 *Salmonella enteritidis* 25-SNP cluster t25.12 outbreak. Glasgow: Health Protection Scotland; 2020. Available at: http://www.shpir.hps.scot.nhs.uk/ [last accessed November 2020].

[16] Moura A, Criscuolo A, Pouseele H, Maury MM, Leclercq A, Tarr C, et al. Whole genome-based population biology and epidemiological surveillance of *Listeria monocytogenes*. Nat Microbiol 2016;2:1–10.

[17] Fox EM, Delappe N, Garvey P, McKeown P, Cormican M, Leonard N, et al. PFGE analysis of *Listeria monocytogenes* isolates of clinical, animal, food and environmental origin from Ireland. J Med Microbiol 2012;61:540–7.

[18] Parcell BJ, Oravcova K, Pinheiro M, Holden MTG, Phillips G, Turton JF, et al. *Pseudomonas aeruginosa* intensive care unit outbreak: winnowing of transmissions with molecular and genomic typing. J Hosp Infect 2018;98:282–8.

[19] Health Protection Scotland. Appendix 13 – NHSScotland list of alert organisms/conditions. Glasgow: Health Protection Scotland; 2017. Available at: http://www.nipcm.hps.scot.nhs.uk/media/1362/2017-03-15-final-appendix-13-alert-organisms.pdf [last accessed November 2020].

[20] Health Protection Scotland. Some information on statistical process control (SPC) charts that may be useful for clinical teams. Version 2.0. Glasgow: Health Protection Scotland; 2012. Available at: http://internalmedicineteaching.org/pdfs/SPC-charts-for-clinicians.pdf [last accessed September 2017].

[21] Raven KE, Gouliouris T, Brodrick H, Coll F, Brown NM, Reynolds R, et al. Complex routes of nosocomial vancomycin-resistant *Enterococcus faecium* transmission revealed by genome sequencing. Clin Infect Dis 2017;64:886–93.

[22] Wang Y, Lv Y, Cai J, Schwarz S, Cui L, Hu Z, et al. A novel gene, *optrA*, that confers transferable resistance to oxazolidinones and phenicols and its presence in *Enterococcus faecalis* and *Enterococcus faecium* of human and animal origin. J Antimicrob Chemother 2015;70:2182–90.

[23] Health Protection Scotland. Oxazolidinone-resistance due to optrA in *Enterococcus faecalis*. HPS Wkly Rep 2016;50:230–1.

[24] Kwong JC, Mccallum N, Sintchenko V, Howden BP. Whole genome sequencing in clinical and public health microbiology. Pathology 2015;47:199–210.

[25] Gargis A, Kalman L, Lubin IM. Assuring the quality of next-generation sequencing in clinical laboratory practice. J Clin Microbiol 2016;54:2857–65.

[26] Kozyreva VK, Truong CH, Greninger A, Crandall J, Mukhopadhyay RCV. Validation and implementation of Clinical Laboratory Improvements Act-compliant whole-genome sequencing in the public health microbiology laboratory. J Clin Microbiol 2017;55:2502–20.

[27] Chen Y, Burall LS, Luo Y, Timme R, Melka D, Muruvanda T, et al. *Listeria monocytogenes* in stone fruits linked to a multistate outbreak: enumeration of cells and whole-genome sequencing. Appl Environ Microbiol 2016;82:7030–40.