BMJ Open

Seeking the source of Pseudomonas aeruginosa infections in a recently opened hospital: a role for whole-genome sequencing

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-006278
Article Type:	Research
Date Submitted by the Author:	01-Aug-2014
Complete List of Authors:	Quick, Josh; Institute of Microbiology and Infection, University of Birmingham; Queen Elizabeth Hospital, NIHR Surgical Reconstruction and Microbiology Research Centre Cumley, Nicola; Queen Elizabeth Hospital, NIHR Surgical Reconstruction and Microbiology Research Centre Wearn, Chris; University Hospital Birmingham Foundation Trust, Healing Foundation Centre for Burns Research; Queen Elizabeth Hospital, NIHR Surgical Reconstruction and Microbiology Research Centre Niebel, Marc; Queen Elizabeth Hospital, NIHR Surgical Reconstruction and Microbiology Research Centre Constantinidou, Chrystala; University of Warwick, Division of Microbiology and Immunology Pallen, Mark; University of Warwick, Division of Microbiology and Immunology Moiemen, Naiem; Queen Elizabeth Hospital, NIHR Surgical Reconstruction and Microbiology Research Centre; University Hospital Birmingham Foundation Trust, Healing Foundation Centre for Burns Research Bamford, Amy; Queen Elizabeth Hospital, NIHR Surgical Reconstruction and Microbiology Research Centre; University Hospital Birmingham Foundation Trust, Healing Foundation Centre for Burns Research Oppenheim, Beryl; Queen Elizabeth Hospital, NIHR Surgical Reconstruction and Microbiology Research Centre Loman, Nicholas; University of Birmingham, Institute of Microbiology and Infection
Primary Subject Heading:	Genetics and genomics
Secondary Subject Heading:	Infectious diseases, Pathology, Public health, Surgery
Keywords:	Microbiology < BASIC SCIENCES, BIOTECHNOLOGY & BIOINFORMATICS, Molecular diagnostics < INFECTIOUS DISEASES, Infection control < INFECTIOUS DISEASES, MICROBIOLOGY, Plastic & reconstructive surgery < SURGERY

Title

Seeking the source of *Pseudomonas aeruginosa* infections in a recently opened hospital: a role for whole-genome sequencing

Authors

Joshua Quick [1,2*] Nicola Cumley [2*], Chris M. Wearn [2,3], Marc Niebel [2], Chrystala Constantinidou [4], Mark J. Pallen [4], Naiem S. Moiemen[2,3], Amy Bamford [2,3], Beryl Oppenheim [2#], Nicholas J. Loman [1#]

Affiliations

[1]Institute of Microbiology and Infection, University of Birmingham, Birmingham, United Kingdom

[2]NIHR Surgical Reconstruction and Microbiology Research Centre, Queen Elizabeth Hospital, Birmingham, United Kingdom

[3]Healing Foundation Centre for Burns Research, University Hospital Birmingham Foundation Trust, Birmingham, United Kingdom

[4]Division of Microbiology and Immunology, University of Warwick, Warwick, United Kingdom

*Contributed equally

Correspondence:

Dr Nicholas James Loman

Institute of Microbiology and Infection, University of Birmingham, Birmingham, B15 2TT United Kingdom

Telephone: +44 (0) 121 414 8849

Email: n.j.loman@bham.ac.uk

Dr Beryl Oppenheim

Clinical Microbiology, University Hospitals Birmingham NHS Foundation Trust

Queen Elizabeth Hospital Birmingham, Mindelsohn Way, Edgbaston

Birmingham, B15 2WB

Telephone: +44 (0) 121 371 6523

Email: Beryl.Oppenheim@uhb.nhs.uk

Keywords

High-Throughput DNA sequencing; Pseudomonas aeruginosa; Burns, Hydrotherapy

Word Count

4265

ABSTRACT

Objectives

Pseudomonas aeruginosa is a common nosocomial pathogen responsible for significant morbidity and mortality internationally. Patients may become colonised or infected with *P. aeruginosa* after exposure to contaminated sources within the hospital environment. The aim of this study was to determine whether whole-genome sequencing (WGS) can be used to determine the source in a cohort of burns patients at high risk of *P. aeruginosa* acquisition.

Study design

An observational prospective cohort study.

Setting

Burns care ward and critical care ward in the United Kingdom.

Participants

Patients with >7% total burns by surface area were recruited into the study.

Methods

All patients were screened for *P. aeruginosa* on admission and samples taken from their immediate environment, including water. Screening patients who subsequently developed a positive *P. aeruginosa* microbiology result were subject to enhanced environmental surveillance. All isolates of *P. aeruginosa* were whole-genome sequenced. Sequence analysis looked at similarity and relatedness between isolates.

Results

Whole-genome sequences for 141 *P. aeruginosa* isolates were obtained from patients, hospital water and the ward environment. Phylogenetic analysis revealed eight distinct clades, with a single clade representing the majority of environmental isolates in the burns unit. Isolates from three patients had identical genotypes compared with water isolates from the same room. There was clear clustering of water isolates by room and outlet, allowing the source of acquisitions to be unambiguously identified. Whole-genome shotgun sequencing of biofilm DNA extracted from a thermostatic mixer valve revealed this was the source of a *P. aeruginosa* subpopulation previously detected in water. In the remaining two cases there was no clear link to the hospital environment.

Conclusions

This study reveals that WGS can be used for source tracking of *P. aeruginosa* in a hospital setting, and that acquisitions can be traced to a specific source within a hospital ward.

ARTICLE SUMMARY

Strengths and limitations of the study

- We have demonstrated that whole-genome sequencing can be used for source tracking of *P. aeruginosa* in a hospital setting.
- We show convincing evidence that transmission has occurred directly from water to patients, but other routes are as likely.
- The main limitation of the study was the sample size, which could be attributable to interventions being carried out during the study.
- Our study focused on a burns unit and critical care unit in a newly built hospital. Modes of *P. aeruginosa* transmission may be different in hospitals with different styles of plumbing and on other augmented care units.

INTRODUCTION

Pseudomonas aeruginosa is a ubiquitous Gram-negative bacterium and an important opportunistic pathogen in the healthcare setting. *P. aeruginosa* particularly affects those with impaired host or mucosal immunity and has a broad range of presentations including respiratory infections in cystic fibrosis and mechanically ventilated patients, bloodstream infections in premature neonates and wounds in burns injuries. Nosocomial *P. aeruginosa* outbreaks are frequently reported and associated with water sources such as taps, showers, mixer valves and flow straighteners, sink traps and drains.[1-10] Other potential routes of transmission include cross-infection, for example carriage on the hands of health care workers, and through contaminated medical equipment such as endoscopic devices.[3,5]

In the UK, the role of water in the transmission of *P. aeruginosa* in healthcare settings has become a matter of urgent concern in response to a recent high-profile outbreak affecting a neonatal critical care unit in Belfast in 2012.[11] This source was eventually determined to be sink taps.[11-13] National guidance is now in place detailing enhanced procedures for routine water sampling on augmented care units, with directed interventions such as disinfection and replacement of high-risk plumbing parts required.[14]

Historical phenotypic typing methods for *P. aeruginosa* such as O-antigen serotyping have more recently been replaced by molecular typing methods such as pulsed-field gel electrophoresis (PFGE), variable number tandem repeat analysis (VNTR) and random amplification of polymorphic DNA (RAPD) and multi-locus sequencing typing (MLST). These methods have been used to investigate outbreaks of *P. aeruginosa* within hospitals.[4,15-17] However, such techniques have important limitations for

source tracking of infections in hospitals as they sample limited numbers of sites in the genome which may result in false clustering of unrelated strains.[18] In the past five years, whole-genome sequencing (WGS) has started to be used to investigate outbreaks in hospitals. WGS is attractive because of its digital, sharable format and ultra-high resolution which is able to discriminate two isolates differing by just a single mutation. WGS has been successfully used to determine likely transmission chains during outbreaks of *Staphylococcus aureus*, *Acinetobacter baumannii* and *Klebsiella pneumoniae*.[18-20] Benchtop sequencing instruments now offer a cost-effective approach for bringing bacterial WGS to the clinical environment. [1,21]

In this study, we explore the utility of WGS to determine the likely sources of *P. aeruginosa* in an at-risk population of burns patients. In the UK and US burns patients receive shower cart hydrotherapy as a mainstay of burns treatment.[22-25] A previous hospital audit suggested that up to one-third of such patients became colonised with *P. aeruginosa*. We hypothesised that this high rate of acquisition may relate to transmission from hospital shower water during therapy. We therefore wished to understand the importance of transmission from water compared with alternative routes such as cross-infection and endogenous carriage.

MATERIALS AND METHODS

Hospital setting

An observational, prospective study design was employed in a burns centre serving approximately 13.7 million people across the Midlands region of England with 300 admissions annually. Opened in June 2010, the burns centre comprises a purpose built 15-bed ward with 11 side-rooms and 2 dual-bedded rooms. Patients requiring mechanical ventilation and organ support are usually treated in two self-contained burns cubicles located within the trauma critical care unit (CCU). Despite the observational nature of the study, sampling was carried out during implementation of interim national guidance on control of *P. aeruginosa* issued by the Department of Health. These guidelines were issued in draft form March 2012, and subsequently revised in March 2013. This meant that parallel water sampling and engineering interventions were being undertaken during the period of study. In addition, some enhanced infection prevention measures were also introduced in response to an outbreak of a multi-drug resistant *A. baumannii.*

Study design and patient selection

Ethical approval was sought and received from the National Research Ethics Service committee in the West Midlands (reference number 12/WM/0181). Patients admitted to the burns unit were eligible for the screening phase of the study if they had burns injuries covering greater than 7% total body surface area (TBSA). Patients were screened as soon as possible after admission after they had given written informed consent. When appropriate, legal consultee advice was sought for patients lacking capacity due to emergency treatment. On admission, recruited patients were screened for carriage of *P. aeruginosa* (wounds, urine and stool) using standard

microbiology techniques. Samples were then taken as part of routine microbiology service during the patients stay. Environmental and water samples were taken after the patient was admitted to the burns centre. If during the period of stay *P. aeruginosa* was isolated from a patient sample the patient was recruited into the second phase of the study. In this phase, patients had wound swabs taken at each dressing change as well as twice-weekly urine samples. The patient's environment and water from outlets in their bed space were sampled weekly for the duration of their stay, and after discharge (post-cleaning). Termination of the study was planned after 30 screening patient admissions, or a year, whichever came soonest, after which 10 patients were expected to acquire *P. aeruginosa*.

Microbiological and molecular methods

P. aeruginosa isolates were obtained from wound swab, urine, stool, environmental and water samples. *P. aeruginosa* was isolated from wound swabs, urine and stool by inoculation onto cysteine lactose electrolyte deficient agar (CLED) and cetrimide agar and incubation for 24 hours at 37°C. Stool samples were cultured overnight in a cetrimide enrichment broth before subculture onto CLED. Identification was confirmed by resistance to C-390 and the VITEK® 2 GN identification card. Antibiotic sensitivity assays were performed using the VITEK® 2 AST N-210 card (bioMérieux, Basingstoke, UK).

The patient's environment (shower head rosette, drain, shower chair or trolley, bedside table, patient chair, instruments in contact with the patient) was sampled over a 10cm^2 area by a Polywipe™ sponge. The sponge was placed in tryptic soy broth incubated for 24 hours at 37°C then sub-cultured onto CLED and cetrimide agar. During water sampling, water was taken from the patient's shower, or tap if a

shower was not present. In duplicate, 100ml of water was filtered through a 0.45 micron filter and the filters placed onto CLED plates and cetrimide agar. Plates were incubated at 37°C for 48 hours and the number of organisms per 100ml quantified.

For storage and DNA extraction a single colony was purified from the primary culture plate. When different colony morphologies were observed, a single colony from each type was purified. Additionally, for a randomly selected water sample, 24 colonies were individually picked from one water-filter primary microbiological plate for sequencing. Isolates were stored on Biobank beads at -20°C prior to DNA extraction. Organisms were resuscitated on CLED agar plates and genome DNA either extracted directly using the MOBIO UltraClean Microbial DNA Kit, or from overnight LB broth culture using a Qiagen Genomic-Tip 100G.

DNA extraction and sequencing

Genomic DNA was prepared from single colony picks using the MIOBIO Ultraclean microbial kit (MOBIO, Carlsbad, USA). 1ng input DNA, as quantified by Qubit (Life Technologies, Carlsbad, USA) was used to prepare genomic libraries for sequencing using the Illumina Nextera XT™ DNA sample kit as per manufacturer's protocol (Illumina, San Diego, USA). Libraries were sequenced on the Illumina MiSeq using a paired-end protocol resulting in read lengths between 150 and 300 bases. A single additional sample, isolate 910, was chosen as a representative member of Clade 5 for long-read sequencing. DNA from this sample was fragmented using a Hydroshear (Digilab, Marlborough, MA) using the recommended protocol for 10kb fragments and further size-selected on a Blue Pippen instrument (Sage Science, MA) with a 7kb minimum size cut-off. The library was sequenced on two zero-mode waveguides (ZMWs) using the Pacific Biosciences RS II instrument at the Norwegian

Sequencing Centre, Oslo. C4-P2 chemistry was chosen because it favours long, more accurate reads for *de novo* assembly.

Stool PCR

For detection of *P. aeruginosa* in stool samples using PCR, a stool sample was collected into a stool collection tube containing stool DNA stabilizer. Total DNA was extracted using the PSP Spin Stool DNA Plus kit (Stratec Molecular). PCR amplification of species specific regions of the 16S rDNA gene was carried out using primers PA-SS-F: GGGGGATCTTCGGACCTCA and PA-SS-R: TCCTTAGAGTGCCCACCCG [12] in the following conditions: 0.5μM of each primer, 1.5mM $MgCl_2$, 0.2mM dNTP's using BIOTAQ^{TM} DNA Polymerase and buffer set. After initial denaturation at 96^{o}C for 2 minutes, 30 cycles of 96^{o}C for 30 seconds, 62^{o} C for 30 seconds and 72^{o}C for 30 seconds were completed with a final extension of 72^{o}C for 5 minutes. Products were visualised for size on an 1.5% agarose gel.

Bioinformatics methods

Illumina MiSeq reads from each isolate were adapter and quality trimmed before use with Trimmomatic.[26] Phylogenetic reconstruction of isolates sequenced in this study were combined with data from a global collection of 55 *P. aeruginosa* strains collected world-wide which have been previously analysed by Stewart et al. [27] For each of the published strains, 600,000 paired-end reads were simulated from the complete or draft genome assembly deposited in Genbank. Read sets were mapped against the *P. aeruginosa* PAO1 reference genome using BWA-MEM 0.7.5a-r405 using default settings.[28] Single nucleotide polymorphisms were called using VarScan 2.3.6 and filtered for regions with an excessive number of variants, which may represent regions of recombination or strong Darwinian selection.[29] FastTree

(version 2.1.7) was used for phylogenetic reconstruction. This software estimates an approximate maximum-likelihood tree under the Jukes-Cantor model of nucleotide evolution with a single rate for each site (CAT).[30] Trees were drawn in FigTree (http://tree.bio.ed.ac.uk/software/figtree/).

For *in silico* MLST prediction, trimmed reads were assembled *de novo* using Velvet [31] with a k-mer size of 81 and searched using nucleotide BLAST against the multi-locus sequence database downloaded from the pubMLST website on 2013-08-05 (http://pubmlst.org/paeruginosa/). For Clade 5 isolates, in order to exhaustively search for discriminatory mutations, a nearly complete reference genome was generated by *de novo* assembly using Pacific Biosciences sequencing data. Reads were assembled using the 'RS_HGAP_Assembly.3' pipeline within SMRT Portal v2.2.0. Illumina reads from the same sample were mapped to this draft genome assembly in order to correct remaining indel errors in the assembly using Pilon (http://www.broadinstitute.org/software/pilon/). Isolates belonging to each clade were mapped individually against either the PacBio reference (Clade 5) or *P. aeruginosa* PAO1 (NC_002516) (Clades 3, 4 and 7).

Variants (single nucleotide polymorphisms and short insertion-deletions) were called using SAMtools mpileup and VarScan with an allele frequency threshold of 80%.[29] Non-informative positions and regions of putative recombination were removed, the later with a variant density filter of more than 3 SNPs every 1000 nucleotides. Analysing samples in each clade individually maximised the number of variants detected by reducing the likelihood of the position being uncovered by a subset of samples. From these variants fine-grained phylogenetic trees were reconstructed for each clade using FastTree. The scripts used to perform this analysis are available at

http://www.github.com/joshquick/snp_calling_scripts. Approximate-maximum-likelihood phylogenetic trees were generated using FastTree and FigTree. For whole-genome shotgun metagenomics analysis, reads were analysed using the Kraken taxonomic classifier software with the supplied *minikraken* database.[32] Reads from the metagenomics dataset were aligned to *P. aeruginosa* PAO1 as in the previous section and phylogenetic placement was carried out using pplacer in conjunction with FastTree.[33] Sequence data is available from the European Nucleotide Archive for the Illumina data (ERP006056) and the corrected Pacific Biosciences assembly (ERP006058).

RESULTS

Study results

Recruitment lasted a period of 300 days, ending according to protocol after the enrolment of 30 screening patients. One additional eligible patient did not consent to enter the study and was excluded. The average age in the study group was 41 years. Males predominated with a male-to-female ration of 2.3:1. Flame burns were the most common mechanism of injury, followed by scalds and mixed flame/flash injuries. The average burn size of the study group was 12.5% of the total body surface area (TBSA) and 27% of patients sustained an inhalation injury. Eight patients required admission to ITU and the majority required surgical treatment of their burns with excision and skin grafting (80%). A large majority of the study group (83%) received shower cart hydrotherapy as a routine part of their wound management to encourage healing through wound debridement and decontamination. The average length of hospital stay (LOS) was 17 days and taking into account burn size, the average was 1.4 days per % TBSA.

The water and environment in burns and critical care units are frequently colonised by *P. aeruginosa*

A total of 282 water and environmental samples were screened for *P. aeruginosa* of which 39/78 (50%) were positive in water samples, 25/96 (26%) were positive from the wet environment and 7/108 (6%) were positive from the dry environment. A total of 83 genome sequences were generated from the 71 positive, as in some cases multiple colony picks were sequenced. When placed in the context of a global collection of *P. aeruginosa* strains, phylogenetic reconstruction demonstrated isolates in our study fell into eight clades (figure 1, Panel A). As has been reported previously, there was no strong association between ecological context and position in the phylogenetic tree.[27] Isolates in this study are most closely related to strains from a variety of settings. The majority of isolates (52%) belong to Clade E (figure 1 Panel B), whose nearest sequenced relative is the Liverpool Epidemic Strain, a clone often isolated from patients in the UK and Canada with cystic fibrosis. [34,35] Isolates from clade E were found in burns unit water and the ward environment, as well as from two patient's wounds. However it was never detected in the critical care unit. Clade E was detected throughout the study in a total of 10 different rooms (figure 2).

Detection of potential transmission events by whole-genome sequencing

Microevolutionary changes occurring over rapid time-scales (i.e. days to months) have been used to detect potential chains of transmission in hospital and community outbreaks.[18-20,36,37] The number of distinct mutations between given isolates has been used to infer whether transmission events are likely to have occurred. Such inferences are aided by prior knowledge of mutation rates in similar

populations. Two patients (1 and 4) in our study both had *P. aeruginosa* from clade E isolated from their wounds. These isolates had an indistinguishable genotype from those present in water and the environment of the room they were nursed within (figure 1 Panel C and figure 3). This genotype was detected in the patient's shower water after initial patient screening, during screening of the second patient admission, twice during the second patient's stay and then 127 days later (days 27, 65, 89 and 216 respectively). When water isolates were positive, the genotype was also detected in wet environment sites (shower drain, shower rosette and patient's trolley) on the same days.

Patient E was nursed on the critical care unit due to concomitant medical problems. *P. aeruginosa* belonging to clade G was isolated from sputum during this time.. Identical genotypes were detected contemporaneously in the water from the associated sink and sink tap handle (see online supplementary appendix 4).

Two further patients (patients 2 and 3) were positive for *P. aeruginosa*. Isolates from these patients belonged to clades C and D respectively. Neither clade was ever isolated from hospital water. In both cases, identical genotypes were detectable in the environment associated with the patient but these were not detected before or after the patients' stay, indicating that the environment was not persistently contaminated. During the course of patient 3's stay, the dry environment such as the bedside table was contaminated, as was the patient's door handle and shower chair. However, after patient discharge, the strain associated with this patient was never seen again during the course of the study in any location.

Whole-genome sequencing permits source tracking of *P. aeruginosa* to individual water outlets

Whole-genome sequencing has been reported previously for source tracking, but never for the detection of transmission events from hospital water.[38] Phylogenetic reconstruction within Clade E, the most commonly detected water clone demonstrated additional diversity within this clone, with a total of 46 mutations detected an average genetic distance between isolates of 4.1 mutations (figure 3). The reconstruction demonstrated clear evidence of clustering of genotypes both by room and outlet (figure 3). When *P. aeruginosa* was detected in the wet environment (e.g. shower rosettes and drains) these genotypes were most often identical to those found in water, indicating that the water was likely the ultimate source of that clone. Genotypic variation was seen between outlets within the same room. For example, tap water sampled from room 11 had a distinct genotype from that sampled from shower water in the same room and this was consistently found over multiple samplings. Notably, isolates from two patients fell within the cluster originating from shower water, indicating that shower hydrotherapy was the most likely source of infection. Two plasmids (designated pBURNS1 and pBURNS2) were detected in this study set, which both demonstrated geographical clustering, with pBURNS1 only being detectable in isolates from room 8 and pBURNS2 only being detectable in isolates from the shower water in room 9.

Rapid evolution of antibiotic resistance associated with treatment

P. aeruginosa is commonly associated with antibiotic resistance due to a number of predisposing features including intrinsic resistance, a repertoire of efflux pumps, antibiotic-inactivating enzymes including beta-lactamases and natural transformability. [39] Three infected patients (2, 3 and 5) received antibiotic therapy, and in each case this was associated with the development of resistance to at least

one therapeutic agent. Associated mutations were detected that were either partially or fully explanatory of the phenotype.

Patient 2 was treated with ciprofloxacin, nitrofurantoin and vancomycin (see online supplementary appendix 11 for full details). 8/21 (38%) tested isolates from this patient were ciprofloxacin resistant. 7/8 isolates (88%) of the ciprofloxacin-resistant strains were distinguishable from the other isolates by a single SNP in *mexS* (annotated as PA2491 in *P. aeruginosa* PAO1) (see online supplementary appendix 1 and 7). This SNP was predicted to result in a non-synonymous amino acid substitution. Disruption of this gene has been shown to cause increased expression of the mexEF-oprN multidrug efflux pump, associated with resistance to quinolones.[40]

Patient 4 was treated with meropenem, piperacillin/tazobactam, flucloxacillin and colistin. Five isolates collected 10 to 18 days after initiation of meropenem showed resistance to imipenem and intermediate resistance to meropenem (see online supplementary appendix 3 and 9). The most likely mutation responsible for this phenotype was detectable in two isolates, both of which had a frame-shift mutation in the gene coding for the membrane porin *oprD.* [41]

Patient 5 had a prolonged stay in ITU and had multiple medical problems including *A. baumannii* infection and was treated with nine antibiotic agents including ciprofloxacin, meropenem and piperacillin-tazobactam. Serial isolates from this patient demonstrated the stepwise acquisition of two mutations. The first was in *nalC*, a probable repressor of the TetR/AcrR family (see online supplementary appendix 10). [42] On inspection of the sequence alignment in this region, a large deletion of 196 nucleotide bases was seen compared to the reference PAO1 strain.

This mutation was seen in association with full resistance to piperacillin-tazobactam, ceftazidine, aztreonam, meropenem and intermediate resistance to ciprofloxacin. This deletion is likely to result in over-expression of efflux pumps involving the *mexAB-oprM* operon. [42,43] Ciprofloxacin resistance in a later isolate corresponded to the stepwise acquisition of a second mutation. This mutation is predicted to affect the well-studied DNA gyrase subunit A gene (*gyrA)* which is strongly associated with ciprofloxacin resistance.[44]

Confirmation of *P. aeruginosa* genotypes in biofilms by whole-genome metagenomic shotgun sequencing

P. aeruginosa is able to produce and survive in biofilms. Plumbing parts such as flow straighteners, shower rosettes, flexible hoses, solenoid valves and thermostatic mixer valves (TMV) are particularly at risk of biofilm formation due to factors including surface areas, convoluted designs and inadequate pasteurisation. [45] To confirm the presence of *P. aeruginosa* in water fittings associated with rooms on the burns unit, we obtained a TMV removed by the hospital estates team from the shower in room nine as part of compliance with UK guidelines for managing *P. aeruginosa* in hospitals. On visual inspection, a biofilm was present which was scraped from the surface with a sterile scalpel. DNA from this biofilm was extracted for whole-genome shotgun sequencing. The majority of reads did not map to any known bacterial taxa. The most abundant taxon identified was *Pseudomonas aeruginosa* (3%). Subsequent alignment to the *P. aeruginosa* PAO1 reference covered 94% of the 6.3 million base reference genome at a median coverage of 5x, confirming that reads were correctly classified to this species and not other environmental *Pseudomonas* species. Alignment to the *P. aeruginosa* clade E reference genome followed by phylogenetic placement of reads demonstrated that it

fell into the same clade as previously recovered isolates from the shower or tap in room 9 (indicated on figure 3, and in online supplementary appendix 6).

DISCUSSION

The hospital environment has been intimately linked with *P. aeruginosa* infection for over 50 years yet hospital acquisitions, clusters and outbreaks remain a common occurrence and understanding precise routes of transmission can be difficult. [45,46] Our results demonstrate that, even in a new hospital, *P. aeruginosa* can become rapidly endemic in hospital plumbing. Furthermore, by linking *P. aeruginosa* genotypes recovered from patients to specific individual water outlets, we offer compelling evidence of unidirectional transmission from water to patients. Further, by sequencing of a biofilm identified in a TMV from a hospital water system, we can identify the likely common source of genotypes found in water and in the hospital environment.

Our results suggest that use of whole-genome sequencing can reduce ambiguity about potential transmission events in hospitals and consequently inform infection prevention efforts about the direction and sequence of transmission. It is notable that the burns unit was colonised by a single clone, meaning that it was very unlikely that water outlets at each bed space were colonised as a result of transmissions from the patient or environment. For this to happen would require multiple transmission events from separate patients with the same clone, for which there is not evidence. Instead we speculate that this clone was introduced to the hospital associated with its commissioning. One hypothesis is that particular plumbing fittings, i.e. the TMV may have been colonised simultaneously by a clone circulating in water. Clone E (ST395) has been frequently reported associated with water, so this remains a

possibility. [47,48] However, it is possible that plumbing fittings are installed 'pre-seeded' with *P. aeruginosa* as has already been proposed by Kelsey.[3,5,45] Investigation of an outbreak in Wales implicated new plumbing parts as a potential source of *P. aeruginosa.* New plumbing components are often tested by companies prior to their supply and it is possible they were contaminated prior to distribution. The limited amount of diversity (average 4 SNPs) seen within this clade is consistent with a single founding genotype coinciding with the opening of the burns unit, based on estimates from a previous study using WGS which reported that mutations accumulate at a rate of approximately one every 3-4 months in a hospital-associated clone.[49] However our results suggest that our isolates accumulate mutations even more slowly. This may be due to reduced growth rates in nutritionally-poor biofilms.[50]

It is notable that antibiotic resistance to multiple first-line agents developed rapidly in response to therapy. These results underline the importance of selecting appropriate antibiotic therapy in *P. aeruginosa* infections. It is reassuring however that antibiotic resistance genotypes selected *in vivo* did not show evidence of persistence in the ward environment or transmission to other patients.

Our study has certain limitations. Based on a previous audit, we expected around one-third of patients screened for *P. aeruginosa* would develop colonisation or clinical infection. In fact, only 5 out of 30 of patients were colonised. This may have been related to guidance and engineering interventions being put in place during the study as detailed in national guidance issued whilst this study was on-going. In addition, infection control policies were revised to address control of an outbreak of a multi-drug resistant *A. baumannii* in this same burns unit. Following these

interventions, only 1 of the last 20 patients recruited was infected with *P. aeruginosa* which may demonstrate the importance of national guidance in reducing transmissions.

By focusing on burns patients who receive hydrotherapy, our study population were at extremely high risk of waterborne infection. In other patient groups it may be that alternative routes of transmission including cross-infection or endogenous carriage play a more important role. Our results suggest that our burns unit is endemically colonised with a distinct clone of *P. aeruginosa* that may have been imported coinciding with the opening of the hospital. Other intensive care units, particularly those which have been open for longer may have harbour a greater diversity of *P. aeruginosa* as a result of increased opportunities for clones to be imported.

One potential application for WGS in infection control would be to determine whether cases are as a result of water transmission, or represent sporadic clones originating from the wider environment. Despite improved guidance concerning improved engineering infection control practices and the introduction o the water safety group in the UK, it may not be realistic to eliminate *P. aeruginosa* from hospitals entirely. In augmented care units such as ITUs, burns units and neonatal wards where *P. aeruginosa* poses a significant risk to vulnerable patients, the increased resolution offered by WGS will justify its use, particularly as the costs continue to fall.

In conclusion, we have identified through WGS clear evidence for transmission of *P. aeruginosa* from specific water outlets to burns patients and offer a forensic-level framework for dealing with outbreaks linked to hospital water. We expect WGS will continue to make inroads into clinical microbiology and become a vital tool for

tracking *P. aeruginosa* in the hospital environment, helping inform targeted control measures to help protect patients at risk of infection.

Acknowledgments

We are grateful to Mark Webber for discussions on antibiotic resistance and to Paul Keim for discussion on phylogenetic placement of metagenomics samples. We thank Chris Thomas for useful discussions about *P. aeruginosa* plasmids. We thank Lex Nederbragt, Ave Tooming-Klunderud and the staff of the Norwegian Sequencing Centre, Oslo for Pacific Biosciences sequencing. We thank Matthew Smith-Banks for laboratory assistance with processing samples. We thank Jimmy Walker for critical reading of the manuscript.

Contributors

MP NM BO conceived the study. CW and AB enrolled patients into study and collected samples. NC collected environmental and water samples. NC CC MN processed samples and performed microbiology. NC CC JQ did sequencing. JQ NC NL analysed the data. NL NC JQ MP BO wrote the paper. All authors commented on the manuscript draft.

Funding

This paper presents independent research funded by the National Institute for Health research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Competing interests

None

Data Sharing Statement

Pacific Biosciences raw data files are available from the corresponding author (Nicholas J Loman, n.j.loman@bham.ac.uk) and will be made available via Data Dryad.

Ethics approval

The study protocol received approval from National Research Ethics Service committee in the West Midlands (reference number 12/WM/0181).

REFERENCES

1 Reuter S, Sigge A, Wiedeck H, *et al.* Analysis of transmission pathways of Pseudomonas aeruginosa between patients and tap water outlets. *Crit Care Med* 2002;**30**:2222–8. doi:10.1097/01.CCM.0000025907.12833.10

2 Struelens MJ, Rost F, Deplano A, *et al.* Pseudomonas aeruginosa and Enterobacteriaceae bacteremia after biliary endoscopy: an outbreak investigation using DNA macrorestriction analysis. *Am J Med* 1993;**95**:489–98.

3 DiazGranados CA, Jones MY, Kongphet-Tran T, *et al.* Outbreak of Pseudomonas aeruginosa infection associated with contamination of a flexible bronchoscope. *Infect Control Hosp Epidemiol* 2009;**30**:550–5. doi:10.1086/597235

4 Crivaro V, Di Popolo A, Caprio A, *et al.* Pseudomonas aeruginosa in a neonatal intensive care unit: molecular epidemiology and infection control measures. *BMC Infect Dis* 2009;**9**:70. doi:10.1186/1471-2334-9-70

5 Moolenaar RL, Crutcher JM, San Joaquin VH, *et al.* A prolonged outbreak of Pseudomonas aeruginosa in a neonatal intensive care unit: did staff fingernails play a role in disease transmission? *Infect Control Hosp Epidemiol* 2000;**21**:80–5. doi:10.1086/501739

6 Trautmann M, Lepper PM, Haller M. Ecology of Pseudomonas aeruginosa in the intensive care unit and the evolving role of water outlets as a reservoir of the organism. *Am J Infect Control* 2005;**33**:S41–9. doi:10.1016/j.ajic.2005.03.006

7 Foca M, Jakob K, Whittier S, *et al.* Endemic Pseudomonas aeruginosa infection in a neonatal intensive care unit. *N Engl J Med* 2000;**343**:695–700. doi:10.1056/NEJM200009073431004

8 Kolmos HJ, Thuesen B, Nielsen SV, *et al.* Outbreak of infection in a

burns unit due to Pseudomonas aeruginosa originating from contaminated tubing used for irrigation of patients. *J Hosp Infect* 1993;**24**:11–21.

9 Widmer AF, Wenzel RP, Trilla A, *et al.* Outbreak of Pseudomonas aeruginosa infections in a surgical intensive care unit: probable transmission via hands of a health care worker. *Clin Infect Dis* 1993;**16**:372–6.

10 Srinivasan A, Wolfenden LL, Song X, *et al.* An outbreak of Pseudomonas aeruginosa infections associated with flexible bronchoscopes. *N Engl J Med* 2003;**348**:221–7. doi:10.1056/NEJMoa021808

11 Wise J. Three babies die in pseudomonas outbreak at Belfast neonatal unit. BMJ. 2012;**344**:e592.

12 Breathnach AS, Cubbon MD, Karunaharan RN, *et al.* Multidrug-resistant Pseudomonas aeruginosa outbreaks in two hospitals: association with contaminated hospital waste-water systems. *J Hosp Infect* 2012;**82**:19–24. doi:10.1016/j.jhin.2012.06.007

13 Walker JT, Jhutty A, Parks S, *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. doi:10.1016/j.jhin.2013.10.003

14 *Health Technical Memorandum*

04-01: Addendum. Department of Health https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/140105/Health_Technical_Memorandum_04-01_Addendum.pdf (accessed 11 Dec2013).

15 Jefferies JMC, Cooper T, Yam T, *et al.* Pseudomonas aeruginosa outbreaks in the neonatal intensive care unit--a systematic review of risk factors and environmental sources. *J Med Microbiol* 2012;**61**:1052–61. doi:10.1099/jmm.0.044818-0

16 Cholley P, Thouverez M, Hocquet D, *et al.* Most multidrug-resistant Pseudomonas aeruginosa isolates from hospitals in eastern France belong to a few clonal types. *J Clin Microbiol* 2011;**49**:2578–83. doi:10.1128/JCM.00102-11

17 Curran B, Jonas D, Grundmann H, *et al.* Development of a multilocus sequence typing scheme for the opportunistic pathogen Pseudomonas aeruginosa. *J Clin Microbiol* 2004;**42**:5644–9. doi:10.1128/JCM.42.12.5644-5649.2004

18 Harris SR, Cartwright EJP, Török ME, *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. doi:10.1016/S1473-

3099(12)70268-2

19 Lewis T, Loman NJ, Bingle L, *et al.* High-throughput whole-genome sequencing to dissect the epidemiology of Acinetobacter baumannii isolates from a hospital outbreak. *J Hosp Infect* 2010;**75**:37–41. doi:10.1016/j.jhin.2010.01.012

20 Snitkin ES, Zelazny AM, Thomas PJ, *et al.* Tracking a hospital outbreak of carbapenem-resistant Klebsiella pneumoniae with whole-genome sequencing. *Sci Transl Med* 2012;**4**:148ra116. doi:10.1126/scitranslmed.3004129

21 Loman NJ, Misra RV, Dallman TJ, *et al.* Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol* 2012;**30**:434–9. doi:10.1038/nbt.2198

22 Tredget EE, Shankowsky HA, Joffe AM, *et al.* Epidemiology of infections with Pseudomonas aeruginosa in burn patients: the role of hydrotherapy. *Clin Infect Dis* 1992;**15**:941–9.

23 Langschmidt J, Caine PL, Wearn CM, *et al.* Hydrotherapy in burn care: A survey of hydrotherapy practices in the UK and Ireland and literature review. *Burns* Published Online First: 10 December 2013. doi:10.1016/j.burns.2013.11.006

24 Tredget EE, Shankowsky HA, Rennie R, *et al.* Pseudomonas infections in the thermally injured patient. *Burns* 2004;**30**:3–26.

25 Davison PG, Loiselle FB, Nickerson D. Survey on current hydrotherapy use among North American burn centers. *J Burn Care Res* 2010;**31**:393–9. doi:10.1097/BCR.0b013e3181db5215

26 Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* Published Online First: 28 April 2014. doi:10.1093/bioinformatics/btu170

27 Stewart L, Ford A, Sangal V, *et al.* Draft genomes of 12 host-adapted and environmental isolates of Pseudomonas aeruginosa and their positions in the core genome phylogeny. *Pathog Dis* Published Online First: 26 October 2013. doi:10.1111/2049-632X.12107

28 Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010;**26**:589–95. doi:10.1093/bioinformatics/btp698

29 Koboldt DC, Zhang Q, Larson DE, *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012;**22**:568–76. doi:10.1101/gr.129684.111

30 Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* 2010;**5**:e9490. doi:10.1371/journal.pone.0009490

31 Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 2008;**18**:821–9. doi:10.1101/gr.074492.107

32 Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 2014;**15**:R46. doi:10.1186/gb-2014-15-3-r46

33 Matsen FA, Kodner RB, Armbrust EV. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* 2010;**11**:538. doi:10.1186/1471-2105-11-538

34 Scott FW, Pitt TL. Identification and characterization of transmissible Pseudomonas aeruginosa strains in cystic fibrosis patients in England and Wales. *J Med Microbiol* 2004;**53**:609–15.

35 Panagea S, Winstanley C, Parsons YN, *et al.* PCR-based detection of a cystic fibrosis epidemic strain of Pseudomonas Aeruginosa. *Mol Diagn* 2003;**7**:195–200.

36 Gardy JL, Johnston JC, Ho Sui SJ, *et al.* Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *N Engl J Med* 2011;**364**:730–9. doi:10.1056/NEJMoa1003176

37 Bryant JM, Grogono DM, Greaves D, *et al.* Whole-genome sequencing to identify transmission of Mycobacterium abscessus between patients with cystic fibrosis: a retrospective cohort study. *Lancet* 2013;**381**:1551–60. doi:10.1016/S0140-6736(13)60632-7

38 Lienau EK, Strain E, Wang C, *et al.* Identification of a salmonellosis outbreak by means of molecular sequencing. *N Engl J Med* 2011;**364**:981–2. doi:10.1056/NEJMc1100443

39 Lambert PA. Mechanisms of antibiotic resistance in Pseudomonas aeruginosa. *J R Soc Med* 2002;**95 Suppl 41**:22–6.

40 Sobel ML, Neshat S, Poole K. Mutations in PA2491 (mexS) promote MexT-dependent mexEF-oprN expression and multidrug resistance in a clinical strain of Pseudomonas aeruginosa. *J Bacteriol* 2005;**187**:1246–53. doi:10.1128/JB.187.4.1246-1253.2005

41 Quinn JP, Dudek EJ, DiVincenzo CA, *et al.* Emergence of resistance to imipenem during therapy for Pseudomonas aeruginosa infections. *J Infect Dis* 1986;**154**:289–94.

42 Cao L, Srikumar R, Poole K. MexAB-OprM hyperexpression in NalC-type multidrug-resistant Pseudomonas aeruginosa: identification and characterization of the nalC gene encoding a repressor of PA3720-PA3719. *Mol Microbiol* 2004;**53**:1423–36. doi:10.1111/j.1365-2958.2004.04210.x

43 Llanes C, Hocquet D, Vogne C, *et al.* Clinical strains of Pseudomonas aeruginosa overproducing MexAB-OprM and MexXY efflux pumps simultaneously. *Antimicrob Agents Chemother* 2004;**48**:1797–802.

44 Cambau E, Perani E, Dib C, *et al.* Role of mutations in DNA gyrase genes in ciprofloxacin resistance of Pseudomonas aeruginosa susceptible or resistant to imipenem. *Antimicrob Agents Chemother* 1995;**39**:2248–52.

45 Kelsey M. Pseudomonas in augmented care: should we worry? *J Antimicrob Chemother* 2013;**68**:2697–700. doi:10.1093/jac/dkt288

46 Rogers DE. The changing pattern of life-threatening microbial disease. *N Engl J Med* 1959;**261**:677–83. doi:10.1056/NEJM195910012611401

47 Martin K, Baddal B, Mustafa N, *et al.* Clusters of genetically similar isolates of Pseudomonas aeruginosa from multiple hospitals in the UK. *J Med Microbiol* 2013;**62**:988–1000. doi:10.1099/jmm.0.054841-0

48 Slekovec C, Plantin J, Cholley P, *et al.* Tracking down antibiotic-resistant Pseudomonas aeruginosa isolates in a wastewater network. *PLoS One* 2012;**7**:e49300. doi:10.1371/journal.pone.0049300

49 Snyder LA, Loman N, Faraj LA, *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**.

50 Donlan RM. Biofilm formation: a clinically relevant microbiological process. *Clin Infect Dis* 2001;**33**:1387–92. doi:10.1086/322972

Figure and Table Legends

Figure 1 Legend

An overview of all samples collected during the study in global phylogenetic context with other sequenced strains of *P. aeruginosa*. Samples collected in this study are widely dispersed in the tree, which contains isolates from different environments (Panel A). Bar plots indicate the numbers of each type of sample collected (Panel C).

Microdiversity within each clade is shown, with the colour bar indicating the source of each sample (Panel C).

Figure 2 Legend

A schematic view of the 300-day study of *P. aeruginosa* in a burns centre and critical care unit. Time in days is shown along the X-axis with location along the Y-axis. Each circular icon indicates a positive isolate of *P. aeruginosa*. The icon's logotype indicates which environment it originated from (wound, urine/sputum, environment or water). The filled colour of the icon indicates the clade it belongs to. Patient icons represent the enrolment of a screening patient into the study and their location. Patient movements around the hospital are noted by dotted lines. The five patients infected with *P. aeruginosa* are denoted by rounded boxes. Boxes are coloured according to the patient number.

Figure 3 Legend

The high-resolution phylogenetic reconstruction of clade E isolates. This demonstrates the clustering of genotypes by bed space. Patient associated samples are contained within the room 11 clade. This clade contains water samples from the shower and environmental samples from the shower, drain and trolley. The water samples from the room 11 tap are in a distinct clade, indicating the biofilm within the tap has a distinct genotype to the shower. This also indicated environmental contamination was more likely to arise from contaminated shower water than tap water. Details of sampling site, days since start of study and presence of pBURNS plasmids are also shown. The likely phylogenetic position of *P. aeruginosa* detected

in a biofilm from a thermostatic mixer valve is shown in the clade associated with room 9 and indicated ‘TMV’.

Appendix 1

Phylogenetic tree, sample origin and antibiotic-resistance phenotype for clade C isolates.

Appendix 2

Phylogenetic tree, sample origin and antibiotic-resistance phenotype for clade D isolates.

Appendix 3

Phylogenetic tree, sample origin and antibiotic-resistance phenotype for clade E isolates.

Appendix 4

Phylogenetic tree, sample origin and antibiotic-resistance phenotype for clade G isolates.

Appendix 5

A physical map of the burns unit, indicating individual patient bed areas, shower areas. The water supply is indicated by a blue line.

Appendix 6

Clade E isolates in phylogenetic context with the metagenomics sample from a tap mixer valve.

Appendix 7

Single nucleotide and small indel variants detected within clade C isolates.

Appendix 8

Single nucleotide and small indel variants detected within clade D isolates.

Appendix 9

Single nucleotide and small indel variants detected within clade E isolates.

Appendix 10

Single nucleotide and small indel variants detected within clade G isolates.

Appendix 11

The type of antibiotics administered to the five patients in this study.

Appendix 12

Mutations predicted to be associated with antibiotic-resistance.

A

B

C

	Strain	MLST	Origin/comment
1	Clade A	667	This study
2	PABL056	298	Blood, USA
3	E2UOS	235	Soil
4	NCGM2.S1	235	Urinary tract infection
5	39016	235	Keratitis
6	Clade B	Novel	This study
7	Clade C	319	This study
8	152504	560	Sputum, non-CF
9	PAb1	560	Frostbite
10	CI27	Novel	Sputum, CF
11	PA14	Novel	Plant, USA
12	J1532		CF, mucoid
13	J1385		CF, progenitor of J1532
14	ATCC 700888		Industrial water system
15	B136.33	1024	Diarrhoea
16	18A	788	CF
17	PAK	693	PAK
18	2192		CF
19	PA17	810	CF, progenitor of PA17SCV
20	PA17SCV	810	Lab evolved from CF Isolate
21	C1426		CF, progenitor of C1433
22	C1433	132	CF, mucoid
23	NCMG1179	1285	Respiratory, Japan
24	ATCC 25324	Novel	Glass-crusher air plate
25	Clade D	348	This study
26	CIG1	348	Sputum, CF
27	DQ8	267	Crude oil, China
28	MRW44.1	Novel	Mutator lineage 44
29	PA01	Novel	Burn
30	PAO579	Novel	Mucoid
31	MSH3		Environmental, Mount St. Helens
32	MSH-10	689	Environmental
33	MSH10 UoS	689	Environmental, Mount St. Helens
34	PA213BR	277	Polymixin resistant, Brazil
35	PA9BR	277	Brazil, non-CF
36	PA19BR	277	Polymixin resistant, Brazil
37	ATCC 14886	260	Soil
38	N002		Crude oil, India
39	AES-1R		Australia, CF
40	E2UoW	41	Plant, USA
41	VRFPA02	111	Eye, India
42	Clade E	395	This study
43	LESB58	146	Liverpool epidemic strain
44	PA62		Soil
45	C1334	439	CF, mucoid, from same patient as C763
46	Clade F	Novel	This study
47	Clade G	17	This study
48	PACS2	1394	CF
49	PA45	266	Blood, Italy
50	XMG	Novel	Soil, China
51	138244	175	Sputum, non-CF
52	PA21_ST175	175	Blood, MDR
53	C763	179	CF, from same patient as C1334
54	Clade H	155	This study
55	AH16	155	China, sputum, non-CF
56	LCT-PA102	155	Blood
57	SJTD-1	Novel	Crude oil, China
58	C3719		Manchester epidemic strain
59	DK2	386	CF, Denmark
60	DK2_CF510	386	CF, Denmark
	M18	1239	Plant isolate, China

Water
Environment
Wound swab
Sputum
Urine

10,000 SNPs

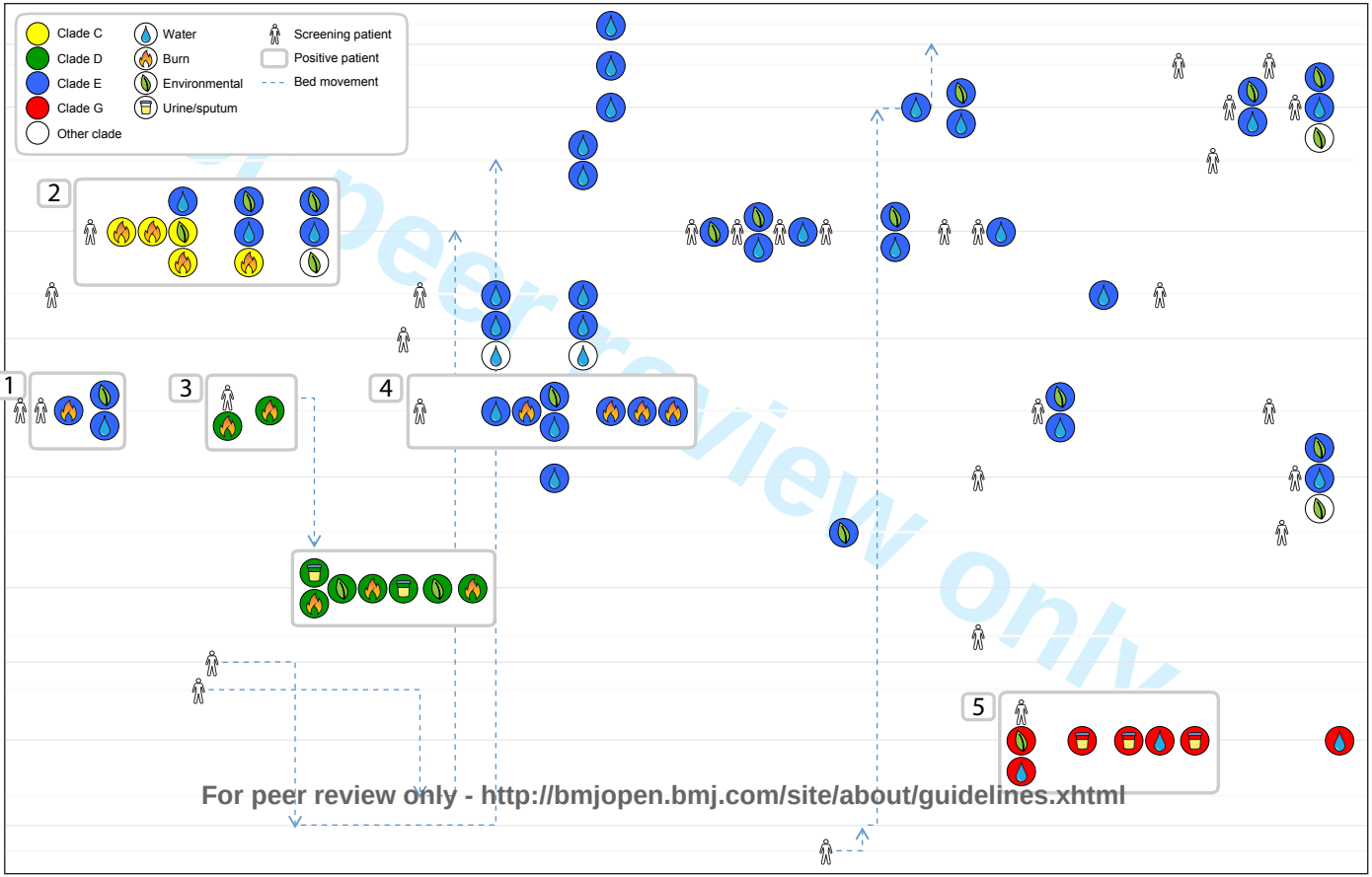
Number of isolates 0 20 40 60

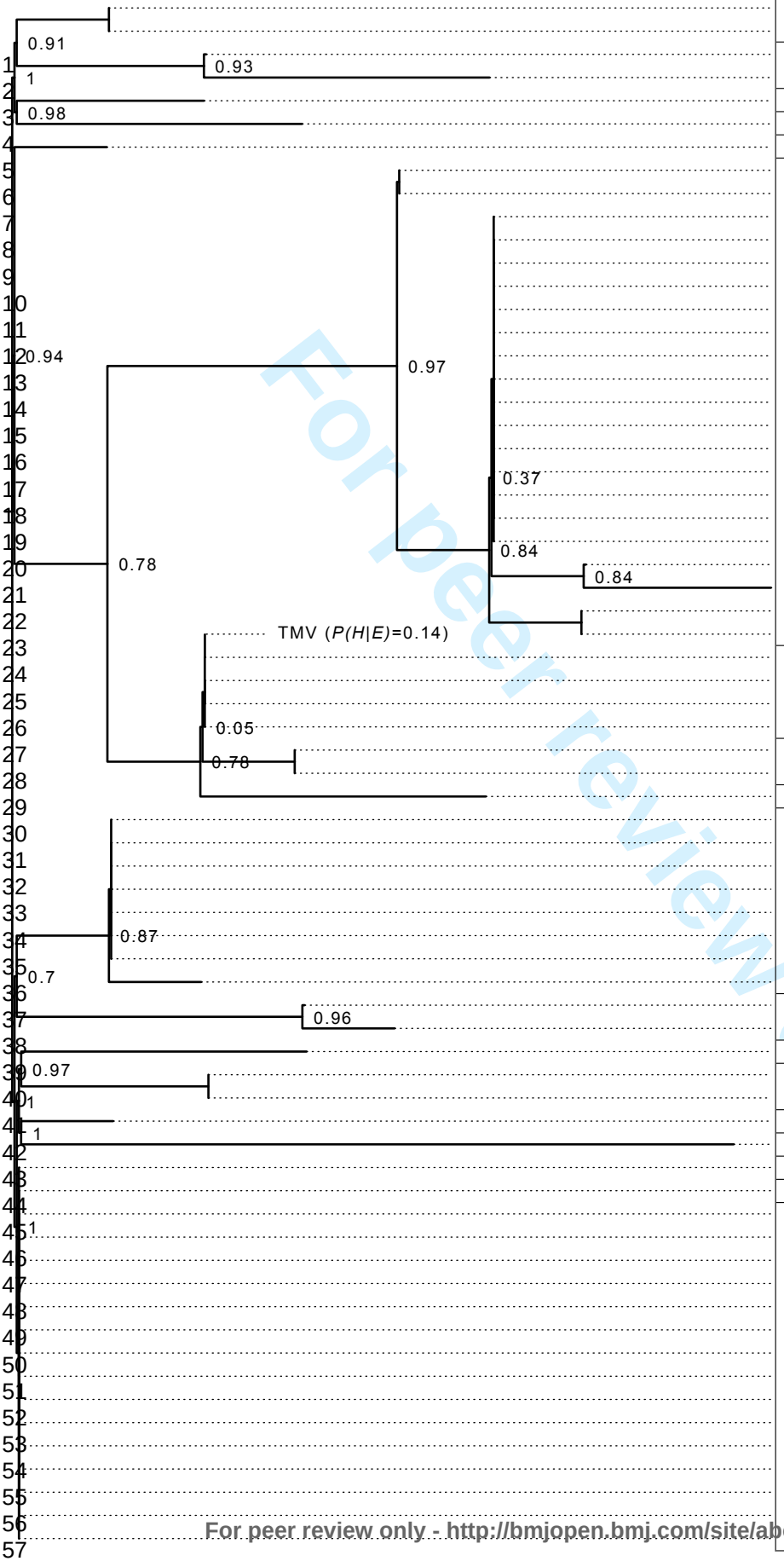
0.5 SNP/Indel

1 SNP/Indel

1 SNP/Indel

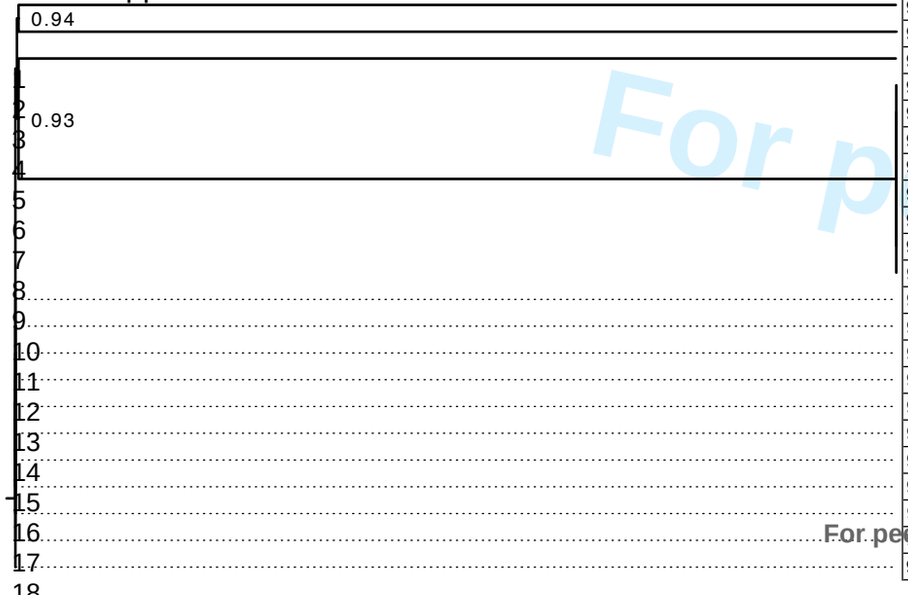
See figure 3



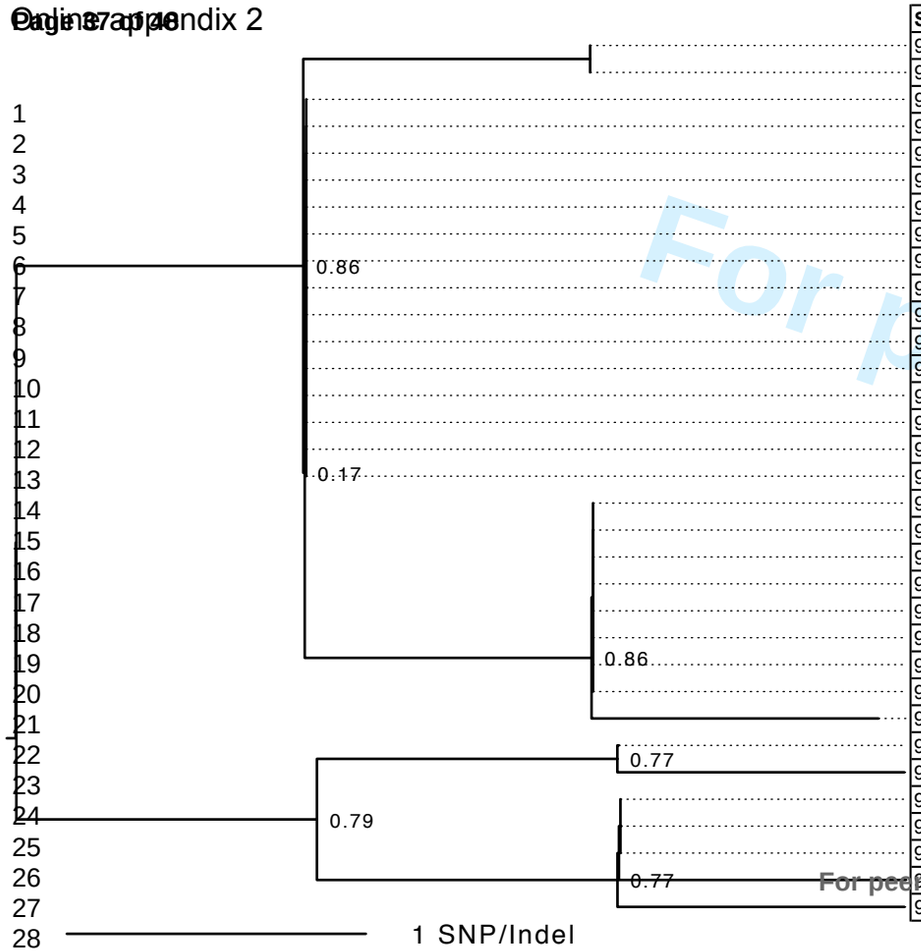


Bed	Type of Specimen	Days	pB1	pB2
12	Shower	292		
	Shower	292		
10	Shower	90		
	Shower	65		
8	Shower	49		
12	Shower	292		
8	Shower	176		
11	Drain	27		
	Shower	27		
	Drain	89		
	Shower	89		
	Shower	216		
	Trolley	27		
	Shower	89		
	Shower	216		
	Shower	89		
	Shower	65		
	Groin	14		
	Anterior right thigh	53		
	Anterior right thigh	103		
	Anterior right shin	53		
	Dorsum of left foot	85		
	Dorsum of left foot	85		
	Dorsum of left foot	85		
	Anterior right thigh	93		
	Anterior right thigh	103		
	Anterior right thigh	91		
	Anterior right thigh	91		
9	Shower	226		
	Shower	90		
	Shower	90		
	Tap	90		
11	Tap	89		
	Tap	89		
9	Shower	65		
6	Shower	292		
	Shower	278		
	Shower	91		
	Shower	292		
	Shower	177		
	Shower	91		
	Shower	278		
	Shower	190		
1	Shower	91		
	Shower	91		
8	Sink	113		
12	Shower	89		
	Shower	89		
6	Shower	190		
8	Shower	155		
5	Tap	91		
7	Shower	90		
8	Drain	146		
	ECG leads	49		
	Shower	49		
	Shower	146		
	Shower	55		
	Shower	55		
	Shower	40		
	Shower	40		
	Shower	49		
	Shower	146		
	Shower	176		
	Shower	40		
	Shower	49		
	Shower	203		
	Shower	55		

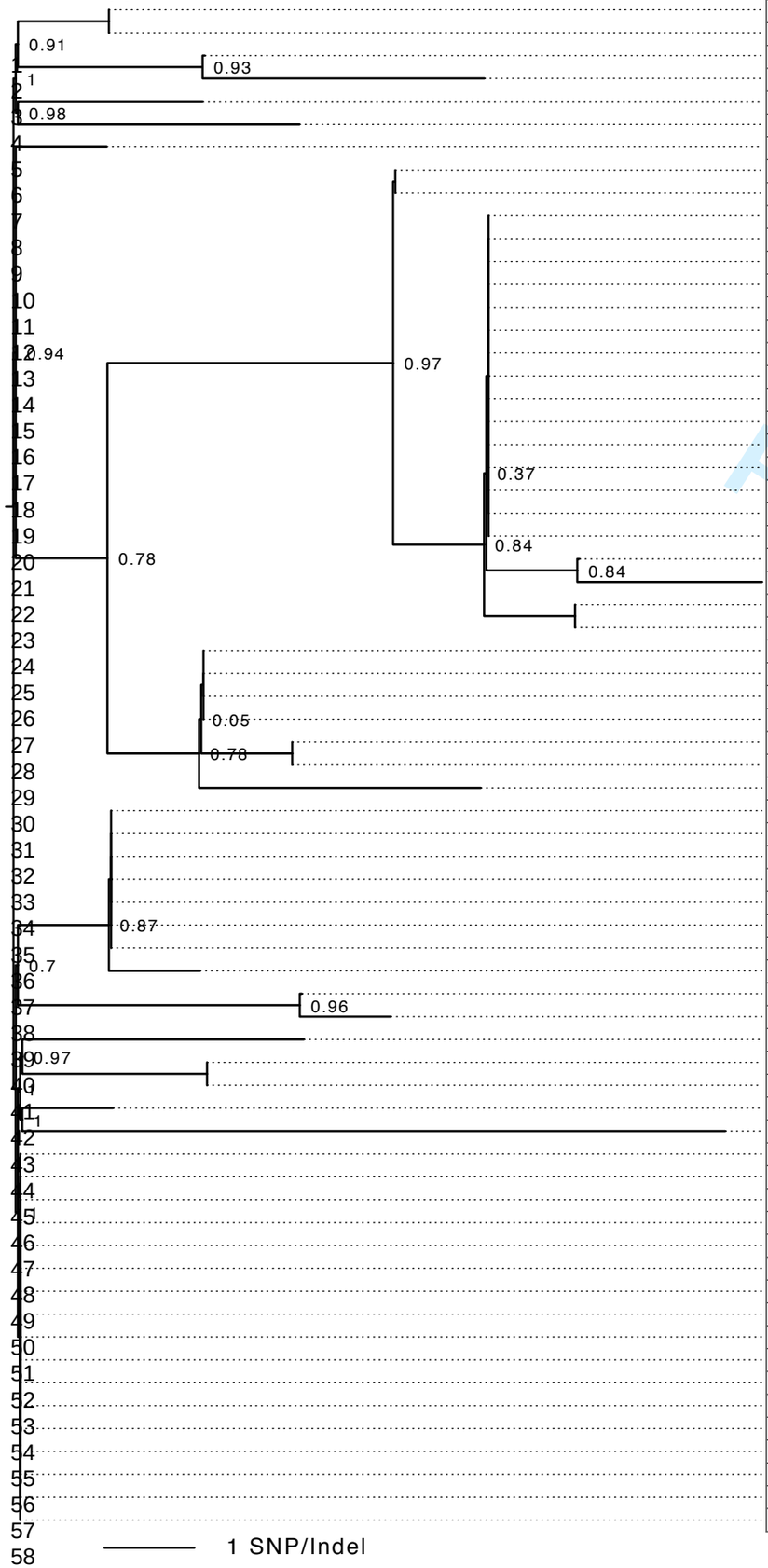
1 SNP/Indel Water Wound Environmental



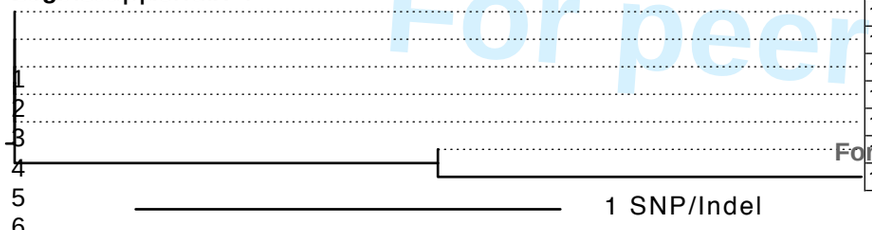
Sample Number	Patient	Ward	Bed	Type of Specimen	Site	Date	TIM	PIP	TAZ	CAZ	AZT	IMI	MER	AMI	GEN	TOB	CIP	COL
922	P02	BCU	8	Wound swab	Upper back	29/10/12	R	I	I	S	I	S	S	S	S	S	S	S
921	P02	BCU	8	Wound swab	Upper back	29/10/12												
913	P02	BCU	8	Environmental	Shower trolley	29/10/12	R	S	S	S	I	S	S	S	S	S	S	S
919	P02	BCU	8	Wound swab	Chest	29/10/12	R	I	I	S	I	I	S	S	S	S	R	S
926	P02	BCU	8	Wound swab	Right palm	02/11/12	R	I	I	S	I	I	S	S	S	S	R	S
909	P02	BCU	8	Wound swab	Anterior left upper-arm	26/10/12	R	I	I	S	I	R	S	S	S	S	R	S
908	P02	BCU	8	Wound swab	Anterior left upper-arm	26/10/12	R	I	I	S	I	S	S	S	S	S	S	S
925	P02	BCU	8	Wound swab	Anterior left upper-arm	02/11/12	R	I	I	S	I	I	S	S	S	S	R	S
932	P02	BCU	8	Wound swab	Back of head	07/11/12	R	I	I	S	I	R	S	S	S	S	R	S
928	P02	BCU	8	Wound swab	Upper back	02/11/12	R	I	I	S	I	I	I	S	S	S	R	S
927	P02	BCU	8	Wound swab	Upper back	02/11/12	R	I	I	S	I	I	I	S	S	S	R	S
915	P02	BCU	8	Environmental	Chair	29/10/12	R	I	I	S	I	S	S	S	S	S	S	S
914	P02	BCU	8	Environmental	Chair	29/10/12	R	S	S	S	I	S	S	S	S	S	S	S
904	P02	BCU	8	Tissue	Anterior right upper-arm	22/10/12	R	S	S	S	I	S	S	S	S	S	S	S
905	P02	BCU	8	Wound swab	Chest	26/10/12	R	I	I	S	I	I	S	S	S	S	S	S
906	P02	BCU	8	Wound swab	Chest	26/10/12	R	I	I	S	I	I	S	S	S	S	S	S
920	P02	BCU	8	Wound swab	Chest	29/10/12	R	I	I	S	I	R	I	S	S	S	R	S
918	P02	BCU	8	Wound swab	Chest	29/10/12	R	I	I	S	I	S	S	S	S	S	S	S
907	P02	BCU	8	Wound swab	Abdomen	26/10/12	R	I	I	S	I	S	S	S	S	S	S	S
911	P02	BCU	8	Wound swab	Anterior left forearm	26/10/12	R	I	I	S	I	I	S	S	S	S	S	S
912	P02	BCU	8	Wound swab	Anterior left forearm	26/10/12	R	I	I	S	I	I	S	S	S	S	S	S
923	P02	BCU	8	Wound swab	Posterior left upper-arm	29/10/12	R	I	I	S	I	S	S	S	S	S	S	S



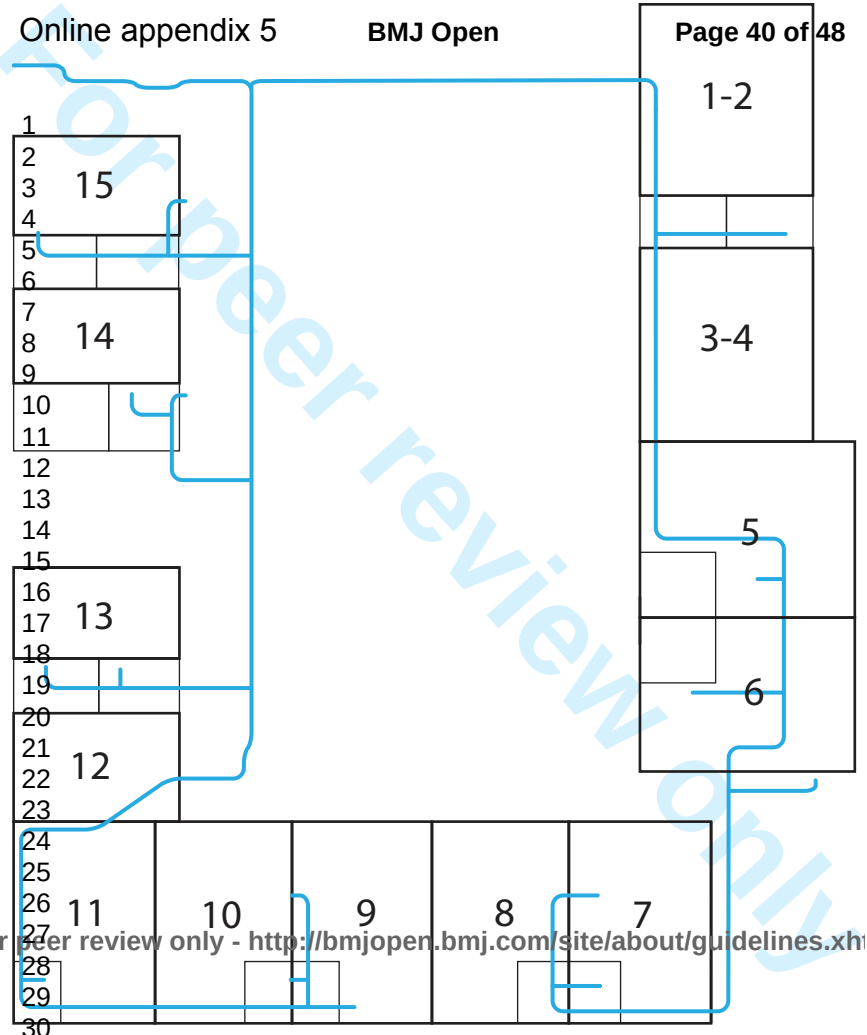
Study Number	Patient	Ward	Bed	Type of Specimen	Site	Date	TIM	PIP	TAZ	CAZ	AZT	IMI	MER	AMI	GEN	TOB	CIP	COL
937	P03	BCU	11	Wound swab	Anterior left thigh	08/11/12												
936	P03	BCU	11	Wound swab	Anterior left thigh	08/11/12												
929	P03	BCU	11	Wound swab	Posterior left thigh	06/11/12	R	S	S	S	I	R	I	S	S	S	S	S
951	P03	BCU	15	Environmental	Bedside table	14/11/12	R	I	I	S	I	R	I	S	S	S	S	S
960	P03	BCU	15	Environmental	Bedside table	20/11/12	R	S	S	S	I	R	I	S	S	S	S	S
961	P03	BCU	15	Environmental	Door handle	20/11/12	R	I	I	S	I	R	I	S	S	S	S	S
958	P03	BCU	15	Environmental	Shower chair	20/11/12	R	I	I	S	I	R	I	S	S	S	S	S
957	P03	BCU	15	Environmental	Shower chair	20/11/12	R	S	S	S	I	R	I	S	S	S	S	S
940	P03	BCU	15	Wound swab	Anterior right thigh	13/11/12	S	S	S	S	I	I	I	S	S	S	S	S
970	P03	BCU	15	Wound swab	Anterior right thigh	22/11/12	R	I	I	S	I	R	I	S	S	S	S	S
944	P03	BCU	15	Wound swab	Anterior right shin	13/11/12	R	I	I	S	I	R	I	S	S	S	S	S
945	P03	BCU	15	Wound swab	Anterior right shin	13/11/12	R	S	S	S	I	R	I	S	S	S	S	S
946	P03	BCU	15	Wound swab	Anterior left thigh	13/11/12	R	S	S	S	I	R	I	S	S	S	S	S
954	P03	BCU	15	Wound swab	Anterior left thigh	15/11/12	R	I	I	S	I	R	I	S	S	S	S	S
947	P03	BCU	15	Wound swab	Anterior left shin	13/11/12	R	S	S	S	I	R	I	S	S	S	S	S
972	P03	BCU	15	Wound swab	Posterior left shin	22/11/12	S	S	S	S	I	R	I	S	S	S	S	S
949	P03	BCU	15	Wound swab	Posterior right thigh	13/11/12	R	I	I	S	I	R	I	S	S	S	S	S
950	P03	BCU	15	Environmental	Bedside table	14/11/12	R	S	S	S	I	R	I	S	S	S	S	S
959	P03	BCU	15	Environmental	Bedside table	20/11/12	R	I	I	S	I	R	I	S	S	S	S	S
953	P03	BCU	15	Environmental	Toilet flush	14/11/12	R	S	S	S	I	I	I	S	S	S	S	S
938	P03	BCU	15	Urine	Urine	13/11/12	R	S	S	S	I	R	I	S	S	S	S	S
939	P03	BCU	15	Urine	Urine	13/11/12	R	S	S	S	I	R	I	S	S	S	S	S
956	P03	BCU	15	Urine	Urine	16/11/12	R	S	S	S	I	R	I	S	S	S	S	S
969	P03	BCU	15	Wound swab	Anterior right thigh	22/11/12	R	S	S	S	I	R	I	S	S	S	S	S
971	P03	BCU	15	Wound swab	Anterior left thigh	22/11/12	R	I	I	S	I	R	I	S	S	S	S	S
973	P03	BCU	15	Wound swab	Posterior right thigh	22/11/12	R	S	S	S	I	R	I	S	S	S	S	S
955	P03	BCU	15	Wound swab	Anterior left thigh	15/11/12	S	S	S	S	I	R	I	S	S	S	S	S
975	P03	BCU	15	Wound swab	Posterior right shin	22/11/12	S	S	S	S	I	R	I	S	S	S	S	S
962	P03	BCU	15	Environmental	Shower chair	20/11/12	S	S	S	S	I	R	I	S	S	S	S	S
948	P03	BCU	15	Wound swab	Posterior right thigh	13/11/12	S	S	S	S	I	R	I	S	S	S	S	S
974	P03	BCU	15	Wound swab	Posterior right thigh	22/11/12	S	S	S	S	I	I	I	S	S	S	S	S
963	P03	BCU	15	Environmental	Bedside table	20/11/12	S	S	S	S	I	I	I	S	S	S	S	S
964	P03	BCU	15	Environmental	Bedside table	20/11/12	S	S	S	S	I	R	I	S	S	S	S	S

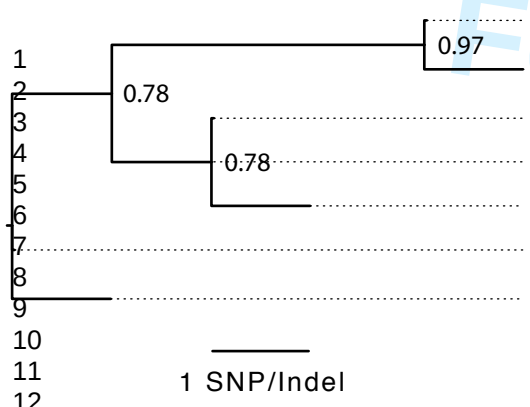


Sample Number	Patient	Ward	Bed	Type of Specimen	Site	Date	TIM	PIP	TAZ	CAZ	AZT	IMI	MER	AMI	GEN	TOB	CIP	COL	pBURNS1	pBURNS2
1067	SP30	BCU	12	Water	Shower (Pre-flush)	08/07/13														
1068	SP30	BCU	12	Water	Shower (Pre-flush)	08/07/13														
991	Water sampling	BCU	10	Water	Shower (Pre-flush)	18/12/12														
966	SP08	BCU	10	Water	Shower (Unknown)	23/11/12	R	S	S	S	I	S	S	S	S	S	S	S		
931	P02	BCU	8	Environmental	Shower (Rose)	07/11/12	R	S	S	S	I	S	S	S	S	S	S	S		
1065	SP30	BCU	12	Environmental	Shower (Rose)	08/07/13														
1034	SP15	BCU	8	Environmental	Shower (Rose)	14/03/13	R	S	S	S	R	S	S	S	S	S	S	S	Yes	
902	P01	BCU	11	Environmental	Drain	16/10/12														
903	P01	BCU	11	Water	Shower (Post-flush)	16/10/12	R	S	S	S	I	S	S	S	S	S	S	S		
981	P04	BCU	11	Environmental	Drain	17/12/12	R	S	S	S	I	S	S	S	S	S	S	S		
980	P04	BCU	11	Environmental	Shower (Rose)	17/12/12	R	S	S	S	I	S	S	S	S	S	S	S		
1048	SP20	BCU	11	Environmental	Shower (Rose)	23/04/13	R	S	S	S	I	S	S	S	S	S	S	S		
901	P01	BCU	11	Environmental	Trolley	16/10/12	R	S	S	S	I	S	S	S	S	S	S	S		
987	P04	BCU	11	Water	Shower (Hose)	17/12/12	R	S	S	S	I	S	S	S	S	S	S	S		
1050	SP21	BCU	11	Water	Shower (Post-flush)	23/04/13														
988	P04	BCU	11	Water	Shower (Pre-flush)	17/12/12	R	S	S	S	I	S	S	S	S	S	S	S		
968	SP10	BCU	11	Water	Shower (Unknown)	23/11/12	R	S	S	S	I	S	S	S	S	S	S	S		
900	P01	BCU	11	Wound swab	Groin	03/10/12	R	S	S	S	I	S	S	S	S	S	S	S		
979	P04	BCU	11	Wound swab	Anterior right thigh	11/11/12	R	S	S	S	I	S	S	S	S	S	S	S		
1008	P04	BCU	11	Wound swab	Anterior right thigh	31/12/12	R	S	S	S	I	S	S	S	S	S	S	S		
976	P04	BCU	11	Wound swab	Anterior right shin	11/11/12	R	S	S	S	I	S	S	S	S	S	S	S		
984	P04	BCU	11	Wound swab	Dorsum of left foot	13/12/12	R	S	S	S	I	R	I	S	S	S	S	S		
983	P04	BCU	11	Wound swab	Dorsum of left foot	13/12/12	R	S	S	S	I	R	I	S	S	S	S	S		
982	P04	BCU	11	Wound swab	Dorsum of left foot	13/12/12	R	S	S	S	I	S	S	S	S	S	S	S		
1007	P04	BCU	11	Wound swab	Anterior right thigh	21/12/12	R	S	S	S	I	R	I	S	S	S	S	S		
1009	P04	BCU	11	Wound swab	Anterior right thigh	31/12/12	R	S	S	S	I	S	S	S	S	S	S	S		
1005	P04	BCU	11	Wound swab	Anterior right thigh	19/12/12	R	S	S	S	I	R	I	S	S	S	S	S		
1006	P04	BCU	11	Wound swab	Anterior right thigh	19/12/12	R	S	S	S	I	R	I	S	S	S	S	S		
1052	SP22	BCU	9	Water	Shower (Hose)	03/05/13														Yes
993	Water sampling	BCU	9	Water	Shower (Hose)	18/12/12	R	S	S	S	I	S	S	S	S	S	S	S		
994	Water sampling	BCU	9	Water	Shower (Pre-flush)	18/12/12	R	S	S	S	I	S	S	S	S	S	S	S		
992	Water sampling	BCU	9	Water	Tap	18/12/12	R	S	S	S	I	S	S	S	S	S	S	S		
985	P04	BCU	11	Water	Tap	17/12/12	R	S	S	S	I	S	S	S	S	S	S	S		
986	P04	BCU	11	Water	Tap	17/12/12														
967	SP09	BCU	9	Water	Shower (Unknown)	23/11/12	R	S	S	S	I	S	S	S	S	S	S	S		Yes
1062	SP29	BCU	6	Environmental	Shower (Rose)	08/07/13														
1058	SP25	BCU	6	Environmental	Shower (Rose)	24/06/13														
1003	Water sampling	BCU	6	Water	Shower (Hose)	19/12/12														
1064	SP29	BCU	6	Water	Shower (Pre-flush)	08/07/13														
1036	SP15	BCU	6	Water	Shower (Pre-flush)	15/03/13														
1004	Water sampling	BCU	6	Water	Shower (Pre-flush)	19/12/12														
1057	SP25	BCU	6	Water	Shower (Pre-flush)	24/06/13	R	S	S	S	I	S	S	S	S	S	S	S		
1041	SP16	BCU	6	Water	Shower (Pre-flush)	28/03/13														
999	Water sampling	BCU	1	Water	Shower (Hose)	19/12/12														
1000	Water sampling	BCU	1	Water	Shower (Pre-flush)	19/12/12														
1010	SP11	BCU	8	Environmental	Sink	10/01/13	S	S	S	S	I	S	S	S	S	S	S	S	Yes	
989	Water sampling	BCU	12	Water	Shower (Hose)	17/12/12	R	S	S	S	I	S	S	S	S	S	S	S		
990	Water sampling	BCU	12	Water	Shower (Pre-flush)	17/12/12	R	S	S	S	I	S	S	S	S	S	S	S		
1040	SP16	BCU	6	Environmental	Shower (Rose)	28/03/13	R	S	S	S	I	S	S	S	S	S	S	S		
1031	SP13	BCU	8	Water	Shower (Hose)	21/02/13													Yes	
1001	Water sampling	BCU	5	Water	Tap	19/12/12														
996	Water sampling	BCU	7	Water	Shower (Pre-flush)	18/12/12	R	S	S	S	I	S	S	S	S	S	S	S		
1029	SP12	BCU	8	Environmental	Drain	12/02/13													Yes	
933	P02	BCU	8	Environmental	EGC	07/11/12	R	S	S	S	I	S	S	S	S	S	S	S		
930	P02	BCU	8	Environmental	Shower (Rose)	07/11/12	R	S	S	S	I	R	S	S	S	S	S	S		
1028	SP12	BCU	8	Environmental	Shower (Rose)	12/02/13													Yes	
942	P02	BCU	8	Environmental	Shower (Rose)	13/11/12	R	S	S	S	I	S	S	S	S	S	S	S		
941	P02	BCU	8	Environmental	Shower (Rose)	13/11/12	R	S	S	S	I	S	S	S	S	S	S	S		
917	P02	BCU	8	Environmental	Shower (Rose)	29/10/12	S	S	S	S	I	S	S	S	S	S	S	S	Yes	
916	P02	BCU	8	Environmental	Shower (Rose)	29/10/12	S	S	S	S	I	S	S	S	S	S	S	S	Yes	
934	P02	BCU	8	Water	Shower (Post-flush)	07/11/12	S	S	S	S	I	S	S	S	S	S	S	S	Yes	
1030	SP12	BCU	8	Water	Shower (post-flush)	12/02/13													Yes	
1035	SP14	BCU	8	Water	Shower (Post-flush)	14/03/13													Yes	
910	P02	BCU	8	Water	Shower (Post-flush)	29/10/12	S	S	S	S	I	S	S	S	S	S	S	S	Yes	
935	P02	BCU	8	Water	Shower (Pre-flush)	07/11/12	R	S	S	S	I	S	S	S	S	S	S	S	Yes	
1046	SP19	BCU	8	Water	Shower (Pre-flush)	10/04/13														
943	P02	BCU	8	Water	Shower (Pre-flush)	13/11/12	S	S	S	S	I	S	S	S	S	S	S	S	Yes	



Study Number	Patient	Ward	Bed	Type of Specimen	Site	Date	TIM	PIP	TAZ	CAZ	AZT	IMI	MER	AMI	GEN	TOB	CIP	COL
1069	Water sampling	WCCB	11	Water	Tap	17/07/13	R	S	S	S	I	R	I	S	S	S	S	S
1045	SP20	WCCB	11	Environmental	Tap handle	11/04/13												
1047	SP20	WCCB	11	Water	Tap	11/04/13	R	S	S	S	I	R	I	I	R	S	S	S
1049	P05	WCCB	11	Sputum	Sputum	28/04/13		I	I	S	I	I	I	I	R	S	S	S
1053	P05	WCCB	11	Water	Tap	08/05/13	R	I	I	S	I	R	S	S	S	S	S	S
1054	P05	WCCB	11	Sputum	Sputum	06/05/13	R	R	R	R	R	R	R	I	R	S	I	S
1056	P05	WCCB	11	Sputum	Sputum	19/05/13		R	R	R	S	R	R		S	S	R	S



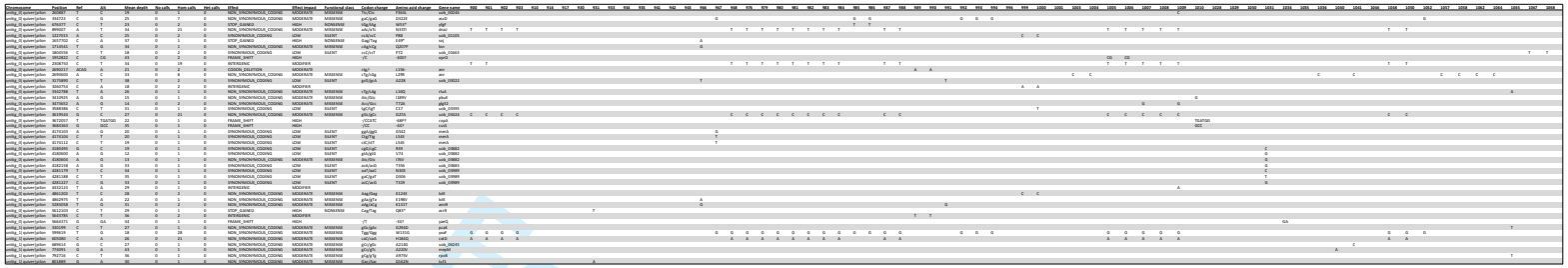


Bed	Type of Specimen	Alignment
11	Shower	TCCTACTCC-CACAGACCTAACCT-----G--ACCGAAATCCTTTTCCG-CGAGGCG
11	Shower	TCCTACTCC-TACAGACCTAACCT-----G--ACCGAAATCCTTTTCCG-CGAGGCG
9	Tap	TGCAACTCC-CACAGACCTAACGT-----G--ACCGAAATCCTTTTCCG-CGCGGCG
9	TMV	TGCA-CT-C-CACAGA---AACGT-----G--ACCG--A-CCTTTTC-G-C--GGCG
11	Tap	TGTAACTCC-CACAGACCTAACGT-----G--ACCGAAATCCTTTTCCG-CGCGGCG
8	Shower	TCCAACTCC-CACAGACCTAACGT-----G--ACCGAAATCCTTTTCCG-CTCGGCG
6	Shower	TCCAACTCC-CACAGCCTAACGT-----G--ACCGAAATCCTTTTCCG-CTCGGCG

Chromosome	Position	Ref	Alt	Mean depth	No calls	Hom calls	Het calls	Effect	Effect impact	Functional class	Codon change	Amino acid change	Gene name	904	905	906	907	908	909	911	912	913	914	915	918	919	920	921	922	923	925	926	927	928	932
NC_002516	1558800	CCATATG	C	40	0	1	0	CODON_DELETION	MODERATE		catatg/-	HM211-	lasR																C						
NC_002516	2806409	G	A	23	0	8	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	Cac/Tac	H321Y	PA2491					A	A							A					A	A	A	A	A
NC_002516	3804666	GCTTGC	G	19	0	1	0	FRAME_SHIFT	HIGH		-	-72	PA3399															G							
NC_002516	4148397	A	T	21	0	1	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	aTt/aAt	I181N	wspD									T													

Online appendix 8

Chromosome	Position	Ref	Alt	Mean depth	No calls	Hom calls	Het calls	Effect	Effect impact	Functional class	Codon change	Amino acid change	Gene name	929	936	937	938	939	940	944	945	946	947	948	949	950	951	953	954	955	956	957	958	959	960	961	962	963	964	969	970	971	972	973	974	975
NC_002516	574989	CCGCAGTTG	C	20	0	9	0	FRAME_SHIFT	HIGH			-37	PA0512				C	C								C		C			C			C					C		C		C			
NC_002516	856001	TGGCCTGG	T	31	0	1	0	FRAME_SHIFT	HIGH			-242	putP																									T			C					
NC_002516	1558520	C	T	41	0	5	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	cCg/cTg	P117L	lasR											T														T	T	T						T		
NC_002516	1558668	C	CTT	36	0	2	0	FRAME_SHIFT	HIGH			-/TT	-167?	lasR																CTT																CTT
NC_002516	2844173	A	C	19	0	26	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	gAc/gCc	D119A	PA2523	C	C	C	C	C	C	C	C	C	C			C	C	C	C	C		C	C	C	C	C	C			C	C	C	C	C			
NC_002516	2926290	G	A	25	0	7	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	cCg/cTg	P48L	gacA											A						A								A	A	A						A	A	
NC_002516	3486673	A	T	25	0	1	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	cTg/cAg	L171Q	PA3106																																			T
NC_002516	3584085	G	A	23	0	2	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	gCc/gTc	A87V	gltR		A	A																																
NC_002516	5796903	C	T	14	0	1	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	gCc/gTc	A317V	mutY																																T					
NC_002516	5853846	A	C	26	0	1	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	Acc/Ccc	T88P	PA5201																													C								



Chromosome	Position	Ref	Alt	Mean depth	No calls	Hom calls	Het calls	Effect	Effect impact	Functional class	Codon change	Amino acid change	Gene name	1045	1047	1049	1053	1054	1056	1069
NC_002516	3558951	G	A	35	0	1	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	aCc/aTc	T83I	gyrA						A	
NC_002516	4166773	A	G	22	0	2	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	Acg/Gcg	T86A	nalC					G	G	

Patient	Antibiotic/antifungal	Start	End	Start	End
2	Ciprofloxacin	16/10/2012	06/11/2012		
	Nitrofurantoin	16/10/2012	11/11/2012		
	Vancomycin	29/10/2012	03/11/2012		
4	Flucloxacillin	23/11/2012	30/11/2012		
	Piperacillin/tazobactam	30/11/2012	03/12/2012		
	Meropenem	03/12/2012	08/12/2012		
	Colistin	15/12/2012	21/12/2012		
5	Gentamycin	12/04/2013	12/04/2013		
	Co-amoxiclav	13/04/2013	14/04/2013		
	Erthromycin	14/04/2013	21/04/2013	11/05/2013	19/05/2013
	Piperacillin/tazobactam	14/04/2013	18/04/2013		
	Meropenem	20/04/2013	08/05/2013	19/05/2013	20/05/2013
	Caspofungin	26/04/2013	14/05/2013	20/05/2013	21/05/2013
	Linezolid	01/05/2013	12/05/2013		
	Ciprofloxacin	06/05/2013	16/05/2013	20/05/2013	21/05/2013
	Colistin	20/05/2013	23/05/2013		

Gene	Mutation type	Effect	AA substitution	Resistance phenotype	Samples
gyrA	SNP	non-synonymous	T83I	Ciprofloxacin	1056
nalC	indel	-	-	Meropenem	1054, 1056
mexS	SNP	non-synonymous	H321Y	Ciprofloxacin	908, 909, 919, 925-928, 932
oprD	indel	frame shift	-400?	Imipenem/meropenem	1005, 1006

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cohort studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	3
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	3-4
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	6
Objectives	3	State specific objectives, including any prespecified hypotheses	7
Methods			
Study design	4	Present key elements of study design early in the paper	8-9
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	8
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	8
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	n/a
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	8-9
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	8-9
Bias	9	Describe any efforts to address potential sources of bias	n/a
Study size	10	Explain how the study size was arrived at	9
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	n/a
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	??
		(*b*) Describe any methods used to examine subgroups and interactions	n/a
		(*c*) Explain how missing data were addressed	n/a
		(*d*) If applicable, explain how loss to follow-up was addressed	n/a
		(*e*) Describe any sensitivity analyses	n/a
Results			

Comment [MN1]: For JAMA it was in eMethods , however not sure where it is going to be now…

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	10
		(b) Give reasons for non-participation at each stage	10
		(c) Consider use of a flow diagram	-
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	10-11
		(b) Indicate number of participants with missing data for each variable of interest	10
		(c) Summarise follow-up time (eg, average and total amount)	n/a
Outcome data	15*	Report numbers of outcome events or summary measures over time	10-11
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	n/a
		(*b*) Report category boundaries when continuous variables were categorized	n/a
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	n/a
Discussion			
Key results	18	Summarise key results with reference to study objectives	16
Limitations			
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	16-18
Generalisability	21	Discuss the generalisability (external validity) of the study results	18
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	19

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Seeking the source of Pseudomonas aeruginosa infections in a recently opened hospital: an observational study using whole-genome sequencing

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-006278.R1
Article Type:	Research
Date Submitted by the Author:	16-Sep-2014
Complete List of Authors:	Quick, Josh; Institute of Microbiology and Infection, University of Birmingham; Queen Elizabeth Hospital, NIHR Surgical Reconstruction and Microbiology Research Centre Cumley, Nicola; Queen Elizabeth Hospital, NIHR Surgical Reconstruction and Microbiology Research Centre Wearn, Chris; University Hospital Birmingham Foundation Trust, Healing Foundation Centre for Burns Research; Queen Elizabeth Hospital, NIHR Surgical Reconstruction and Microbiology Research Centre Niebel, Marc; Queen Elizabeth Hospital, NIHR Surgical Reconstruction and Microbiology Research Centre Constantinidou, Chrystala; University of Warwick, Division of Microbiology and Immunology Thomas, Chris; University of Birmingham, Institute of Microbiology and Infection Pallen, Mark; University of Warwick, Division of Microbiology and Immunology Moiemen, Naiem; Queen Elizabeth Hospital, NIHR Surgical Reconstruction and Microbiology Research Centre; University Hospital Birmingham Foundation Trust, Healing Foundation Centre for Burns Research Bamford, Amy; Queen Elizabeth Hospital, NIHR Surgical Reconstruction and Microbiology Research Centre; University Hospital Birmingham Foundation Trust, Healing Foundation Centre for Burns Research Oppenheim, Beryl; Queen Elizabeth Hospital, NIHR Surgical Reconstruction and Microbiology Research Centre Loman, Nicholas; University of Birmingham, Institute of Microbiology and Infection
Primary Subject Heading:	Genetics and genomics
Secondary Subject Heading:	Infectious diseases, Pathology, Public health, Surgery
Keywords:	Microbiology < BASIC SCIENCES, BIOTECHNOLOGY & BIOINFORMATICS, Molecular diagnostics < INFECTIOUS DISEASES, Infection control < INFECTIOUS DISEASES, MICROBIOLOGY, Plastic & reconstructive surgery < SURGERY

Title

Seeking the source of *Pseudomonas aeruginosa* infections in a recently opened hospital: an observational study using whole-genome sequencing

Authors

Joshua Quick [1,2*], Nicola Cumley [2*], Christopher M. Wearn [2,3], Marc Niebel [2], Chrystala Constantinidou [4], Chris M Thomas [1], Mark J. Pallen [4], Naiem S. Moiemen [2,3], Amy Bamford [2,3], Beryl Oppenheim [2#], Nicholas J. Loman [1#]

Affiliations

[1]Institute of Microbiology and Infection, University of Birmingham, Birmingham, United Kingdom

[2]NIHR Surgical Reconstruction and Microbiology Research Centre, Queen Elizabeth Hospital, Birmingham, United Kingdom

[3]Healing Foundation Centre for Burns Research, University Hospital Birmingham Foundation Trust, Birmingham, United Kingdom

[4]Division of Microbiology and Immunology, University of Warwick, Warwick, United Kingdom

*Contributed equally

Correspondence:

Dr Nicholas James Loman

Institute of Microbiology and Infection, University of Birmingham, Birmingham, B15 2TT United Kingdom

Telephone: +44 (0) 121 414 8849

Email: n.j.loman@bham.ac.uk

Dr Beryl Oppenheim

Clinical Microbiology, University Hospitals Birmingham NHS Foundation Trust

Queen Elizabeth Hospital Birmingham, Mindelsohn Way, Edgbaston

Birmingham, B15 2WB

Telephone: +44 (0) 121 371 6523

Email: Beryl.Oppenheim@uhb.nhs.uk

Keywords

High-Throughput DNA sequencing; Pseudomonas aeruginosa; Burns, Hydrotherapy

Word Count

4539

ABSTRACT

Objectives

Pseudomonas aeruginosa is a common nosocomial pathogen responsible for significant morbidity and mortality internationally. Patients may become colonised or infected with *P. aeruginosa* after exposure to contaminated sources within the hospital environment. The aim of this study was to determine whether whole-genome sequencing (WGS) can be used to determine the source in a cohort of burns patients at high risk of *P. aeruginosa* acquisition.

Study design

An observational prospective cohort study.

Setting

Burns care ward and critical care ward in the United Kingdom.

Participants

Patients with >7% total burns by surface area were recruited into the study.

Methods

All patients were screened for *P. aeruginosa* on admission and samples taken from their immediate environment, including water. Screening patients who subsequently developed a positive *P. aeruginosa* microbiology result were subject to enhanced environmental surveillance. All isolates of *P. aeruginosa* were whole-genome sequenced. Sequence analysis looked at similarity and relatedness between isolates.

Results

Whole-genome sequences for 141 *P. aeruginosa* isolates were obtained from patients, hospital water and the ward environment. Phylogenetic analysis revealed eight distinct clades, with a single clade representing the majority of environmental isolates in the burns unit. Isolates from three patients had identical genotypes compared with water isolates from the same room. There was clear clustering of water isolates by room and outlet, allowing the source of acquisitions to be unambiguously identified. Whole-genome shotgun sequencing of biofilm DNA extracted from a thermostatic mixer valve revealed this was the source of a *P. aeruginosa* subpopulation previously detected in water. In the remaining two cases there was no clear link to the hospital environment.

Conclusions

This study reveals that WGS can be used for source tracking of *P. aeruginosa* in a hospital setting, and that acquisitions can be traced to a specific source within a hospital ward.

ARTICLE SUMMARY

Strengths and limitations of the study

- We have demonstrated that whole-genome sequencing can be used for source tracking of *P. aeruginosa* in a hospital setting.
- We show convincing evidence that transmission has occurred directly from water to patients, but other routes are as likely.
- The main limitation of the study was the sample size, which could be attributable to interventions being carried out during the study.
- Our study focused on a burns unit and critical care unit in a newly built hospital. Modes of *P. aeruginosa* transmission may be different in hospitals with different styles of plumbing and on other augmented care units.

INTRODUCTION

Pseudomonas aeruginosa is a ubiquitous Gram-negative bacterium and an important opportunistic pathogen in the healthcare setting. *P. aeruginosa* particularly affects those with impaired host or mucosal immunity and has a broad range of presentations including respiratory infections in cystic fibrosis and mechanically ventilated patients, bloodstream infections in premature neonates and wounds in burns injuries. Nosocomial *P. aeruginosa* outbreaks are frequently reported and associated with water sources such as taps, showers, mixer valves and flow straighteners, sink traps and drains.[1-10] Other potential routes of transmission include cross-infection, for example carriage on the hands of health care workers, and through contaminated medical equipment such as endoscopic devices.[3,5]

In the UK, the role of water in the transmission of *P. aeruginosa* in healthcare settings has become a matter of urgent concern in response to a recent high-profile outbreak affecting a neonatal critical care unit in Belfast in 2012.[11] This source was eventually determined to be sink taps.[11-13] National guidance is now in place detailing enhanced procedures for routine water sampling on augmented care units, with directed interventions such as disinfection and replacement of high-risk plumbing parts required.[14]

Historical phenotypic typing methods for *P. aeruginosa* such as O-antigen serotyping have more recently been replaced by molecular typing methods such as pulsed-field gel electrophoresis (PFGE), variable number tandem repeat analysis (VNTR) and random amplification of polymorphic DNA (RAPD) and multi-locus sequencing typing (MLST).[15] These methods have been used to investigate outbreaks of *P. aeruginosa* within hospitals.[4,16-18] However, such techniques have important

limitations for source tracking of infections in hospitals as they sample limited numbers of sites in the genome which may result in false clustering of unrelated strains.[19] In the past five years, whole-genome sequencing (WGS) has started to be used to investigate outbreaks in hospitals. WGS is attractive because of its digital, sharable format and ultra-high resolution, which is able to discriminate two isolates differing by just a single mutation. WGS has been successfully used to determine likely transmission chains during outbreaks of *Staphylococcus aureus*, *Acinetobacter baumannii* and *Klebsiella pneumoniae*.[19-21] Benchtop sequencing instruments now offer a cost-effective approach for bringing bacterial WGS to the clinical environment. [1,22]

In this study, we explore the utility of WGS to determine the likely sources of *P. aeruginosa* in an at-risk population of burns patients. In the UK and US burns patients receive shower cart hydrotherapy as a mainstay of burns treatment.[23-26] A previous hospital audit suggested that up to one-third of such patients became colonised with *P. aeruginosa*. We hypothesised that this high rate of acquisition may relate to transmission from hospital shower water during therapy. We therefore wished to understand the importance of transmission from water compared with alternative routes such as cross-infection and endogenous carriage.

MATERIALS AND METHODS

Hospital setting

An observational, prospective study design was employed in a burns centre serving approximately 13.7 million people across the Midlands region of England with 300 admissions annually. Opened in June 2010, the burns centre comprises a purpose built 15-bed ward with 11 side-rooms and 2 dual-bedded rooms. Patients requiring mechanical ventilation and organ support are usually treated in two self-contained burns cubicles located within the trauma critical care unit (CCU). Despite the observational nature of the study, sampling was carried out during implementation of interim national guidance on control of *P. aeruginosa* issued by the Department of Health. These guidelines were issued in draft form March 2012, and subsequently revised in March 2013. This meant that parallel water sampling and engineering interventions were being undertaken during the period of study. In addition, some enhanced infection prevention measures were also introduced in response to an outbreak of a multi-drug resistant *A. baumannii.*

Study design and patient selection

Ethical approval was sought and received from the National Research Ethics Service committee in the West Midlands (reference number 12/WM/0181). Patients admitted to the burns unit were eligible for the screening phase of the study if they had burns injuries covering greater than 7% total body surface area (TBSA). Patients were screened as soon as possible after admission after they had given written informed consent. When appropriate, legal consultee advice was sought for patients lacking capacity due to emergency treatment. On admission, recruited patients were screened for carriage of *P. aeruginosa* (wounds, urine and stool) using standard

microbiology techniques. Samples were then taken as part of routine microbiology service during the patients stay. Environmental and water samples were taken after the patient was admitted to the burns centre. If during the period of stay *P. aeruginosa* was isolated from a patient sample the patient was recruited into the second phase of the study. In this phase, patients had wound swabs taken at each dressing change as well as twice-weekly urine samples. The patient's environment and water from outlets in their bed space were sampled weekly for the duration of their stay, and after discharge (post-cleaning). Termination of the study was planned after 30 screening patient admissions, or a year, whichever came soonest, after which 10 patients were expected to acquire *P. aeruginosa*. This prediction was based on a previous local audit which suggested about one-thirds of burns patients became colonised with *P. aeruginosa*..

Microbiological and molecular methods

P. aeruginosa isolates were obtained from wound swab, urine, stool, environmental and water samples. *P. aeruginosa* was isolated from wound swabs, urine and stool by inoculation onto cysteine lactose electrolyte deficient agar (CLED) and cetrimide agar and incubation for 24 hours at 37°C. Stool samples were cultured overnight in a cetrimide enrichment broth before subculture onto CLED. Identification was confirmed by resistance to C-390 and the VITEK® 2 GN identification card. Antibiotic sensitivity assays were performed using the VITEK® 2 AST N-210 card (bioMérieux, Basingstoke, UK).

The patient's environment (shower head rosette, drain, shower chair or trolley, bedside table, patient chair, instruments in contact with the patient) was sampled over a 10cm^2 area by a Polywipe™ sponge. The sponge was placed in tryptic soy

broth incubated for 24 hours at 37°C then sub-cultured onto CLED and cetrimide agar. During water sampling, water was taken from the patient's shower, or tap if a shower was not present. Shower heads were not removed for water sampling. At least 200ml of water was collected into a vessel containing sodium thiosulphate as a neutraliser. In duplicate, 100ml of water was filtered through a 0.45 micron filter and the filters placed onto CLED plates and cetrimide agar. Plates were incubated at 37°C for 48 hours and the number of organisms per 100ml quantified.

For storage and DNA extraction a single colony was purified from the primary culture plate. When different colony morphologies were observed, a single colony from each type was purified. Additionally, for a randomly selected water sample, 24 colonies were individually picked from one water-filter primary microbiological plate for sequencing. Isolates were stored on Biobank beads at -20°C prior to DNA extraction. Organisms were resuscitated on CLED agar plates and genome DNA either extracted directly using the MOBIO UltraClean Microbial DNA Kit, or from overnight LB broth culture using a Qiagen Genomic-Tip 100G.

DNA extraction and sequencing

Genomic DNA was prepared from single colony picks using the MIOBIO Ultraclean microbial kit (MOBIO, Carlsbad, USA). 1ng input DNA, as quantified by Qubit (Life Technologies, Carlsbad, USA) was used to prepare genomic libraries for sequencing using the Illumina Nextera XT™ DNA sample kit as per manufacturer's protocol (Illumina, San Diego, USA). Libraries were sequenced on the Illumina MiSeq using a paired-end protocol resulting in read lengths between 150 and 300 bases. A single additional sample, isolate 910, was chosen as a representative member of Clade 5 for long-read sequencing. DNA from this sample was fragmented using a

Hydroshear (Digilab, Marlborough, MA) using the recommended protocol for 10kb fragments and further size-selected on a Blue Pippen instrument (Sage Science, MA) with a 7kb minimum size cut-off. The library was sequenced on two SMRT Cells using the Pacific Biosciences RS II instrument at the Norwegian Sequencing Centre, Oslo. C4-P2 chemistry was chosen because it favours long, more accurate reads for *de novo* assembly.

Stool PCR

For simple presence/absence detection of *P. aeruginosa* in stool samples using PCR, a stool sample was collected into a stool collection tube containing stool DNA stabilizer. Total DNA was extracted using the PSP Spin Stool DNA Plus kit (Stratec Molecular). PCR amplification of species specific regions of the 16S rDNA gene was carried out using primers PA-SS-F: GGGGGATCTTCGGACCTCA and PA-SS-R: TCCTTAGAGTGCCCACCCG [12] in the following conditions: 0.5µM of each primer, 1.5mM $MgCl_2$, 0.2mM dNTP's using BIOTAQ™ DNA Polymerase and buffer set. After initial denaturation at 96°C for 2 minutes, 30 cycles of 96°C for 30 seconds, 62° C for 30 seconds and 72°C for 30 seconds were completed with a final extension of 72°C for 5 minutes. Products were visualised for size on an 1.5% agarose gel.

Bioinformatics methods

Illumina MiSeq reads from each isolate were adapter and quality trimmed before use with Trimmomatic.[27] Phylogenetic reconstruction of isolates sequenced in this study were combined with data from a global collection of 55 *P. aeruginosa* strains collected world-wide which have been previously analysed by Stewart et al. [28] For each of the published strains, 600,000 paired-end reads of length 250 bases were simulated using wgsim (https://github.com/lh3/wgsim) from the complete or draft

genome assembly deposited in Genbank. Read sets were mapped against the *P. aeruginosa* PAO1 reference genome using BWA-MEM 0.7.5a-r405 using default settings.[29] Single nucleotide polymorphisms were called using VarScan 2.3.6 and filtered for regions with an excessive number of variants. These may represent regions of recombination, misalignments or strong Darwinian selection.[30] FastTree (version 2.1.7) was used for phylogenetic reconstruction. This software estimates an approximate maximum-likelihood tree under the Jukes-Cantor model of nucleotide evolution with a single rate for each site (CAT).[31] Trees were drawn in FigTree (http://tree.bio.ed.ac.uk/software/figtree/).

For *in silico* MLST prediction, trimmed reads were assembled *de novo* using Velvet [32] with a k-mer size of 81 and searched using nucleotide BLAST against the multi-locus sequence database downloaded from the pubMLST website on 2013-08-05 (http://pubmlst.org/paeruginosa/) [33]. For Clade E isolates, in order to exhaustively search for discriminatory mutations, a nearly complete reference genome was generated by *de novo* assembly using Pacific Biosciences sequencing data. Reads were assembled using the 'RS_HGAP_Assembly.3' pipeline within SMRT Portal v2.2.0. Illumina reads from the same sample were mapped to this draft genome assembly in order to correct remaining indel errors in the assembly using Pilon (http://www.broadinstitute.org/software/pilon/). Isolates belonging to each clade were mapped individually against either the PacBio reference (Clade E) or *P. aeruginosa* PAO1 (NC_002516) (Clades C, D and G).

Variants (single nucleotide polymorphisms and short insertion-deletions) were called using SAMtools mpileup and VarScan with an allele frequency threshold of 80%.**[30]** Non-informative positions and regions of putative recombination were removed, the

later with a variant density filter of more than 3 SNPs every 1000 nucleotides. Analysing samples in each clade individually maximised the number of variants detected by reducing the likelihood of the position being uncovered by a subset of samples. From these variants fine-grained phylogenetic trees were reconstructed for each clade using FastTree. The scripts used to perform this analysis are available at http://www.github.com/joshquick/snp_calling_scripts. Approximate-maximum-likelihood phylogenetic trees were generated using FastTree and visualised in FigTree. For whole-genome shotgun metagenomics analysis, reads were analysed using the Kraken taxonomic classifier software with the supplied *minikraken* database.**[34]** Reads from the metagenomics dataset were aligned to *P. aeruginosa* PAO1 as in the previous section and phylogenetic placement was carried out using pplacer in conjunction with FastTree.**[35]** Sequence data is available from the European Nucleotide Archive for the Illumina data (ERP006056) and the corrected Pacific Biosciences assembly (ERP006058).

RESULTS

Study results

Recruitment lasted a period of 300 days, ending according to protocol after the enrolment of 30 screening patients. In total, we detected *P. aeruginosa* in five patients. Of these patients, three had *P. aeruginosa* detected only in burns wound swabs, one had *P. aeruginosa* detected in both their burns wound and in their urine, and one had *P. aeruginosa* in their sputum. One additional eligible patient did not consent to enter the study and was excluded. The average age in the study group was 41 years. Males predominated with a male-to-female ration of 2.3:1. Flame burns were the most common mechanism of injury, followed by scalds and mixed flame/flash injuries. The average burn size of the study group was 12.5% of the total

body surface area (TBSA) and 27% of patients sustained an inhalation injury. Eight patients required admission to ITU and the majority required surgical treatment of their burns with excision and skin grafting (80%). A large majority of the study group (83%) received shower cart hydrotherapy as a routine part of their wound management to encourage healing through wound debridement and decontamination. The average length of hospital stay (LOS) was 17 days and taking into account burn size, the average was 1.4 days per % TBSA.

The water and environment in burns and critical care units are frequently colonised by *P. aeruginosa*

A total of 282 water and environmental samples were screened for *P. aeruginosa* of which 39/78 (50%) were positive in water samples, 25/96 (26%) were positive from the wet environment and 7/108 (6%) were positive from the dry environment. A total of 86 genome sequences were generated from the 71 positives, as in some cases multiple colony picks were sequenced. 78 patient samples were screened for *P. aeruginosa* of which 39 (50%) were positive. A total of 55 genome sequences were generated, as in some cases multiple colony picks were sequenced. In total, 141 genomes were sequenced; water and environmental (n=86) and patient (n=55). Genomes were sequenced to a mean coverage of 24.4x, with the minimum coverage of a sample being 14x and highest 64.7x.

When placed in the context of a global collection of *P. aeruginosa* strains, phylogenetic reconstruction demonstrated isolates in our study fell into eight clades (figure 1, Panel A). As has been reported previously, there was no strong association between ecological context and position in the phylogenetic tree.[28] Isolates in this study are most closely related to strains from a variety of settings. The majority of

isolates (52%) belong to Clade E (figure 1 Panel B), whose nearest sequenced relative is the Liverpool Epidemic Strain, a clone often isolated from patients in the UK and Canada with cystic fibrosis. [36,37] Isolates from Clade E were found in the burns unit's water and the ward environment, as well as from two patient's wounds. However it was never detected in the critical care unit. Clade E was detected throughout the study in a total of 10 different rooms (figure 2).

Inferring potential transmission events by whole-genome sequencing

Microevolutionary changes occurring over rapid time-scales (i.e. days to months) have been used to detect potential chains of transmission in hospital and community outbreaks.[19-21,38,39] The number of distinct mutations between given isolates has been used to infer whether transmission events are likely to have occurred. Such inferences are aided by prior knowledge of mutation rates in similar populations. Two patients (1 and 4) in our study both had *P. aeruginosa* from Clade E isolated from their wounds. These isolates had an indistinguishable genotype from those present in water and the environment of the room they were nursed within (figure 1 Panel C and figure 3). This genotype was detected in the patient's shower water after initial patient screening, during screening of the second patient admission, twice during the second patient's stay and then 127 days later (days 27, 65, 89 and 216 respectively). When water isolates were positive, the genotype was also detected in wet environment sites (shower drain, shower rosette and patient's trolley) on the same days.

Patient 5 was nursed on the critical care unit due to concomitant medical problems. *P. aeruginosa* belonging to Clade G was isolated from sputum during this time.

Identical genotypes were detected contemporaneously in the water from the associated sink and sink tap handle (see online supplementary appendix 4).

Two further patients (patients 2 and 3) were positive for *P. aeruginosa*. Isolates from these patients belonged to Clade C and D respectively. Neither clade was ever isolated from hospital water. In both cases, identical genotypes were detectable in the environment associated with the patient but these were not detected before or after the patients' stay, indicating that the environment was not persistently contaminated. During the course of patient 3's stay, the dry environment such as the bedside table was contaminated, as was the patient's door handle and shower chair. However, after patient discharge, the strain associated with this patient was never seen again during the course of the study in any location.

Whole-genome sequencing permits source tracking of *P. aeruginosa* to individual water outlets

Whole-genome sequencing has been reported previously for source tracking, but never for the detection of transmission events from hospital water.[40] Phylogenetic reconstruction within Clade E, the most commonly detected water clone demonstrated additional diversity within this clone, with a total of 46 mutations detected an average genetic distance between isolates of 4.1 mutations (figure 3). The reconstruction demonstrated clear evidence of clustering of genotypes both by room and outlet (figure 3). When *P. aeruginosa* was detected in the wet environment (e.g. shower rosettes and drains) these genotypes were most often identical to those found in water, indicating that the water was likely the ultimate source of that clone. Genotypic variation was seen between outlets within the same room. For example, tap water sampled from room 11 had a distinct genotype from that sampled from

shower water in the same room and this was consistently found over multiple samplings. Notably, isolates from two patients fell within the cluster originating from shower water, indicating that shower hydrotherapy was the most likely source of infection. Two plasmids (designated pBURNS1 and pBURNS2) were detected in this study set, which both demonstrated geographical clustering, with pBURNS1 only being detectable in isolates from room 8 and pBURNS2 only being detectable in isolates from the shower water in room 9.

Rapid evolution of antibiotic resistance associated with treatment

P. aeruginosa is commonly associated with antibiotic resistance due to a number of predisposing features including intrinsic resistance, a repertoire of efflux pumps and antibiotic-inactivating enzymes including beta-lactamases. [41] Three infected patients (2, 3 and 5) received antibiotic therapy, and in each case this was associated with the development of resistance to at least one therapeutic agent. Associated mutations were detected that were either partially or fully explanatory of the phenotype (online supplementary appendix 12).

Patient 2 was treated with ciprofloxacin, nitrofurantoin and vancomycin (see online supplementary appendix 11 for full details). 8/21 (38%) tested isolates from this patient were ciprofloxacin resistant. 7/8 isolates (88%) of the ciprofloxacin-resistant strains were distinguishable from the other isolates by a single SNP in *mexS* (annotated as PA2491 in *P. aeruginosa* PAO1) (see online supplementary appendix 1 and 7). This SNP was predicted to result in a non-synonymous amino acid substitution. Disruption of this gene has been shown to cause increased expression of the mexEF-oprN multidrug efflux pump, associated with resistance to quinolones.[42]

Patient 3 was not treated with antibiotics, but isolates associated with this patient demonstrated differences in resistance to timentin and piperacillin-tazobactam. These changes were associated with non-synonymous mutations in *gacA*, the response regulator of the GacA/GacS two-component system and in *lasR*, a transcriptional activator required for transcription of elastase and LasA protease (online supplementary appendices 2 and 8).

Patient 4 was treated with meropenem, piperacillin/tazobactam, flucloxacillin and colistin. Five isolates collected 10 to 18 days after initiation of meropenem showed resistance to imipenem and intermediate resistance to meropenem (see online supplementary appendix 3 and 9). The most likely mutation responsible for this phenotype was detectable in two isolates, both of which had a frame-shift mutation in the gene coding for the membrane porin *oprD.* [43]

Patient 5 had a prolonged stay in ITU and had multiple medical problems including *A. baumannii* infection and was treated with nine antibiotic agents including ciprofloxacin, meropenem and piperacillin-tazobactam. Serial isolates from this patient demonstrated the stepwise acquisition of two mutations (online supplementary appendix 4). The first was in *nalC*, a probable repressor of the TetR/AcrR family (online supplementary appendix 10). [44] On inspection of the sequence alignment in this region, a large deletion of 196 nucleotide bases was seen compared to the reference PAO1 strain. This mutation was seen in association with full resistance to piperacillin-tazobactam, ceftazidine, aztreonam, meropenem and intermediate resistance to ciprofloxacin. This deletion is likely to result in over-expression of efflux pumps involving the *mexAB-oprM* operon. [44,45] Ciprofloxacin resistance in a later isolate corresponded to the stepwise acquisition of a second

mutation. This mutation is predicted to affect the well-studied DNA gyrase subunit A gene (*gyrA)* which is strongly associated with ciprofloxacin resistance.[46]

Confirmation of *P. aeruginosa* genotypes in biofilms by whole-genome metagenomic shotgun sequencing

P. aeruginosa is able to produce and survive in biofilms. Plumbing parts such as flow straighteners, shower rosettes, flexible hoses, solenoid valves and thermostatic mixer valves (TMV) are particularly at risk of biofilm formation due to factors including surface areas, convoluted designs and inadequate pasteurisation. [47] To confirm the presence of *P. aeruginosa* in water fittings associated with rooms on the burns unit, we obtained a TMV removed by the hospital estates team from the shower in room nine as part of compliance with UK guidelines for managing *P. aeruginosa* in hospitals. On visual inspection, a biofilm was present which was scraped from the surface with a sterile scalpel. DNA from this biofilm was extracted for whole-genome shotgun sequencing. The majority of reads did not map to any known bacterial taxa. The most abundant taxon identified was *Pseudomonas aeruginosa* (3%). Subsequent alignment to the *P. aeruginosa* PAO1 reference covered 94% of the 6.3 million base reference genome at a median coverage of 5x, confirming that reads were correctly classified to this species and not other environmental *Pseudomonas* species. Alignment to the *P. aeruginosa* Clade E reference genome followed by phylogenetic placement of reads demonstrated that it fell into the same clade as previously recovered isolates from the shower or tap in room 9 (indicated on figure 3, and in online supplementary appendix 6).

DISCUSSION

The hospital environment has been intimately linked with *P. aeruginosa* infection for over 50 years yet hospital acquisitions, clusters and outbreaks remain a common occurrence and understanding precise routes of transmission can be difficult. [47,48] Our results demonstrate that, even in a new hospital, *P. aeruginosa* can become rapidly endemic in hospital plumbing. Furthermore, by linking *P. aeruginosa* genotypes recovered from patients to specific individual water outlets, we offer compelling evidence of unidirectional transmission from water to patients. Further, by sequencing of a biofilm identified in a TMV from a hospital water system, we can identify the likely common source of genotypes found in water and in the hospital environment.

Our results suggest that use of whole-genome sequencing can reduce ambiguity about potential transmission events in hospitals and consequently inform infection prevention efforts about the direction and sequence of transmission. Typing schemes such as MLST and PFGE are much lower resolution methods and would not be able to provide sufficient information to permit such inferences to be made. It is notable that the burns unit was colonised by a single clone, meaning that it was very unlikely that water outlets at each bed space were colonised as a result of transmissions from the patient or environment. For this to happen would require multiple transmission events from separate patients with the same clone, for which there is not evidence. Instead we speculate that this clone was introduced to the hospital associated with its commissioning. One hypothesis is that particular plumbing fittings, i.e. the TMV may have been colonised simultaneously by a clone circulating in water. Clade E (ST395) has been frequently reported associated with water, so this remains a possibility. [49,50] However, it is possible that plumbing fittings are installed ‘pre-seeded’ with *P. aeruginosa* as has already been proposed

by Kelsey.[3,5,47] Investigation of an outbreak in Wales implicated new plumbing parts as a potential source of *P. aeruginosa.* New plumbing components are often tested by companies prior to their supply and it is possible they were contaminated prior to distribution. The limited amount of diversity (average 4 SNPs) seen within this clade is consistent with a single founding genotype coinciding with the opening of the burns unit, based on estimates from a previous study using WGS which reported that mutations accumulate at a rate of approximately one every 3-4 months in a hospital-associated clone.[51] However our results suggest that our isolates accumulate mutations even more slowly. This may be due to reduced growth rates in nutritionally-poor biofilms.[52]

It is notable that antibiotic resistance to multiple first-line agents developed rapidly in response to therapy. These results underline the importance of selecting appropriate antibiotic therapy in *P. aeruginosa* infections. It is reassuring however that antibiotic resistance genotypes selected *in vivo* did not show evidence of persistence in the ward environment or transmission to other patients.

Our study has certain limitations. Based on a previous audit, we expected around one-third of patients screened for *P. aeruginosa* would develop colonisation or clinical infection. In fact, only 5 out of 30 of patients were colonised. This may have been related to guidance and engineering interventions being put in place during the study as detailed in national guidance issued whilst this study was on-going. In addition, infection control policies were revised to address control of an outbreak of a multi-drug resistant *A. baumannii* in this same burns unit. Following these interventions, only 1 of the last 20 patients recruited was infected with *P. aeruginosa*

which may demonstrate the importance of national guidance in reducing transmissions.

By focusing on burns patients who receive hydrotherapy, our study population were at extremely high risk of waterborne infection. In other patient groups it may be that alternative routes of transmission including cross-infection or endogenous carriage play a more important role. Our results suggest that our burns unit is endemically colonised with a distinct clone of *P. aeruginosa* that may have been imported coinciding with the opening of the hospital. Other intensive care units, particularly those which have been open for longer may have harbour a greater diversity of *P. aeruginosa* as a result of increased opportunities for clones to be imported.

One potential application for WGS in infection control would be to determine whether cases are as a result of water transmission, or represent sporadic clones originating from the wider environment. Despite improved guidance concerning improved engineering infection control practices and the introduction o the water safety group in the UK, it may not be realistic to eliminate *P. aeruginosa* from hospitals entirely. In augmented care units such as ITUs, burns units and neonatal wards where *P. aeruginosa* poses a significant risk to vulnerable patients, the increased resolution offered by WGS will justify its use, particularly as the costs continue to fall.

In conclusion, we have identified through WGS clear evidence for transmission of *P. aeruginosa* from specific water outlets to burns patients and offer a forensic-level framework for dealing with outbreaks linked to hospital water. We expect WGS will continue to make inroads into clinical microbiology and become a vital tool for tracking *P. aeruginosa* in the hospital environment, helping inform targeted control measures to help protect patients at risk of infection.

Acknowledgments

We are grateful to Mark Webber for discussions on antibiotic resistance and to Paul Keim for discussion on phylogenetic placement of metagenomics samples. We thank Lex Nederbragt, Ave Tooming-Klunderud and the staff of the Norwegian Sequencing Centre, Oslo for Pacific Biosciences sequencing. We thank Matthew Smith-Banks for laboratory assistance with processing samples. We thank Jimmy Walker for critical reading of the manuscript. We thank Drs David Baltrus, Thomas Connor, Jennifer Gardy and Alan McNally for their helpful comments and suggestions to help improve the manuscript made during the open peer review process.

Contributors

MP NM BO conceived the study. CW and AB enrolled patients into study and collected samples. NC collected environmental and water samples. NC CC MN processed samples and performed microbiology. NC CC JQ did sequencing. JQ NC CT NL analysed the data. NL NC JQ MP BO wrote the paper. All authors commented on the manuscript draft.

Funding

This paper presents independent research funded by the National Institute for Health research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Competing interests

None

Data Sharing Statement

Pacific Biosciences raw data files are available from the corresponding author (Nicholas J Loman, n.j.loman@bham.ac.uk) and will be made available via Data Dryad.

Ethics approval

The study protocol received approval from National Research Ethics Service committee in the West Midlands (reference number 12/WM/0181).

Figure and Table Legends

Figure 1 Legend

An overview of all samples collected during the study in global phylogenetic context with other sequenced strains of *P. aeruginosa* from the set of Stewart *et al*.[28] Samples collected in this study are widely dispersed in the tree, which contains isolates from different environments (Panel A). Bar plots indicate the numbers of each type of sample collected (Panel B). Microdiversity within each clade is shown, with the colour bar indicating the source of each sample (Panel C).

Figure 2 Legend

A schematic view of the 300-day study of *P. aeruginosa* in a burns centre and critical care unit. Time in days is shown along the X-axis with bed numbers in the critical care unit and burns unit along the Y-axis. Each circular icon indicates a positive isolate of *P. aeruginosa*. The icon's logotype indicates which environment it originated from (wound, urine/sputum, environment or water). The filled colour of the icon indicates the clade it belongs to. Patient icons represent the enrolment of a screening patient into the study and their location. Patient movements around the hospital are noted by dotted lines. The five patients infected with *P. aeruginosa* are denoted by rounded boxes. Boxes are coloured according to the patient number. In the event two or more isolates of the same source and clade were collected on the same day, these have been collapsed into a single circular icon.

Figure 3 Legend

The high-resolution phylogenetic reconstruction of Clade E isolates. This demonstrates the clustering of genotypes by bed space. Patient associated samples are contained within the room 11 clade. This clade contains water samples from the shower and environmental samples from the shower, drain and trolley. The water samples from the room 11 tap are in a distinct clade, indicating the biofilm within the tap has a distinct genotype to the shower. This suggests environmental contamination was more likely to arise from contaminated shower water than tap water. Details of sampling site, days since start of study and presence of pBURNS plasmids are also shown. The likely phylogenetic position of *P. aeruginosa* detected in a biofilm from a thermostatic mixer valve is shown in the clade associated with room 9 and indicated 'TMV'.

Appendix 1

Phylogenetic tree, sample origin and antibiotic-resistance phenotype for Clade C isolates. The patient column refers to the sampling event. If a sample was collected from a patient colonised with *P. aeruginosa* or that patient's environment the patient number if marked. Patient numbers starting with SP relate to screening patients who were not colonised by *P. aeruginosa*.

Appendix 2

Phylogenetic tree, sample origin and antibiotic-resistance phenotype for Clade D isolates.

Appendix 3

Phylogenetic tree, sample origin and antibiotic-resistance phenotype for Clade E isolates.

Appendix 4

Phylogenetic tree, sample origin and antibiotic-resistance phenotype for Clade G isolates.

Appendix 5

A physical map of the burns unit, indicating individual patient bed areas, shower areas. The water supply is indicated by a blue line.

Appendix 6

Clade E isolates in phylogenetic context with the metagenomics sample from a tap mixer valve.

Appendix 7

Single nucleotide and small indel variants detected within Clade C isolates.

Appendix 8

Single nucleotide and small indel variants detected within Clade D isolates.

Appendix 9

Single nucleotide and small indel variants detected within Clade E isolates.

Appendix 10

Single nucleotide and small indel variants detected within Clade G isolates.

Appendix 11

The type of antibiotics administered to the five patients in this study.

Appendix 12

Mutations predicted to be associated with antibiotic-resistance.

REFERENCES

1 Reuter S, Sigge A, Wiedeck H, *et al.* Analysis of transmission pathways of Pseudomonas aeruginosa between patients and tap water outlets. *Crit Care Med* 2002;**30**:2222–8. doi:10.1097/01.CCM.0000025907.12833.10

2 Struelens MJ, Rost F, Deplano A, *et al.* Pseudomonas aeruginosa and Enterobacteriaceae bacteremia after biliary endoscopy: an outbreak investigation using DNA macrorestriction analysis. *Am J Med* 1993;**95**:489–98.

3 DiazGranados CA, Jones MY, Kongphet-Tran T, *et al.* Outbreak of Pseudomonas aeruginosa infection associated with contamination of a flexible bronchoscope. *Infect Control Hosp Epidemiol* 2009;**30**:550–5. doi:10.1086/597235

4 Crivaro V, Di Popolo A, Caprio A, *et al.* Pseudomonas aeruginosa in a neonatal intensive care unit: molecular epidemiology and infection control measures. *BMC Infect Dis* 2009;**9**:70. doi:10.1186/1471-2334-9-70

5 Moolenaar RL, Crutcher JM, San Joaquin VH, *et al.* A prolonged outbreak of Pseudomonas aeruginosa in a neonatal intensive care unit: did staff fingernails play a role in disease transmission? *Infect Control Hosp Epidemiol* 2000;**21**:80–5. doi:10.1086/501739

6 Trautmann M, Lepper PM, Haller M. Ecology of Pseudomonas aeruginosa in the intensive care unit and the evolving role of water outlets as a reservoir of the organism. *Am J Infect Control* 2005;**33**:S41–9. doi:10.1016/j.ajic.2005.03.006

7 Foca M, Jakob K, Whittier S, *et al.* Endemic Pseudomonas aeruginosa infection in a neonatal intensive care unit. *N Engl J Med* 2000;**343**:695–700. doi:10.1056/NEJM200009073431004

8 Kolmos HJ, Thuesen B, Nielsen SV, *et al.* Outbreak of infection in a burns unit due to Pseudomonas aeruginosa originating from contaminated tubing used for irrigation of patients. *J Hosp Infect* 1993;**24**:11–21.

9 Widmer AF, Wenzel RP, Trilla A, *et al.* Outbreak of Pseudomonas aeruginosa infections in a surgical intensive care unit: probable transmission via hands of a health care worker. *Clin Infect Dis* 1993;**16**:372–6.

10 Srinivasan A, Wolfenden LL, Song X, *et al.* An outbreak of Pseudomonas aeruginosa infections associated with flexible bronchoscopes. *N Engl J Med* 2003;**348**:221–7. doi:10.1056/NEJMoa021808

11 Wise J. Three babies die in pseudomonas outbreak at Belfast neonatal unit. BMJ. 2012;**344**:e592.

12 Breathnach AS, Cubbon MD, Karunaharan RN, *et al.* Multidrug-resistant Pseudomonas aeruginosa outbreaks in two hospitals: association with contaminated hospital waste-water systems. *J Hosp Infect* 2012;**82**:19–24. doi:10.1016/j.jhin.2012.06.007

13 Walker JT, Jhutty A, Parks S, *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. doi:10.1016/j.jhin.2013.10.003

14 *Health Technical Memorandum 04-01: Addendum*. Department of Health https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/140105/Health_Technical_Memorandum_04-01_Addendum.pdf (accessed 11 Dec2013).

15 Maiden MC, Bygraves JA, Feil E, *et al.* Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci USA* 1998;**95**:3140–5.

16 Jefferies JMC, Cooper T, Yam T, *et al.* Pseudomonas aeruginosa outbreaks in the neonatal intensive care unit--a systematic review of risk factors and environmental sources. *J Med Microbiol* 2012;**61**:1052–61. doi:10.1099/jmm.0.044818-0

17 Cholley P, Thouverez M, Hocquet D, *et al.* Most multidrug-resistant Pseudomonas aeruginosa isolates from hospitals in eastern France belong to a few clonal types. *J Clin Microbiol* 2011;**49**:2578–83. doi:10.1128/JCM.00102-11

18 Curran B, Jonas D, Grundmann H, *et al.* Development of a multilocus sequence typing scheme for the opportunistic pathogen Pseudomonas aeruginosa. *J Clin Microbiol* 2004;**42**:5644–9. doi:10.1128/JCM.42.12.5644-5649.2004

19 Harris SR, Cartwright EJP, Török ME, *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. doi:10.1016/S1473-3099(12)70268-2

20 Lewis T, Loman NJ, Bingle L, *et al.* High-throughput whole-genome sequencing to dissect the epidemiology of Acinetobacter baumannii isolates from a hospital outbreak. *J Hosp Infect* 2010;**75**:37–41. doi:10.1016/j.jhin.2010.01.012

21 Snitkin ES, Zelazny AM, Thomas PJ, *et al.* Tracking a hospital outbreak of carbapenem-resistant Klebsiella pneumoniae with whole-genome sequencing. *Sci Transl Med* 2012;**4**:148ra116. doi:10.1126/scitranslmed.3004129

22 Loman NJ, Misra RV, Dallman TJ, *et al.* Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol*

2012;**30**:434–9. doi:10.1038/nbt.2198

23 Tredget EE, Shankowsky HA, Joffe AM, *et al.* Epidemiology of infections with Pseudomonas aeruginosa in burn patients: the role of hydrotherapy. *Clin Infect Dis* 1992;**15**:941–9.

24 Langschmidt J, Caine PL, Wearn CM, *et al.* Hydrotherapy in burn care: A survey of hydrotherapy practices in the UK and Ireland and literature review. *Burns* Published Online First: 10 December 2013. doi:10.1016/j.burns.2013.11.006

25 Tredget EE, Shankowsky HA, Rennie R, *et al.* Pseudomonas infections in the thermally injured patient. *Burns* 2004;**30**:3–26.

26 Davison PG, Loiselle FB, Nickerson D. Survey on current hydrotherapy use among North American burn centers. *J Burn Care Res* 2010;**31**:393–9. doi:10.1097/BCR.0b013e3181db5215

27 Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* Published Online First: 28 April 2014. doi:10.1093/bioinformatics/btu170

28 Stewart L, Ford A, Sangal V, *et al.* Draft genomes of 12 host-adapted and environmental isolates of Pseudomonas aeruginosa and their positions in the core genome phylogeny. *Pathog Dis* Published Online First: 26 October 2013. doi:10.1111/2049-632X.12107

29 Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010;**26**:589–95. doi:10.1093/bioinformatics/btp698

30 Koboldt DC, Zhang Q, Larson DE, *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012;**22**:568–76. doi:10.1101/gr.129684.111

31 Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* 2010;**5**:e9490. doi:10.1371/journal.pone.0009490

32 Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 2008;**18**:821–9. doi:10.1101/gr.074492.107

33 Jolley KA, Maiden MC. BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* 2010;**11**:595–5. doi:10.1186/1471-2105-11-595

34 Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 2014;**15**:R46. doi:10.1186/gb-2014-15-3-r46

35 Matsen FA, Kodner RB, Armbrust EV. pplacer: linear time maximum-

likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* 2010;**11**:538. doi:10.1186/1471-2105-11-538

36 Scott FW, Pitt TL. Identification and characterization of transmissible Pseudomonas aeruginosa strains in cystic fibrosis patients in England and Wales. *J Med Microbiol* 2004;**53**:609–15.

37 Panagea S, Winstanley C, Parsons YN, *et al.* PCR-based detection of a cystic fibrosis epidemic strain of Pseudomonas Aeruginosa. *Mol Diagn* 2003;**7**:195–200.

38 Gardy JL, Johnston JC, Ho Sui SJ, *et al.* Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *N Engl J Med* 2011;**364**:730–9. doi:10.1056/NEJMoa1003176

39 Bryant JM, Grogono DM, Greaves D, *et al.* Whole-genome sequencing to identify transmission of Mycobacterium abscessus between patients with cystic fibrosis: a retrospective cohort study. *Lancet* 2013;**381**:1551–60. doi:10.1016/S0140-6736(13)60632-7

40 Lienau EK, Strain E, Wang C, *et al.* Identification of a salmonellosis outbreak by means of molecular sequencing. *N Engl J Med* 2011;**364**:981–2. doi:10.1056/NEJMc1100443

41 Lambert PA. Mechanisms of antibiotic resistance in Pseudomonas aeruginosa. *J R Soc Med* 2002;**95 Suppl 41**:22–6.

42 Sobel ML, Neshat S, Poole K. Mutations in PA2491 (mexS) promote MexT-dependent mexEF-oprN expression and multidrug resistance in a clinical strain of Pseudomonas aeruginosa. *J Bacteriol* 2005;**187**:1246–53. doi:10.1128/JB.187.4.1246-1253.2005

43 Quinn JP, Dudek EJ, DiVincenzo CA, *et al.* Emergence of resistance to imipenem during therapy for Pseudomonas aeruginosa infections. *J Infect Dis* 1986;**154**:289–94.

44 Cao L, Srikumar R, Poole K. MexAB-OprM hyperexpression in NalC-type multidrug-resistant Pseudomonas aeruginosa: identification and characterization of the nalC gene encoding a repressor of PA3720-PA3719. *Mol Microbiol* 2004;**53**:1423–36. doi:10.1111/j.1365-2958.2004.04210.x

45 Llanes C, Hocquet D, Vogne C, *et al.* Clinical strains of Pseudomonas aeruginosa overproducing MexAB-OprM and MexXY efflux pumps simultaneously. *Antimicrob Agents Chemother* 2004;**48**:1797–802.

46 Cambau E, Perani E, Dib C, *et al.* Role of mutations in DNA gyrase genes in ciprofloxacin resistance of Pseudomonas aeruginosa susceptible or resistant to imipenem. *Antimicrob Agents Chemother* 1995;**39**:2248–52.

47 Kelsey M. Pseudomonas in augmented care: should we worry? *J Antimicrob Chemother* 2013;**68**:2697–700. doi:10.1093/jac/dkt288

48 Rogers DE. The changing pattern of life-threatening microbial disease. *N Engl J Med* 1959;**261**:677–83. doi:10.1056/NEJM195910012611401

49 Martin K, Baddal B, Mustafa N, *et al.* Clusters of genetically similar isolates of Pseudomonas aeruginosa from multiple hospitals in the UK. *J Med Microbiol* 2013;**62**:988–1000. doi:10.1099/jmm.0.054841-0

50 Slekovec C, Plantin J, Cholley P, *et al.* Tracking down antibiotic-resistant Pseudomonas aeruginosa isolates in a wastewater network. *PLoS One* 2012;**7**:e49300. doi:10.1371/journal.pone.0049300

51 Snyder LA, Loman N, Faraj LA, *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**.

52 Donlan RM. Biofilm formation: a clinically relevant microbiological process. *Clin Infect Dis* 2001;**33**:1387–92. doi:10.1086/322972

Title

Seeking the source of *Pseudomonas aeruginosa* infections in a recently opened hospital: an observational study using whole-genome sequencing

Authors

Joshua Quick [1,2*], Nicola Cumley [2*], Christopher M. Wearn [2,3], Marc Niebel [2], Chrystala Constantinidou [4], Chris M Thomas [1], Mark J. Pallen [4], Naiem S. Moiemen [2,3], Amy Bamford [2,3], Beryl Oppenheim [2#], Nicholas J. Loman [1#]

Affiliations

[1]Institute of Microbiology and Infection, University of Birmingham, Birmingham, United Kingdom

[2]NIHR Surgical Reconstruction and Microbiology Research Centre, Queen Elizabeth Hospital, Birmingham, United Kingdom

[3]Healing Foundation Centre for Burns Research, University Hospital Birmingham Foundation Trust, Birmingham, United Kingdom

[4]Division of Microbiology and Immunology, University of Warwick, Warwick, United Kingdom

*Contributed equally

Correspondence:

Dr Nicholas James Loman

Institute of Microbiology and Infection, University of Birmingham, Birmingham, B15 2TT United Kingdom

Telephone: +44 (0) 121 414 8849

Email: n.j.loman@bham.ac.uk

Dr Beryl Oppenheim

Clinical Microbiology, University Hospitals Birmingham NHS Foundation Trust

Queen Elizabeth Hospital Birmingham, Mindelsohn Way, Edgbaston

Birmingham, B15 2WB

Telephone: +44 (0) 121 371 6523

Email: Beryl.Oppenheim@uhb.nhs.uk

Keywords

High-Throughput DNA sequencing; Pseudomonas aeruginosa; Burns, Hydrotherapy

Word Count

4539

ABSTRACT

Objectives

Pseudomonas aeruginosa is a common nosocomial pathogen responsible for significant morbidity and mortality internationally. Patients may become colonised or infected with *P. aeruginosa* after exposure to contaminated sources within the hospital environment. The aim of this study was to determine whether whole-genome sequencing (WGS) can be used to determine the source in a cohort of burns patients at high risk of *P. aeruginosa* acquisition.

Study design

An observational prospective cohort study.

Setting

Burns care ward and critical care ward in the United Kingdom.

Participants

Patients with >7% total burns by surface area were recruited into the study.

Methods

All patients were screened for *P. aeruginosa* on admission and samples taken from their immediate environment, including water. Screening patients who subsequently developed a positive *P. aeruginosa* microbiology result were subject to enhanced environmental surveillance. All isolates of *P. aeruginosa* were whole-genome sequenced. Sequence analysis looked at similarity and relatedness between isolates.

Results

Whole-genome sequences for 141 *P. aeruginosa* isolates were obtained from patients, hospital water and the ward environment. Phylogenetic analysis revealed eight distinct clades, with a single clade representing the majority of environmental isolates in the burns unit. Isolates from three patients had identical genotypes compared with water isolates from the same room. There was clear clustering of water isolates by room and outlet, allowing the source of acquisitions to be unambiguously identified. Whole-genome shotgun sequencing of biofilm DNA extracted from a thermostatic mixer valve revealed this was the source of a *P. aeruginosa* subpopulation previously detected in water. In the remaining two cases there was no clear link to the hospital environment.

Conclusions

This study reveals that WGS can be used for source tracking of *P. aeruginosa* in a hospital setting, and that acquisitions can be traced to a specific source within a hospital ward.

ARTICLE SUMMARY

Strengths and limitations of the study

- We have demonstrated that whole-genome sequencing can be used for source tracking of *P. aeruginosa* in a hospital setting.
- We show convincing evidence that transmission has occurred directly from water to patients, but other routes are as likely.
- The main limitation of the study was the sample size, which could be attributable to interventions being carried out during the study.
- Our study focused on a burns unit and critical care unit in a newly built hospital. Modes of *P. aeruginosa* transmission may be different in hospitals with different styles of plumbing and on other augmented care units.

INTRODUCTION

Pseudomonas aeruginosa is a ubiquitous Gram-negative bacterium and an important opportunistic pathogen in the healthcare setting. *P. aeruginosa* particularly affects those with impaired host or mucosal immunity and has a broad range of presentations including respiratory infections in cystic fibrosis and mechanically ventilated patients, bloodstream infections in premature neonates and wounds in burns injuries. Nosocomial *P. aeruginosa* outbreaks are frequently reported and associated with water sources such as taps, showers, mixer valves and flow straighteners, sink traps and drains.[1-10] Other potential routes of transmission include cross-infection, for example carriage on the hands of health care workers, and through contaminated medical equipment such as endoscopic devices.[3,5]

In the UK, the role of water in the transmission of *P. aeruginosa* in healthcare settings has become a matter of urgent concern in response to a recent high-profile outbreak affecting a neonatal critical care unit in Belfast in 2012.[11] This source was eventually determined to be sink taps.[11-13] National guidance is now in place detailing enhanced procedures for routine water sampling on augmented care units, with directed interventions such as disinfection and replacement of high-risk plumbing parts required.[14]

Historical phenotypic typing methods for *P. aeruginosa* such as O-antigen serotyping have more recently been replaced by molecular typing methods such as pulsed-field gel electrophoresis (PFGE), variable number tandem repeat analysis (VNTR) and random amplification of polymorphic DNA (RAPD) and multi-locus sequencing typing (MLST).[15] These methods have been used to investigate outbreaks of *P. aeruginosa* within hospitals.[4,16-18] However, such techniques have important

limitations for source tracking of infections in hospitals as they sample limited numbers of sites in the genome which may result in false clustering of unrelated strains.[19] In the past five years, whole-genome sequencing (WGS) has started to be used to investigate outbreaks in hospitals. WGS is attractive because of its digital, sharable format and ultra-high resolution, which is able to discriminate two isolates differing by just a single mutation. WGS has been successfully used to determine likely transmission chains during outbreaks of *Staphylococcus aureus*, *Acinetobacter baumannii* and *Klebsiella pneumoniae*.[19-21] Benchtop sequencing instruments now offer a cost-effective approach for bringing bacterial WGS to the clinical environment. [1,22]

In this study, we explore the utility of WGS to determine the likely sources of *P. aeruginosa* in an at-risk population of burns patients. In the UK and US burns patients receive shower cart hydrotherapy as a mainstay of burns treatment.[23-26] A previous hospital audit suggested that up to one-third of such patients became colonised with *P. aeruginosa*. We hypothesised that this high rate of acquisition may relate to transmission from hospital shower water during therapy. We therefore wished to understand the importance of transmission from water compared with alternative routes such as cross-infection and endogenous carriage.

MATERIALS AND METHODS

Hospital setting

An observational, prospective study design was employed in a burns centre serving approximately 13.7 million people across the Midlands region of England with 300 admissions annually. Opened in June 2010, the burns centre comprises a purpose built 15-bed ward with 11 side-rooms and 2 dual-bedded rooms. Patients requiring mechanical ventilation and organ support are usually treated in two self-contained burns cubicles located within the trauma critical care unit (CCU). Despite the observational nature of the study, sampling was carried out during implementation of interim national guidance on control of *P. aeruginosa* issued by the Department of Health. These guidelines were issued in draft form March 2012, and subsequently revised in March 2013. This meant that parallel water sampling and engineering interventions were being undertaken during the period of study. In addition, some enhanced infection prevention measures were also introduced in response to an outbreak of a multi-drug resistant *A. baumannii.*

Study design and patient selection

Ethical approval was sought and received from the National Research Ethics Service committee in the West Midlands (reference number 12/WM/0181). Patients admitted to the burns unit were eligible for the screening phase of the study if they had burns injuries covering greater than 7% total body surface area (TBSA). Patients were screened as soon as possible after admission after they had given written informed consent. When appropriate, legal consultee advice was sought for patients lacking capacity due to emergency treatment. On admission, recruited patients were screened for carriage of *P. aeruginosa* (wounds, urine and stool) using standard

microbiology techniques. Samples were then taken as part of routine microbiology service during the patients stay. Environmental and water samples were taken after the patient was admitted to the burns centre. If during the period of stay *P. aeruginosa* was isolated from a patient sample the patient was recruited into the second phase of the study. In this phase, patients had wound swabs taken at each dressing change as well as twice-weekly urine samples. The patient's environment and water from outlets in their bed space were sampled weekly for the duration of their stay, and after discharge (post-cleaning). Termination of the study was planned after 30 screening patient admissions, or a year, whichever came soonest, after which 10 patients were expected to acquire *P. aeruginosa*. This prediction was based on a previous local audit which suggested about one-thirds of burns patients became colonised with *P. aeruginosa*..

Microbiological and molecular methods

P. aeruginosa isolates were obtained from wound swab, urine, stool, environmental and water samples. *P. aeruginosa* was isolated from wound swabs, urine and stool by inoculation onto cysteine lactose electrolyte deficient agar (CLED) and cetrimide agar and incubation for 24 hours at 37°C. Stool samples were cultured overnight in a cetrimide enrichment broth before subculture onto CLED. Identification was confirmed by resistance to C-390 and the VITEK® 2 GN identification card. Antibiotic sensitivity assays were performed using the VITEK® 2 AST N-210 card (bioMérieux, Basingstoke, UK).

The patient's environment (shower head rosette, drain, shower chair or trolley, bedside table, patient chair, instruments in contact with the patient) was sampled over a 10cm^2 area by a Polywipe™ sponge. The sponge was placed in tryptic soy

broth incubated for 24 hours at 37˚C then sub-cultured onto CLED and cetrimide agar. During water sampling, water was taken from the patient's shower, or tap if a shower was not present. Shower heads were not removed for water sampling. At least 200ml of water was collected into a vessel containing sodium thiosulphate as a neutraliser. In duplicate, 100ml of water was filtered through a 0.45 micron filter and the filters placed onto CLED plates and cetrimide agar. Plates were incubated at 37˚C for 48 hours and the number of organisms per 100ml quantified.

For storage and DNA extraction a single colony was purified from the primary culture plate. When different colony morphologies were observed, a single colony from each type was purified. Additionally, for a randomly selected water sample, 24 colonies were individually picked from one water-filter primary microbiological plate for sequencing. Isolates were stored on Biobank beads at -20˚C prior to DNA extraction. Organisms were resuscitated on CLED agar plates and genome DNA either extracted directly using the MOBIO UltraClean Microbial DNA Kit, or from overnight LB broth culture using a Qiagen Genomic-Tip 100G.

DNA extraction and sequencing

Genomic DNA was prepared from single colony picks using the MIOBIO Ultraclean microbial kit (MOBIO, Carlsbad, USA). 1ng input DNA, as quantified by Qubit (Life Technologies, Carlsbad, USA) was used to prepare genomic libraries for sequencing using the Illumina Nextera XT™ DNA sample kit as per manufacturer's protocol (Illumina, San Diego, USA). Libraries were sequenced on the Illumina MiSeq using a paired-end protocol resulting in read lengths between 150 and 300 bases. A single additional sample, isolate 910, was chosen as a representative member of Clade 5 for long-read sequencing. DNA from this sample was fragmented using a

Hydroshear (Digilab, Marlborough, MA) using the recommended protocol for 10kb fragments and further size-selected on a Blue Pippen instrument (Sage Science, MA) with a 7kb minimum size cut-off. The library was sequenced on two SMRT Cells using the Pacific Biosciences RS II instrument at the Norwegian Sequencing Centre, Oslo. C4-P2 chemistry was chosen because it favours long, more accurate reads for *de novo* assembly.

Stool PCR

For simple presence/absence detection of *P. aeruginosa* in stool samples using PCR, a stool sample was collected into a stool collection tube containing stool DNA stabilizer. Total DNA was extracted using the PSP Spin Stool DNA Plus kit (Stratec Molecular). PCR amplification of species specific regions of the 16S rDNA gene was carried out using primers PA-SS-F: GGGGGATCTTCGGACCTCA and PA-SS-R: TCCTTAGAGTGCCCACCCG [12] in the following conditions: 0.5µM of each primer, 1.5mM $MgCl_2$, 0.2mM dNTP's using BIOTAQ™ DNA Polymerase and buffer set. After initial denaturation at 96°C for 2 minutes, 30 cycles of 96°C for 30 seconds, 62° C for 30 seconds and 72°C for 30 seconds were completed with a final extension of 72°C for 5 minutes. Products were visualised for size on an 1.5% agarose gel.

Bioinformatics methods

Illumina MiSeq reads from each isolate were adapter and quality trimmed before use with Trimmomatic.[27] Phylogenetic reconstruction of isolates sequenced in this study were combined with data from a global collection of 55 *P. aeruginosa* strains collected world-wide which have been previously analysed by Stewart et al. [28] For each of the published strains, 600,000 paired-end reads of length 250 bases were simulated using wgsim (https://github.com/lh3/wgsim) from the complete or draft

genome assembly deposited in Genbank. Read sets were mapped against the *P. aeruginosa* PAO1 reference genome using BWA-MEM 0.7.5a-r405 using default settings.[29] Single nucleotide polymorphisms were called using VarScan 2.3.6 and filtered for regions with an excessive number of variants. These may represent regions of recombination, misalignments or strong Darwinian selection.[30] FastTree (version 2.1.7) was used for phylogenetic reconstruction. This software estimates an approximate maximum-likelihood tree under the Jukes-Cantor model of nucleotide evolution with a single rate for each site (CAT).[31] Trees were drawn in FigTree (http://tree.bio.ed.ac.uk/software/figtree/).

For *in silico* MLST prediction, trimmed reads were assembled *de novo* using Velvet [32] with a k-mer size of 81 and searched using nucleotide BLAST against the multi-locus sequence database downloaded from the pubMLST website on 2013-08-05 (http://pubmlst.org/paeruginosa/) [33]. For Clade E isolates, in order to exhaustively search for discriminatory mutations, a nearly complete reference genome was generated by *de novo* assembly using Pacific Biosciences sequencing data. Reads were assembled using the 'RS_HGAP_Assembly.3' pipeline within SMRT Portal v2.2.0. Illumina reads from the same sample were mapped to this draft genome assembly in order to correct remaining indel errors in the assembly using Pilon (http://www.broadinstitute.org/software/pilon/). Isolates belonging to each clade were mapped individually against either the PacBio reference (Clade E) or *P. aeruginosa* PAO1 (NC_002516) (Clades C, D and G).

Variants (single nucleotide polymorphisms and short insertion-deletions) were called using SAMtools mpileup and VarScan with an allele frequency threshold of 80%.**[30]** Non-informative positions and regions of putative recombination were removed, the

later with a variant density filter of more than 3 SNPs every 1000 nucleotides. Analysing samples in each clade individually maximised the number of variants detected by reducing the likelihood of the position being uncovered by a subset of samples. From these variants fine-grained phylogenetic trees were reconstructed for each clade using FastTree. The scripts used to perform this analysis are available at http://www.github.com/joshquick/snp_calling_scripts. Approximate-maximum-likelihood phylogenetic trees were generated using FastTree and visualised in FigTree. For whole-genome shotgun metagenomics analysis, reads were analysed using the Kraken taxonomic classifier software with the supplied *minikraken* database.**[34]** Reads from the metagenomics dataset were aligned to *P. aeruginosa* PAO1 as in the previous section and phylogenetic placement was carried out using pplacer in conjunction with FastTree.**[35]** Sequence data is available from the European Nucleotide Archive for the Illumina data (ERP006056) and the corrected Pacific Biosciences assembly (ERP006058).

RESULTS

Study results

Recruitment lasted a period of 300 days, ending according to protocol after the enrolment of 30 screening patients. In total, we detected *P. aeruginosa* in five patients. Of these patients, three had *P. aeruginosa* detected only in burns wound swabs, one had *P. aeruginosa* detected in both their burns wound and in their urine, and one had *P. aeruginosa* in their sputum. One additional eligible patient did not consent to enter the study and was excluded. The average age in the study group was 41 years. Males predominated with a male-to-female ration of 2.3:1. Flame burns were the most common mechanism of injury, followed by scalds and mixed flame/flash injuries. The average burn size of the study group was 12.5% of the total

body surface area (TBSA) and 27% of patients sustained an inhalation injury. Eight patients required admission to ITU and the majority required surgical treatment of their burns with excision and skin grafting (80%). A large majority of the study group (83%) received shower cart hydrotherapy as a routine part of their wound management to encourage healing through wound debridement and decontamination. The average length of hospital stay (LOS) was 17 days and taking into account burn size, the average was 1.4 days per % TBSA.

The water and environment in burns and critical care units are frequently colonised by *P. aeruginosa*

A total of 282 water and environmental samples were screened for *P. aeruginosa* of which 39/78 (50%) were positive in water samples, 25/96 (26%) were positive from the wet environment and 7/108 (6%) were positive from the dry environment. A total of 86 genome sequences were generated from the 71 positives, as in some cases multiple colony picks were sequenced. 78 patient samples were screened for *P. aeruginosa* of which 39 (50%) were positive. A total of 55 genome sequences were generated, as in some cases multiple colony picks were sequenced. In total, 141 genomes were sequenced; water and environmental (n=86) and patient (n=55). Genomes were sequenced to a mean coverage of 24.4x, with the minimum coverage of a sample being 14x and highest 64.7x.

When placed in the context of a global collection of *P. aeruginosa* strains, phylogenetic reconstruction demonstrated isolates in our study fell into eight clades (figure 1, Panel A). As has been reported previously, there was no strong association between ecological context and position in the phylogenetic tree.[28] Isolates in this study are most closely related to strains from a variety of settings. The majority of

isolates (52%) belong to Clade E (figure 1 Panel B), whose nearest sequenced relative is the Liverpool Epidemic Strain, a clone often isolated from patients in the UK and Canada with cystic fibrosis. [36,37] Isolates from Clade E were found in the burns unit's water and the ward environment, as well as from two patient's wounds. However it was never detected in the critical care unit. Clade E was detected throughout the study in a total of 10 different rooms (figure 2).

Inferring potential transmission events by whole-genome sequencing

Microevolutionary changes occurring over rapid time-scales (i.e. days to months) have been used to detect potential chains of transmission in hospital and community outbreaks.[19-21,38,39] The number of distinct mutations between given isolates has been used to infer whether transmission events are likely to have occurred. Such inferences are aided by prior knowledge of mutation rates in similar populations. Two patients (1 and 4) in our study both had *P. aeruginosa* from Clade E isolated from their wounds. These isolates had an indistinguishable genotype from those present in water and the environment of the room they were nursed within (figure 1 Panel C and figure 3). This genotype was detected in the patient's shower water after initial patient screening, during screening of the second patient admission, twice during the second patient's stay and then 127 days later (days 27, 65, 89 and 216 respectively). When water isolates were positive, the genotype was also detected in wet environment sites (shower drain, shower rosette and patient's trolley) on the same days.

Patient 5 was nursed on the critical care unit due to concomitant medical problems. *P. aeruginosa* belonging to Clade G was isolated from sputum during this time.

Identical genotypes were detected contemporaneously in the water from the associated sink and sink tap handle (see online supplementary appendix 4).

Two further patients (patients 2 and 3) were positive for *P. aeruginosa*. Isolates from these patients belonged to Clade C and D respectively. Neither clade was ever isolated from hospital water. In both cases, identical genotypes were detectable in the environment associated with the patient but these were not detected before or after the patients' stay, indicating that the environment was not persistently contaminated. During the course of patient 3's stay, the dry environment such as the bedside table was contaminated, as was the patient's door handle and shower chair. However, after patient discharge, the strain associated with this patient was never seen again during the course of the study in any location.

Whole-genome sequencing permits source tracking of *P. aeruginosa* to individual water outlets

Whole-genome sequencing has been reported previously for source tracking, but never for the detection of transmission events from hospital water.[40] Phylogenetic reconstruction within Clade E, the most commonly detected water clone demonstrated additional diversity within this clone, with a total of 46 mutations detected an average genetic distance between isolates of 4.1 mutations (figure 3). The reconstruction demonstrated clear evidence of clustering of genotypes both by room and outlet (figure 3). When *P. aeruginosa* was detected in the wet environment (e.g. shower rosettes and drains) these genotypes were most often identical to those found in water, indicating that the water was likely the ultimate source of that clone. Genotypic variation was seen between outlets within the same room. For example, tap water sampled from room 11 had a distinct genotype from that sampled from

shower water in the same room and this was consistently found over multiple samplings. Notably, isolates from two patients fell within the cluster originating from shower water, indicating that shower hydrotherapy was the most likely source of infection. Two plasmids (designated pBURNS1 and pBURNS2) were detected in this study set, which both demonstrated geographical clustering, with pBURNS1 only being detectable in isolates from room 8 and pBURNS2 only being detectable in isolates from the shower water in room 9.

Rapid evolution of antibiotic resistance associated with treatment

P. aeruginosa is commonly associated with antibiotic resistance due to a number of predisposing features including intrinsic resistance, a repertoire of efflux pumps and antibiotic-inactivating enzymes including beta-lactamases. [41] Three infected patients (2, 3 and 5) received antibiotic therapy, and in each case this was associated with the development of resistance to at least one therapeutic agent. Associated mutations were detected that were either partially or fully explanatory of the phenotype (online supplementary appendix 12).

Patient 2 was treated with ciprofloxacin, nitrofurantoin and vancomycin (see online supplementary appendix 11 for full details). 8/21 (38%) tested isolates from this patient were ciprofloxacin resistant. 7/8 isolates (88%) of the ciprofloxacin-resistant strains were distinguishable from the other isolates by a single SNP in *mexS* (annotated as PA2491 in *P. aeruginosa* PAO1) (see online supplementary appendix 1 and 7). This SNP was predicted to result in a non-synonymous amino acid substitution. Disruption of this gene has been shown to cause increased expression of the mexEF-oprN multidrug efflux pump, associated with resistance to quinolones.[42]

Patient 3 was not treated with antibiotics, but isolates associated with this patient demonstrated differences in resistance to timentin and piperacillin-tazobactam. These changes were associated with non-synonymous mutations in *gacA*, the response regulator of the GacA/GacS two-component system and in *lasR*, a transcriptional activator required for transcription of elastase and LasA protease (online supplementary appendices 2 and 8).

Patient 4 was treated with meropenem, piperacillin/tazobactam, flucloxacillin and colistin. Five isolates collected 10 to 18 days after initiation of meropenem showed resistance to imipenem and intermediate resistance to meropenem (see online supplementary appendix 3 and 9). The most likely mutation responsible for this phenotype was detectable in two isolates, both of which had a frame-shift mutation in the gene coding for the membrane porin *oprD.* [43]

Patient 5 had a prolonged stay in ITU and had multiple medical problems including *A. baumannii* infection and was treated with nine antibiotic agents including ciprofloxacin, meropenem and piperacillin-tazobactam. Serial isolates from this patient demonstrated the stepwise acquisition of two mutations (online supplementary appendix 4). The first was in *nalC*, a probable repressor of the TetR/AcrR family (online supplementary appendix 10). [44] On inspection of the sequence alignment in this region, a large deletion of 196 nucleotide bases was seen compared to the reference PAO1 strain. This mutation was seen in association with full resistance to piperacillin-tazobactam, ceftazidine, aztreonam, meropenem and intermediate resistance to ciprofloxacin. This deletion is likely to result in over-expression of efflux pumps involving the *mexAB-oprM* operon. [44,45] Ciprofloxacin resistance in a later isolate corresponded to the stepwise acquisition of a second

mutation. This mutation is predicted to affect the well-studied DNA gyrase subunit A gene (*gyrA)* which is strongly associated with ciprofloxacin resistance.[46]

Confirmation of *P. aeruginosa* genotypes in biofilms by whole-genome metagenomic shotgun sequencing

P. aeruginosa is able to produce and survive in biofilms. Plumbing parts such as flow straighteners, shower rosettes, flexible hoses, solenoid valves and thermostatic mixer valves (TMV) are particularly at risk of biofilm formation due to factors including surface areas, convoluted designs and inadequate pasteurisation. [47] To confirm the presence of *P. aeruginosa* in water fittings associated with rooms on the burns unit, we obtained a TMV removed by the hospital estates team from the shower in room nine as part of compliance with UK guidelines for managing *P. aeruginosa* in hospitals. On visual inspection, a biofilm was present which was scraped from the surface with a sterile scalpel. DNA from this biofilm was extracted for whole-genome shotgun sequencing. The majority of reads did not map to any known bacterial taxa. The most abundant taxon identified was *Pseudomonas aeruginosa* (3%). Subsequent alignment to the *P. aeruginosa* PAO1 reference covered 94% of the 6.3 million base reference genome at a median coverage of 5x, confirming that reads were correctly classified to this species and not other environmental *Pseudomonas* species. Alignment to the *P. aeruginosa* Clade E reference genome followed by phylogenetic placement of reads demonstrated that it fell into the same clade as previously recovered isolates from the shower or tap in room 9 (indicated on figure 3, and in online supplementary appendix 6).

DISCUSSION

The hospital environment has been intimately linked with *P. aeruginosa* infection for over 50 years yet hospital acquisitions, clusters and outbreaks remain a common occurrence and understanding precise routes of transmission can be difficult. [47,48] Our results demonstrate that, even in a new hospital, *P. aeruginosa* can become rapidly endemic in hospital plumbing. Furthermore, by linking *P. aeruginosa* genotypes recovered from patients to specific individual water outlets, we offer compelling evidence of unidirectional transmission from water to patients. Further, by sequencing of a biofilm identified in a TMV from a hospital water system, we can identify the likely common source of genotypes found in water and in the hospital environment.

Our results suggest that use of whole-genome sequencing can reduce ambiguity about potential transmission events in hospitals and consequently inform infection prevention efforts about the direction and sequence of transmission. Typing schemes such as MLST and PFGE are much lower resolution methods and would not be able to provide sufficient information to permit such inferences to be made. It is notable that the burns unit was colonised by a single clone, meaning that it was very unlikely that water outlets at each bed space were colonised as a result of transmissions from the patient or environment. For this to happen would require multiple transmission events from separate patients with the same clone, for which there is not evidence. Instead we speculate that this clone was introduced to the hospital associated with its commissioning. One hypothesis is that particular plumbing fittings, i.e. the TMV may have been colonised simultaneously by a clone circulating in water. Clade E (ST395) has been frequently reported associated with water, so this remains a possibility. [49,50] However, it is possible that plumbing fittings are installed ‘pre-seeded’ with *P. aeruginosa* as has already been proposed

by Kelsey.[3,5,47] Investigation of an outbreak in Wales implicated new plumbing parts as a potential source of *P. aeruginosa.* New plumbing components are often tested by companies prior to their supply and it is possible they were contaminated prior to distribution. The limited amount of diversity (average 4 SNPs) seen within this clade is consistent with a single founding genotype coinciding with the opening of the burns unit, based on estimates from a previous study using WGS which reported that mutations accumulate at a rate of approximately one every 3-4 months in a hospital-associated clone.[51] However our results suggest that our isolates accumulate mutations even more slowly. This may be due to reduced growth rates in nutritionally-poor biofilms.[52]

It is notable that antibiotic resistance to multiple first-line agents developed rapidly in response to therapy. These results underline the importance of selecting appropriate antibiotic therapy in *P. aeruginosa* infections. It is reassuring however that antibiotic resistance genotypes selected *in vivo* did not show evidence of persistence in the ward environment or transmission to other patients.

Our study has certain limitations. Based on a previous audit, we expected around one-third of patients screened for *P. aeruginosa* would develop colonisation or clinical infection. In fact, only 5 out of 30 of patients were colonised. This may have been related to guidance and engineering interventions being put in place during the study as detailed in national guidance issued whilst this study was on-going. In addition, infection control policies were revised to address control of an outbreak of a multi-drug resistant *A. baumannii* in this same burns unit. Following these interventions, only 1 of the last 20 patients recruited was infected with *P. aeruginosa*

which may demonstrate the importance of national guidance in reducing transmissions.

By focusing on burns patients who receive hydrotherapy, our study population were at extremely high risk of waterborne infection. In other patient groups it may be that alternative routes of transmission including cross-infection or endogenous carriage play a more important role. Our results suggest that our burns unit is endemically colonised with a distinct clone of *P. aeruginosa* that may have been imported coinciding with the opening of the hospital. Other intensive care units, particularly those which have been open for longer may have harbour a greater diversity of *P. aeruginosa* as a result of increased opportunities for clones to be imported.

One potential application for WGS in infection control would be to determine whether cases are as a result of water transmission, or represent sporadic clones originating from the wider environment. Despite improved guidance concerning improved engineering infection control practices and the introduction o the water safety group in the UK, it may not be realistic to eliminate *P. aeruginosa* from hospitals entirely. In augmented care units such as ITUs, burns units and neonatal wards where *P. aeruginosa* poses a significant risk to vulnerable patients, the increased resolution offered by WGS will justify its use, particularly as the costs continue to fall.

In conclusion, we have identified through WGS clear evidence for transmission of *P. aeruginosa* from specific water outlets to burns patients and offer a forensic-level framework for dealing with outbreaks linked to hospital water. We expect WGS will continue to make inroads into clinical microbiology and become a vital tool for tracking *P. aeruginosa* in the hospital environment, helping inform targeted control measures to help protect patients at risk of infection.

Acknowledgments

We are grateful to Mark Webber for discussions on antibiotic resistance and to Paul Keim for discussion on phylogenetic placement of metagenomics samples. ~~We thank Chris Thomas for useful discussions about *P. aeruginosa* plasmids.~~ We thank Lex Nederbragt, Ave Tooming-Klunderud and the staff of the Norwegian Sequencing Centre, Oslo for Pacific Biosciences sequencing. We thank Matthew Smith-Banks for laboratory assistance with processing samples. We thank Jimmy Walker for critical reading of the manuscript. We thank Drs David Baltrus, Thomas Connor, Jennifer Gardy and Alan McNally for their helpful comments and suggestions to help improve the manuscript made during the open peer review process.

Contributors

MP NM BO conceived the study. CW and AB enrolled patients into study and collected samples. NC collected environmental and water samples. NC CC MN processed samples and performed microbiology. NC CC JQ did sequencing. JQ NC CT NL analysed the data. NL NC JQ MP BO wrote the paper. All authors commented on the manuscript draft.

Funding

This paper presents independent research funded by the National Institute for Health research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Competing interests

None

Data Sharing Statement

Pacific Biosciences raw data files are available from the corresponding author (Nicholas J Loman, n.j.loman@bham.ac.uk) and will be made available via Data Dryad.

Ethics approval

The study protocol received approval from National Research Ethics Service committee in the West Midlands (reference number 12/WM/0181).

REFERENCES

1 Reuter S, Sigge A, Wiedeck H, *et al.* Analysis of transmission pathways of Pseudomonas aeruginosa between patients and tap water outlets. *Crit Care Med* 2002;**30**:2222–8. doi:10.1097/01.CCM.0000025907.12833.10

2 Struelens MJ, Rost F, Deplano A, *et al.* Pseudomonas aeruginosa and Enterobacteriaceae bacteremia after biliary endoscopy: an outbreak investigation using DNA macrorestriction analysis. *Am J Med* 1993;**95**:489–98.

3 DiazGranados CA, Jones MY, Kongphet-Tran T, *et al.* Outbreak of Pseudomonas aeruginosa infection associated with contamination of a flexible bronchoscope. *Infect Control Hosp Epidemiol* 2009;**30**:550–5. doi:10.1086/597235

4 Crivaro V, Di Popolo A, Caprio A, *et al.* Pseudomonas aeruginosa in a neonatal intensive care unit: molecular epidemiology and infection control measures. *BMC Infect Dis* 2009;**9**:70. doi:10.1186/1471-2334-9-70

5 Moolenaar RL, Crutcher JM, San Joaquin VH, *et al.* A prolonged outbreak of Pseudomonas aeruginosa in a neonatal intensive care unit: did staff fingernails play a role in disease transmission? *Infect Control Hosp Epidemiol* 2000;**21**:80–5. doi:10.1086/501739

6 Trautmann M, Lepper PM, Haller M. Ecology of Pseudomonas aeruginosa in the intensive care unit and the evolving role of water outlets as a reservoir of the organism. *Am J Infect Control* 2005;**33**:S41–9. doi:10.1016/j.ajic.2005.03.006

7 Foca M, Jakob K, Whittier S, *et al.* Endemic Pseudomonas aeruginosa infection in a neonatal intensive care unit. *N Engl J Med* 2000;**343**:695–700. doi:10.1056/NEJM200009073431004

8 Kolmos HJ, Thuesen B, Nielsen SV, *et al.* Outbreak of infection in a

burns unit due to Pseudomonas aeruginosa originating from contaminated tubing used for irrigation of patients. *J Hosp Infect* 1993;**24**:11–21.

9 Widmer AF, Wenzel RP, Trilla A, *et al.* Outbreak of Pseudomonas aeruginosa infections in a surgical intensive care unit: probable transmission via hands of a health care worker. *Clin Infect Dis* 1993;**16**:372–6.

10 Srinivasan A, Wolfenden LL, Song X, *et al.* An outbreak of Pseudomonas aeruginosa infections associated with flexible bronchoscopes. *N Engl J Med* 2003;**348**:221–7. doi:10.1056/NEJMoa021808

11 Wise J. Three babies die in pseudomonas outbreak at Belfast neonatal unit. BMJ. 2012;**344**:e592.

12 Breathnach AS, Cubbon MD, Karunaharan RN, *et al.* Multidrug-resistant Pseudomonas aeruginosa outbreaks in two hospitals: association with contaminated hospital waste-water systems. *J Hosp Infect* 2012;**82**:19–24. doi:10.1016/j.jhin.2012.06.007

13 Walker JT, Jhutty A, Parks S, *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. doi:10.1016/j.jhin.2013.10.003

14 *Health Technical Memorandum 04-01: Addendum*. Department of Health https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/140105/Health_Technical_Memorandum_04-01_Addendum.pdf (accessed 11 Dec2013).

15 Maiden MC, Bygraves JA, Feil E, *et al.* Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci USA* 1998;**95**:3140–5.

16 Jefferies JMC, Cooper T, Yam T, *et al.* Pseudomonas aeruginosa outbreaks in the neonatal intensive care unit--a systematic review of risk factors and environmental sources. *J Med Microbiol* 2012;**61**:1052–61. doi:10.1099/jmm.0.044818-0

17 Cholley P, Thouverez M, Hocquet D, *et al.* Most multidrug-resistant Pseudomonas aeruginosa isolates from hospitals in eastern France belong to a few clonal types. *J Clin Microbiol* 2011;**49**:2578–83. doi:10.1128/JCM.00102-11

18 Curran B, Jonas D, Grundmann H, *et al.* Development of a multilocus sequence typing scheme for the opportunistic pathogen Pseudomonas aeruginosa. *J Clin Microbiol* 2004;**42**:5644–9. doi:10.1128/JCM.42.12.5644-5649.2004

19 Harris SR, Cartwright EJP, Török ME, *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. doi:10.1016/S1473-3099(12)70268-2

20 Lewis T, Loman NJ, Bingle L, *et al.* High-throughput whole-genome sequencing to dissect the epidemiology of Acinetobacter baumannii isolates from a hospital outbreak. *J Hosp Infect* 2010;**75**:37–41. doi:10.1016/j.jhin.2010.01.012

21 Snitkin ES, Zelazny AM, Thomas PJ, *et al.* Tracking a hospital outbreak of carbapenem-resistant Klebsiella pneumoniae with whole-genome sequencing. *Sci Transl Med* 2012;**4**:148ra116. doi:10.1126/scitranslmed.3004129

22 Loman NJ, Misra RV, Dallman TJ, *et al.* Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol* 2012;**30**:434–9. doi:10.1038/nbt.2198

23 Tredget EE, Shankowsky HA, Joffe AM, *et al.* Epidemiology of infections with Pseudomonas aeruginosa in burn patients: the role of hydrotherapy. *Clin Infect Dis* 1992;**15**:941–9.

24 Langschmidt J, Caine PL, Wearn CM, *et al.* Hydrotherapy in burn care: A survey of hydrotherapy practices in the UK and Ireland and literature review. *Burns* Published Online First: 10 December 2013. doi:10.1016/j.burns.2013.11.006

25 Tredget EE, Shankowsky HA, Rennie R, *et al.* Pseudomonas infections in the thermally injured patient. *Burns* 2004;**30**:3–26.

26 Davison PG, Loiselle FB, Nickerson D. Survey on current hydrotherapy use among North American burn centers. *J Burn Care Res* 2010;**31**:393–9. doi:10.1097/BCR.0b013e3181db5215

27 Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* Published Online First: 28 April 2014. doi:10.1093/bioinformatics/btu170

28 Stewart L, Ford A, Sangal V, *et al.* Draft genomes of 12 host-adapted and environmental isolates of Pseudomonas aeruginosa and their positions in the core genome phylogeny. *Pathog Dis* Published Online First: 26 October 2013. doi:10.1111/2049-632X.12107

29 Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010;**26**:589–95. doi:10.1093/bioinformatics/btp698

30 Koboldt DC, Zhang Q, Larson DE, *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012;**22**:568–76. doi:10.1101/gr.129684.111

31 Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-

likelihood trees for large alignments. *PLoS One* 2010;**5**:e9490. doi:10.1371/journal.pone.0009490

32 Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 2008;**18**:821–9. doi:10.1101/gr.074492.107

33 Jolley KA, Maiden MC. BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* 2010;**11**:595–5. doi:10.1186/1471-2105-11-595

34 Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 2014;**15**:R46. doi:10.1186/gb-2014-15-3-r46

35 Matsen FA, Kodner RB, Armbrust EV. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* 2010;**11**:538. doi:10.1186/1471-2105-11-538

36 Scott FW, Pitt TL. Identification and characterization of transmissible Pseudomonas aeruginosa strains in cystic fibrosis patients in England and Wales. *J Med Microbiol* 2004;**53**:609–15.

37 Panagea S, Winstanley C, Parsons YN, *et al.* PCR-based detection of a cystic fibrosis epidemic strain of Pseudomonas Aeruginosa. *Mol Diagn* 2003;**7**:195–200.

38 Gardy JL, Johnston JC, Ho Sui SJ, *et al.* Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *N Engl J Med* 2011;**364**:730–9. doi:10.1056/NEJMoa1003176

39 Bryant JM, Grogono DM, Greaves D, *et al.* Whole-genome sequencing to identify transmission of Mycobacterium abscessus between patients with cystic fibrosis: a retrospective cohort study. *Lancet* 2013;**381**:1551–60. doi:10.1016/S0140-6736(13)60632-7

40 Lienau EK, Strain E, Wang C, *et al.* Identification of a salmonellosis outbreak by means of molecular sequencing. *N Engl J Med* 2011;**364**:981–2. doi:10.1056/NEJMc1100443

41 Lambert PA. Mechanisms of antibiotic resistance in Pseudomonas aeruginosa. *J R Soc Med* 2002;**95 Suppl 41**:22–6.

42 Sobel ML, Neshat S, Poole K. Mutations in PA2491 (mexS) promote MexT-dependent mexEF-oprN expression and multidrug resistance in a clinical strain of Pseudomonas aeruginosa. *J Bacteriol* 2005;**187**:1246–53. doi:10.1128/JB.187.4.1246-1253.2005

43 Quinn JP, Dudek EJ, DiVincenzo CA, *et al.* Emergence of resistance to imipenem during therapy for Pseudomonas aeruginosa infections. *J Infect Dis* 1986;**154**:289–94.

44 Cao L, Srikumar R, Poole K. MexAB-OprM hyperexpression in NalC-type multidrug-resistant Pseudomonas aeruginosa: identification and characterization of the nalC gene encoding a repressor of PA3720-PA3719. *Mol Microbiol* 2004;**53**:1423–36. doi:10.1111/j.1365-2958.2004.04210.x

45 Llanes C, Hocquet D, Vogne C, *et al.* Clinical strains of Pseudomonas aeruginosa overproducing MexAB-OprM and MexXY efflux pumps simultaneously. *Antimicrob Agents Chemother* 2004;**48**:1797–802.

46 Cambau E, Perani E, Dib C, *et al.* Role of mutations in DNA gyrase genes in ciprofloxacin resistance of Pseudomonas aeruginosa susceptible or resistant to imipenem. *Antimicrob Agents Chemother* 1995;**39**:2248–52.

47 Kelsey M. Pseudomonas in augmented care: should we worry? *J Antimicrob Chemother* 2013;**68**:2697–700. doi:10.1093/jac/dkt288

48 Rogers DE. The changing pattern of life-threatening microbial disease. *N Engl J Med* 1959;**261**:677–83. doi:10.1056/NEJM195910012611401

49 Martin K, Baddal B, Mustafa N, *et al.* Clusters of genetically similar isolates of Pseudomonas aeruginosa from multiple hospitals in the UK. *J Med Microbiol* 2013;**62**:988–1000. doi:10.1099/jmm.0.054841-0

50 Slekovec C, Plantin J, Cholley P, *et al.* Tracking down antibiotic-resistant Pseudomonas aeruginosa isolates in a wastewater network. *PLoS One* 2012;**7**:e49300. doi:10.1371/journal.pone.0049300

51 Snyder LA, Loman N, Faraj LA, *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**.

52 Donlan RM. Biofilm formation: a clinically relevant microbiological process. *Clin Infect Dis* 2001;**33**:1387–92. doi:10.1086/322972

Figure and Table Legends

Figure 1 Legend

An overview of all samples collected during the study in global phylogenetic context with other sequenced strains of *P. aeruginosa* from the set of Stewart *et al*.[28]

Samples collected in this study are widely dispersed in the tree, which contains isolates from different environments (Panel A). Bar plots indicate the numbers of each type of sample collected (Panel B). Microdiversity within each clade is shown, with the colour bar indicating the source of each sample (Panel C).

Figure 2 Legend

A schematic view of the 300-day study of *P. aeruginosa* in a burns centre and critical care unit. Time in days is shown along the X-axis with bed numbers in the critical care unit and burns unit along the Y-axis. Each circular icon indicates a positive isolate of *P. aeruginosa*. The icon's logotype indicates which environment it originated from (wound, urine/sputum, environment or water). The filled colour of the icon indicates the clade it belongs to. Patient icons represent the enrolment of a screening patient into the study and their location. Patient movements around the hospital are noted by dotted lines. The five patients infected with *P. aeruginosa* are denoted by rounded boxes. Boxes are coloured according to the patient number. In the event two or more isolates of the same source and clade were collected on the same day, these have been collapsed into a single circular icon.

Figure 3 Legend

The high-resolution phylogenetic reconstruction of Clade E isolates. This demonstrates the clustering of genotypes by bed space. Patient associated samples are contained within the room 11 clade. This clade contains water samples from the shower and environmental samples from the shower, drain and trolley. The water samples from the room 11 tap are in a distinct clade, indicating the biofilm within the tap has a distinct genotype to the shower. This suggests environmental

contamination was more likely to arise from contaminated shower water than tap water. Details of sampling site, days since start of study and presence of pBURNS plasmids are also shown. The likely phylogenetic position of *P. aeruginosa* detected in a biofilm from a thermostatic mixer valve is shown in the clade associated with room 9 and indicated 'TMV'.

Appendix 1

Phylogenetic tree, sample origin and antibiotic-resistance phenotype for Clade C isolates. The patient column refers to the sampling event. If a sample was collected from a patient colonised with *P. aeruginosa* or that patient's environment the patient number if marked. Patient numbers starting with SP relate to screening patients who were not colonised by *P. aeruginosa*.

Appendix 2

Phylogenetic tree, sample origin and antibiotic-resistance phenotype for Clade D isolates.

Appendix 3

Phylogenetic tree, sample origin and antibiotic-resistance phenotype for Clade E isolates.

Appendix 4

Phylogenetic tree, sample origin and antibiotic-resistance phenotype for Clade G isolates.

Appendix 5

A physical map of the burns unit, indicating individual patient bed areas, shower areas. The water supply is indicated by a blue line.

Appendix 6

Clade E isolates in phylogenetic context with the metagenomics sample from a tap mixer valve.

Appendix 7

Single nucleotide and small indel variants detected within Clade C isolates.

Appendix 8

Single nucleotide and small indel variants detected within Clade D isolates.

Appendix 9

Single nucleotide and small indel variants detected within Clade E isolates.

Appendix 10

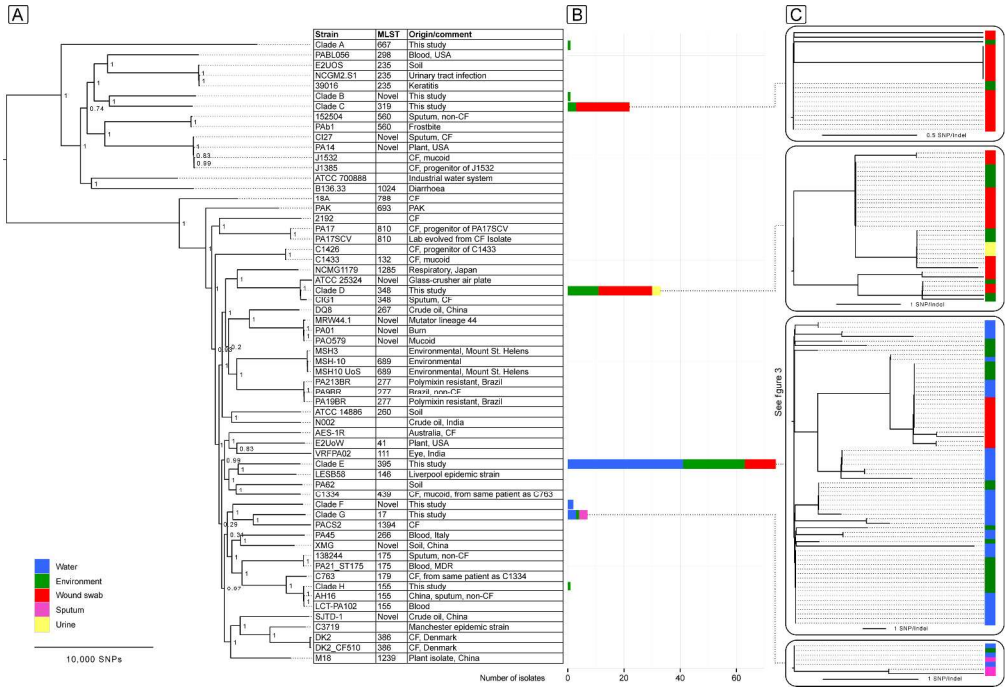
Single nucleotide and small indel variants detected within Clade G isolates.

Appendix 11

The type of antibiotics administered to the five patients in this study.

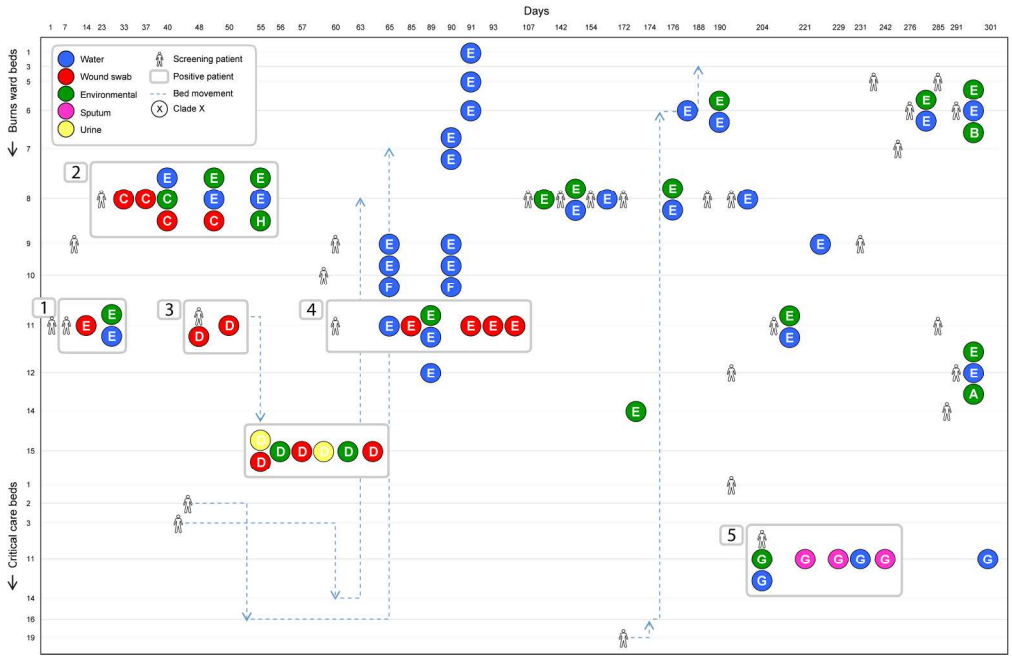
Appendix 12

Mutations predicted to be associated with antibiotic-resistance.

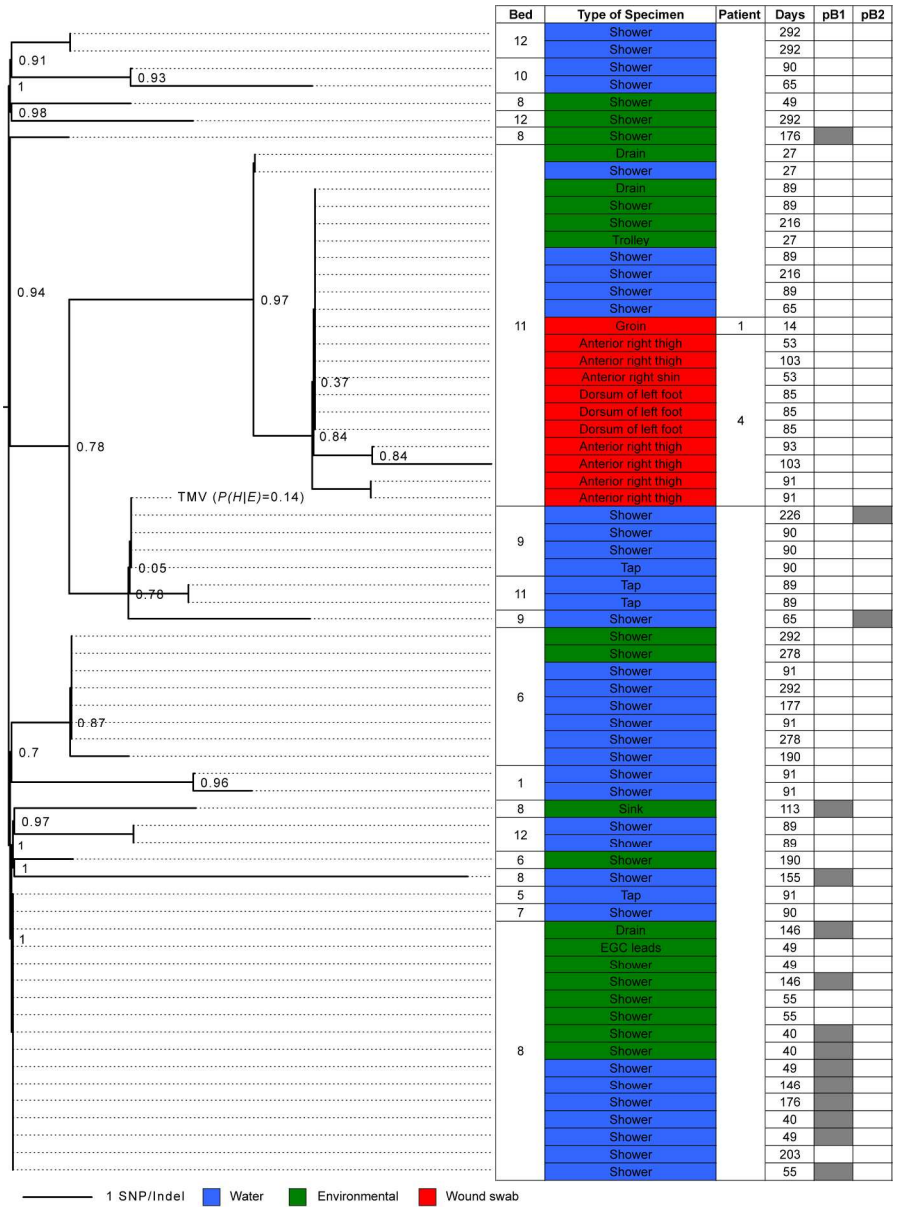


Strain	MLST	Origin/comment
Clade A	667	This study
PABL056	298	Blood, USA
E2UOS	235	Soil
NCGM2.S1	235	Urinary tract infection
39016	235	Keratitis
Clade B	Novel	This study
Clade C	319	This study
152504	560	Sputum, non-CF
PAb1	560	Frostbite
CI27	Novel	Sputum, CF
PA14	Novel	Plant, USA
J1532		CF, mucoid
J1385		CF, progenitor of J1532
ATCC 700888		Industrial water system
B136.33	1024	Diarrhoea
18A	788	CF
PAK	693	PAK
2192		CF
PA17	810	CF, progenitor of PA17SCV
PA17SCV	810	Lab evolved from CF Isolate
C1426		CF, progenitor of C1433
C1433	132	CF, mucoid
NCMG1179	1285	Respiratory, Japan
ATCC 25324	Novel	Glass-crusher air plate
Clade D	348	This study
CIG1	348	Sputum, CF
DQ8	267	Crude oil, China
MRVI44.1	Novel	Mutator lineage 44
PA01	Novel	Burn
PAO579	Novel	Mucoid
MSH3		Environmental, Mount St. Helens
MSH-10	689	Environmental
MSH10 UoS	689	Environmental, Mount St. Helens
PA213BR	277	Polymixin resistant, Brazil
PA9BR	277	Brazil, non-CF
PA19BR	277	Polymixin resistant, Brazil
ATCC 14886	260	Soil
N002		Crude oil, India
AES-1R		Australia, CF
E2UoW	41	Plant, USA
VRFPA02	111	Eye, India
Clade E	395	This study
LESB58	146	Liverpool epidemic strain
PA82		Soil
C1334	439	CF, mucoid, from same patient as C763
Clade F	Novel	This study
Clade G	17	This study
PACS2	1394	CF
PA45	266	Blood, Italy
XMG	Novel	Soil, China
138244	175	Sputum, non-CF
PA21_ST175	175	Blood, MDR
C763	179	CF, from same patient as C1334
Clade H	155	This study
AH16	155	China, sputum, non-CF
LCT-PA102	155	Blood
SJTD-1	Novel	Crude oil, China
C3719		Manchester epidemic strain
DK2	386	CF, Denmark
DK2_CF510	386	CF, Denmark
M18	1239	Plant isolate, China

386x263mm (300 x 300 DPI)

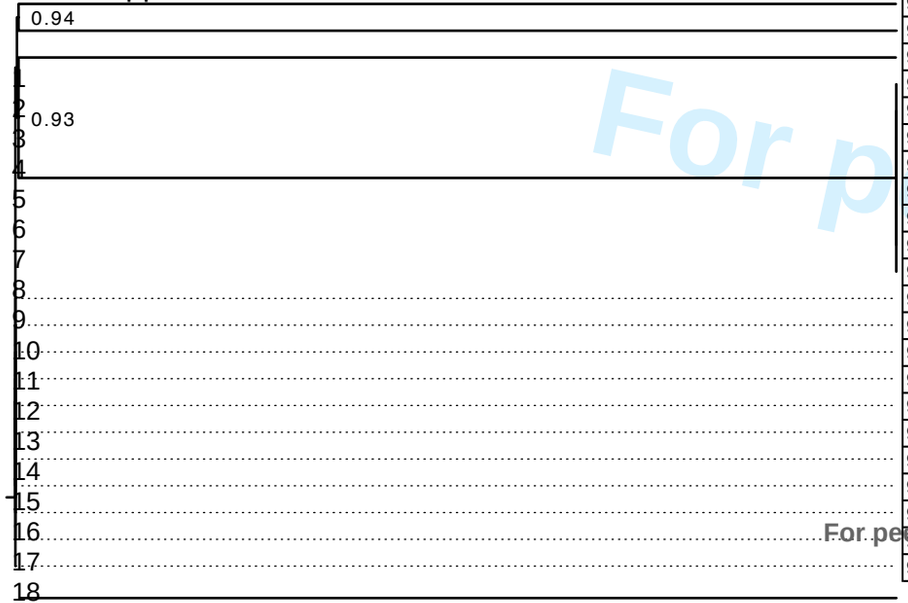


189x122mm (300 x 300 DPI)

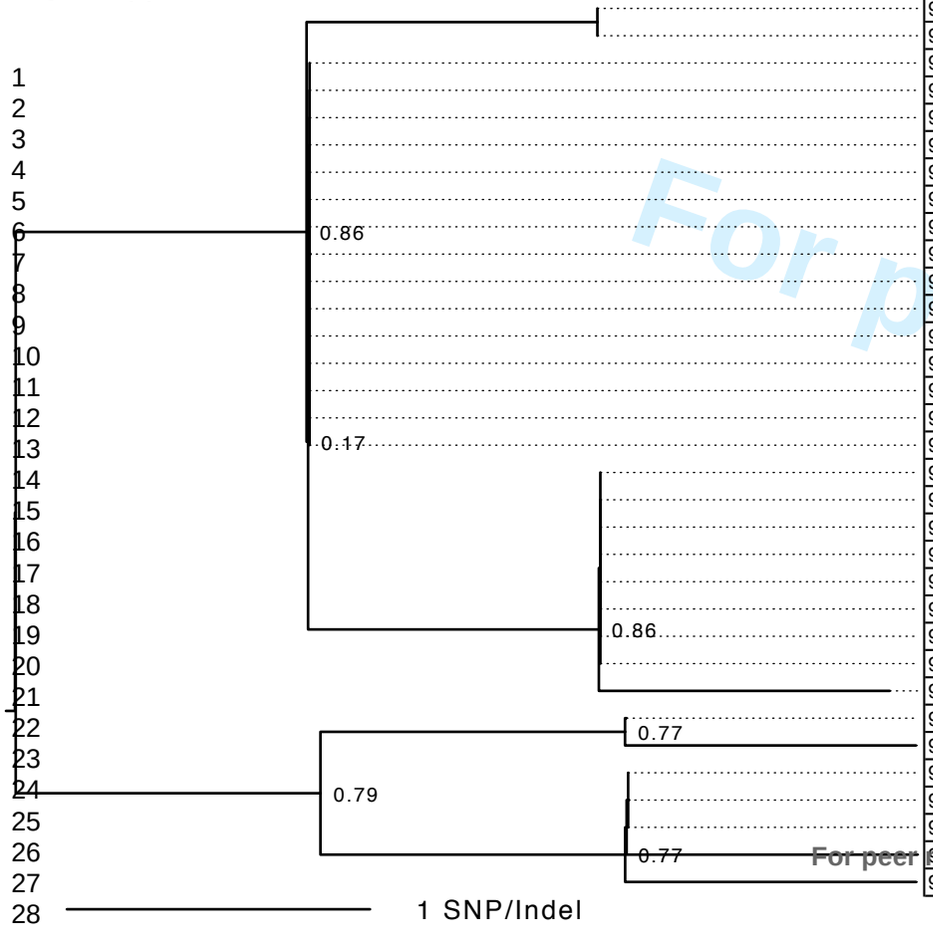


Bed	Type of Specimen	Patient	Days	pB1	pB2
12	Shower		292		
	Shower		292		
10	Shower		90		
	Shower		65		
8	Shower		49		
12	Shower		292		
8	Shower		176		
11	Drain		27		
	Shower		27		
	Drain		89		
	Shower		89		
	Shower		216		
	Trolley		27		
	Shower		89		
	Shower		216		
	Shower		89		
	Shower		65		
	Groin	1	14		
	Anterior right thigh	4	53		
	Anterior right thigh		103		
	Anterior right shin		53		
	Dorsum of left foot		85		
	Dorsum of left foot		85		
	Dorsum of left foot		85		
	Anterior right thigh		93		
	Anterior right thigh		103		
	Anterior right thigh		91		
	Anterior right thigh		91		
9	Shower		226		
	Shower		90		
	Shower		90		
	Tap		90		
11	Tap		89		
	Tap		89		
9	Shower		65		
6	Shower		292		
	Shower		278		
	Shower		91		
	Shower		292		
	Shower		177		
	Shower		91		
	Shower		278		
	Shower		190		
1	Shower		91		
	Shower		91		
8	Sink		113		
12	Shower		89		
	Shower		89		
6	Shower		190		
8	Shower		155		
5	Tap		91		
7	Shower		90		
8	Drain		146		
	EGC leads		49		
	Shower		49		
	Shower		146		
	Shower		55		
	Shower		55		
	Shower		40		
	Shower		40		
	Shower		49		
	Shower		146		
	Shower		176		
	Shower		40		
	Shower		49		
	Shower		203		
	Shower		55		

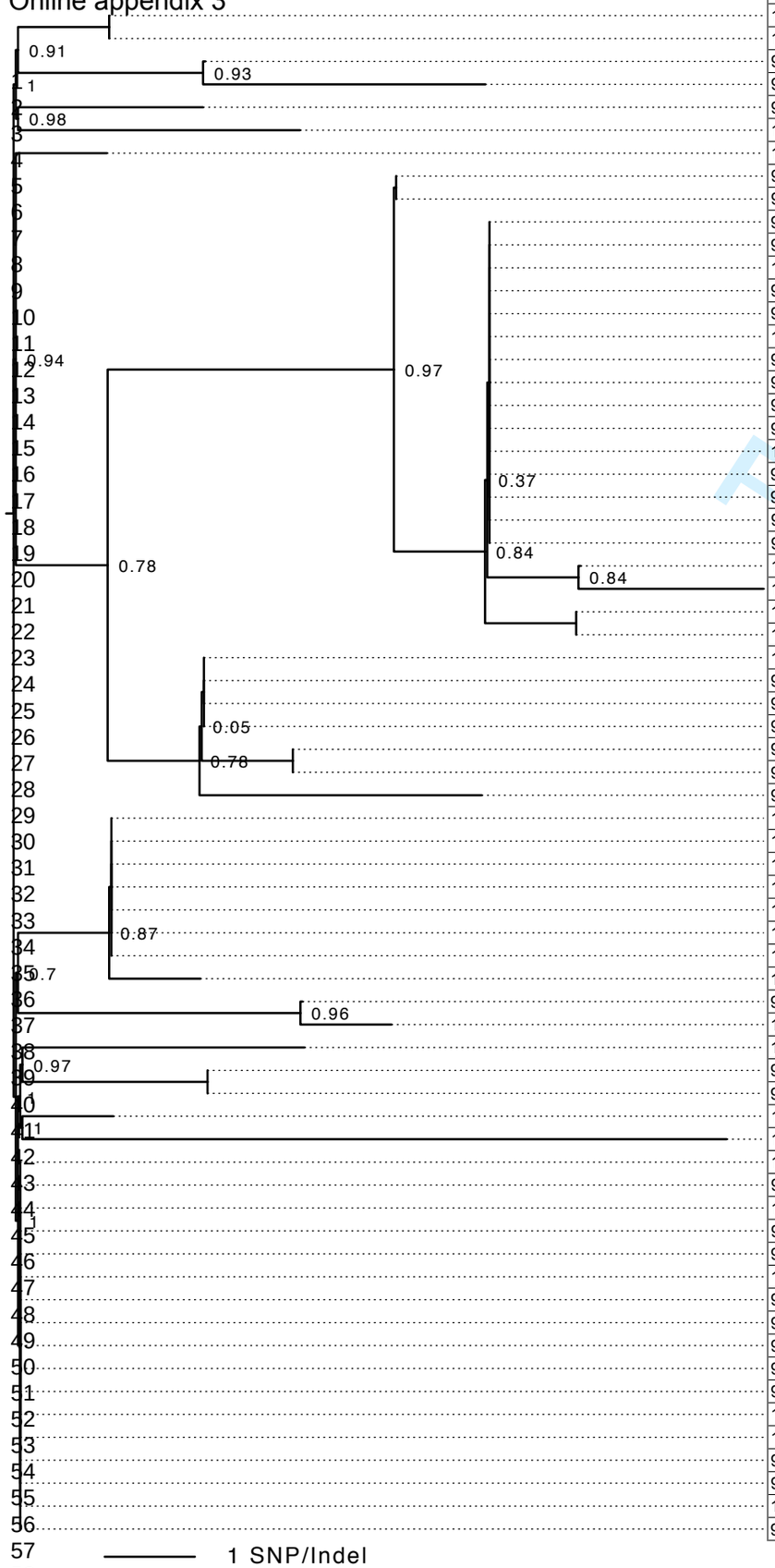
186x250mm (300 x 300 DPI)



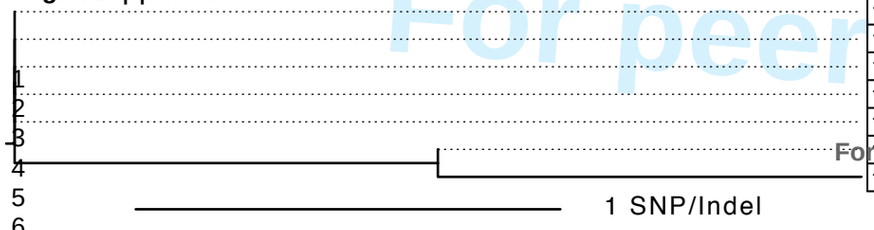
Sample Number	Patient	Ward	Bed	Type of Specimen	Site	Days	TIM	PIP	TAZ	CAZ	AZT	IMI	MER	AMI	GEN	TOB	CIP	COL
922	P02	BCU	8	Wound swab	Upper back	40	R	I	I	S	I	S	S	S	S	S	S	S
921	P02	BCU	8	Wound swab	Upper back	40												
913	P02	BCU	8	Environmental	Shower trolley	40	R	S	S	S	I	S	S	S	S	S	S	S
919	P02	BCU	8	Wound swab	Chest	40	R	I	I	S	I	I	S	S	S	S	R	S
926	P02	BCU	8	Wound swab	Right palm	44	R	I	I	S	I	I	S	S	S	S	R	S
909	P02	BCU	8	Wound swab	Anterior left upper-arm	37	R	I	I	S	I	R	S	S	S	S	R	S
908	P02	BCU	8	Wound swab	Anterior left upper-arm	37	R	I	I	S	I	S	S	S	S	S	S	S
925	P02	BCU	8	Wound swab	Anterior left upper-arm	44	R	I	I	S	I	I	S	S	S	S	R	S
932	P02	BCU	8	Wound swab	Back of head	49	R	I	I	S	I	R	S	S	S	S	R	S
928	P02	BCU	8	Wound swab	Upper back	44	R	I	I	S	I	I	I	S	S	S	R	S
927	P02	BCU	8	Wound swab	Upper back	44	R	I	I	S	I	I	I	S	S	S	R	S
915	P02	BCU	8	Environmental	Chair	40	R	I	I	S	I	S	S	S	S	S	S	S
914	P02	BCU	8	Environmental	Chair	40	R	S	S	S	I	S	S	S	S	S	S	S
904	P02	BCU	8	Tissue	Anterior right upper-arm	33	R	S	S	S	I	S	S	S	S	S	S	S
905	P02	BCU	8	Wound swab	Chest	37	R	I	I	S	I	I	S	S	S	S	S	S
906	P02	BCU	8	Wound swab	Chest	37	R	I	I	S	I	I	S	S	S	S	S	S
920	P02	BCU	8	Wound swab	Chest	40	R	I	I	S	I	R	I	S	S	S	R	S
918	P02	BCU	8	Wound swab	Chest	40	R	I	I	S	I	S	S	S	S	S	S	S
907	P02	BCU	8	Wound swab	Abdomen	37	R	I	I	S	I	S	S	S	S	S	S	S
911	P02	BCU	8	Wound swab	Anterior left forearm	37	R	I	I	S	I	I	S	S	S	S	S	S
912	P02	BCU	8	Wound swab	Anterior left forearm	37	R	I	I	S	I	I	S	S	S	S	S	S
923	P02	BCU	8	Wound swab	Posterior left upper-arm	40	R	I	I	S	I	S	S	S	S	S	S	S



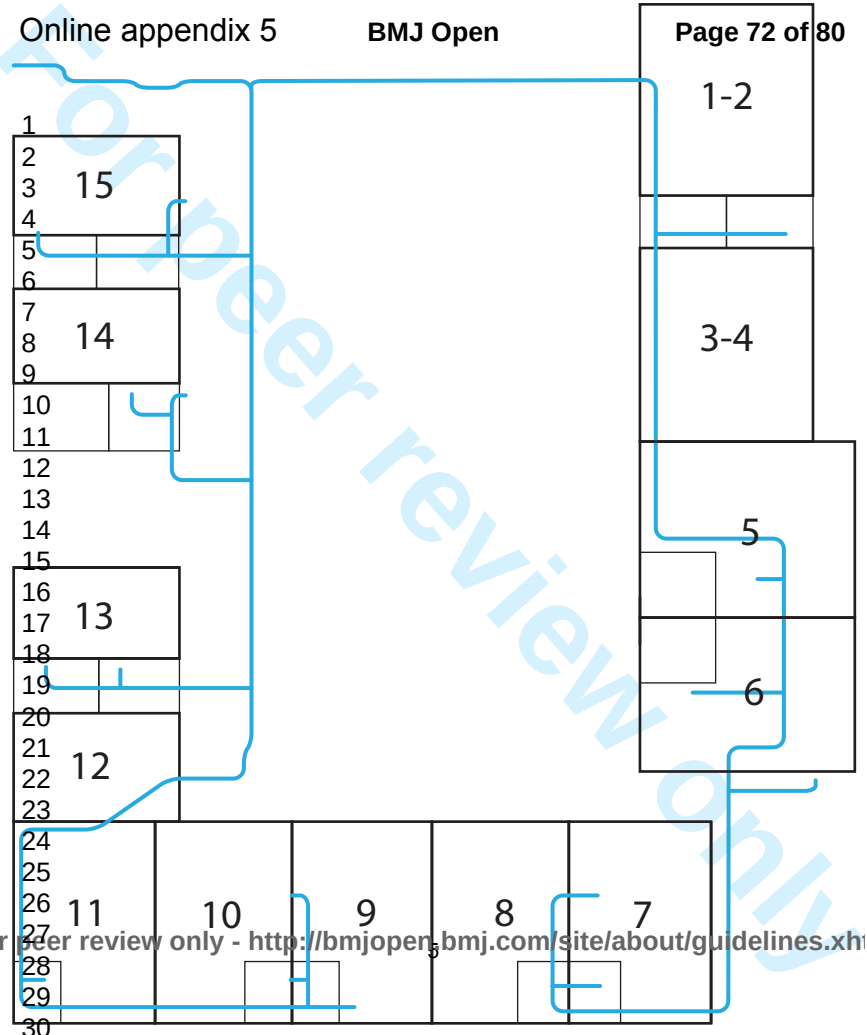
Study Number	Patient	Ward	Bed	Type of Specimen	Site	Days	TIM	PIP	TAZ	CAZ	AZT	IMI	MER	AMI	GEN	TOB	CIP	COL
937	P03	BCU	11	Wound swab	Anterior left thigh	50												
936	P03	BCU	11	Wound swab	Anterior left thigh	50												
929	P03	BCU	11	Wound swab	Posterior left thigh	48	R	S	S	S	I	R	I	S	S	S	S	S
951	P03	BCU	15	Environmental	Bedside table	56	R	I	I	S	I	R	I	S	S	S	S	S
960	P03	BCU	15	Environmental	Bedside table	62	R	S	S	S	I	R	I	S	S	S	S	S
961	P03	BCU	15	Environmental	Door handle	62	R	I	I	S	I	R	I	S	S	S	S	S
958	P03	BCU	15	Environmental	Shower chair	62	R	I	I	S	I	R	I	S	S	S	S	S
957	P03	BCU	15	Environmental	Shower chair	62	R	S	S	S	I	R	I	S	S	S	S	S
940	P03	BCU	15	Wound swab	Anterior right thigh	55	S	S	S	S	I	I	I	S	S	S	S	S
970	P03	BCU	15	Wound swab	Anterior right thigh	64	R	I	I	S	I	R	I	S	S	S	S	S
944	P03	BCU	15	Wound swab	Anterior right shin	55	R	I	I	S	I	R	I	S	S	S	S	S
945	P03	BCU	15	Wound swab	Anterior right shin	55	R	S	S	S	I	R	I	S	S	S	S	S
946	P03	BCU	15	Wound swab	Anterior left thigh	55	R	S	S	S	I	R	I	S	S	S	S	S
954	P03	BCU	15	Wound swab	Anterior left thigh	57	R	I	I	S	I	R	I	S	S	S	S	S
947	P03	BCU	15	Wound swab	Anterior left shin	55	R	S	S	S	I	R	I	S	S	S	S	S
972	P03	BCU	15	Wound swab	Posterior left shin	64	S	S	S	S	I	R	I	S	S	S	S	S
949	P03	BCU	15	Wound swab	Posterior right thigh	55	R	I	I	S	I	R	I	S	S	S	S	S
950	P03	BCU	15	Environmental	Bedside table	56	R	S	S	S	I	R	I	S	S	S	S	S
959	P03	BCU	15	Environmental	Bedside table	62	R	I	I	S	I	R	I	S	S	S	S	S
953	P03	BCU	15	Environmental	Toilet flush	56	R	S	S	S	I	I	I	S	S	S	S	S
938	P03	BCU	15	Urine	Urine	55	R	S	S	S	I	R	I	S	S	S	S	S
939	P03	BCU	15	Urine	Urine	55	R	S	S	S	I	R	I	S	S	S	S	S
956	P03	BCU	15	Urine	Urine	58	R	S	S	S	I	R	I	S	S	S	S	S
969	P03	BCU	15	Wound swab	Anterior right thigh	64	R	S	S	S	I	R	I	S	S	S	S	S
971	P03	BCU	15	Wound swab	Anterior left thigh	64	R	I	I	S	I	R	I	S	S	S	S	S
973	P03	BCU	15	Wound swab	Posterior right thigh	64	R	S	S	S	I	R	I	S	S	S	S	S
955	P03	BCU	15	Wound swab	Anterior left thigh	57	S	S	S	S	I	R	I	S	S	S	S	S
975	P03	BCU	15	Wound swab	Posterior right shin	64	S	S	S	S	I	R	I	S	S	S	S	S
962	P03	BCU	15	Environmental	Shower chair	62	S	S	S	S	I	R	I	S	S	S	S	S
948	P03	BCU	15	Wound swab	Posterior right thigh	55	S	S	S	S	I	R	I	S	S	S	S	S
974	P03	BCU	15	Wound swab	Posterior right thigh	64	S	S	S	S	I	I	I	S	S	S	S	S
963	P03	BCU	15	Environmental	Bedside table	62	S	S	S	S	I	I	I	S	S	S	S	S
964	P03	BCU	15	Environmental	Bedside table	62	S	S	S	S	I	R	I	S	S	S	S	S

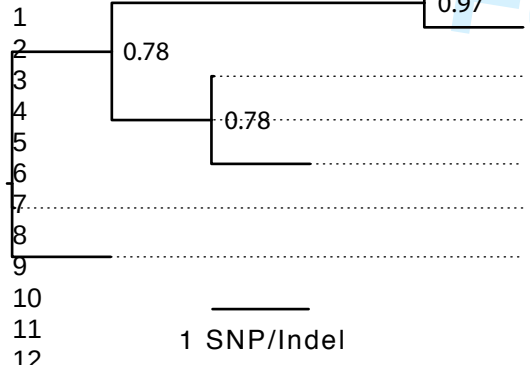


Sample Number	Patient	Ward	Bed	Type of Specimen	Site	Days	TIM	PIP	TAZ	CAZ	AZT	IMI	MER	AMI	GEN	TOB	CIP	COL	pBURNS1	pBURNS2
1067	SP30	BCU	12	Water	Shower (Pre-flush)	292														
1068	SP30	BCU	12	Water	Shower (Pre-flush)	292														
991	Water sampling	BCU	10	Water	Shower (Pre-flush)	90														
966	SP08	BCU	10	Water	Shower (Unknown)	65	R	S	S	S	I	S	S	S	S	S	S	S		
931	P02	BCU	8	Environmental	Shower (Rose)	49	R	S	S	S	I	S	S	S	S	S	S	S		
1065	SP30	BCU	12	Environmental	Shower (Rose)	292														
1034	SP15	BCU	8	Environmental	Shower (Rose)	176	R	S	S	S	R	S	S	S	S	S	S	S	Yes	
902	P01	BCU	11	Environmental	Drain	27														
903	P01	BCU	11	Water	Shower (Post-flush)	27	R	S	S	S	I	S	S	S	S	S	S	S		
981	P04	BCU	11	Environmental	Drain	89	R	S	S	S	I	S	S	S	S	S	S	S		
980	P04	BCU	11	Environmental	Shower (Rose)	89	R	S	S	S	I	S	S	S	S	S	S	S		
1048	SP20	BCU	11	Environmental	Shower (Rose)	216	R	S	S	S	I	S	S	S	S	S	S	S		
901	P01	BCU	11	Environmental	Trolley	27	R	S	S	S	I	S	S	S	S	S	S	S		
987	P04	BCU	11	Water	Shower (Hose)	89	R	S	S	S	I	S	S	S	S	S	S	S		
1050	SP21	BCU	11	Water	Shower (Post-flush)	216														
988	P04	BCU	11	Water	Shower (Pre-flush)	89	R	S	S	S	I	S	S	S	S	S	S	S		
968	SP10	BCU	11	Water	Shower (Unknown)	65	R	S	S	S	I	S	S	S	S	S	S	S		
900	P01	BCU	11	Wound swab	Groin	14	R	S	S	S	I	S	S	S	S	S	S	S		
979	P04	BCU	11	Wound swab	Anterior right thigh	53	R	S	S	S	I	S	S	S	S	S	S	S		
1008	P04	BCU	11	Wound swab	Anterior right thigh	103	R	S	S	S	I	S	S	S	S	S	S	S		
976	P04	BCU	11	Wound swab	Anterior right shin	53	R	S	S	S	I	S	S	S	S	S	S	S		
984	P04	BCU	11	Wound swab	Dorsum of left foot	85	R	S	S	S	I	R	I	S	S	S	S	S		
983	P04	BCU	11	Wound swab	Dorsum of left foot	85	R	S	S	S	I	R	I	S	S	S	S	S		
982	P04	BCU	11	Wound swab	Dorsum of left foot	85	R	S	S	S	I	S	S	S	S	S	S	S		
1007	P04	BCU	11	Wound swab	Anterior right thigh	93	R	S	S	S	I	R	I	S	S	S	S	S		
1009	P04	BCU	11	Wound swab	Anterior right thigh	103	R	S	S	S	I	S	S	S	S	S	S	S		
1005	P04	BCU	11	Wound swab	Anterior right thigh	91	R	S	S	S	I	R	I	S	S	S	S	S		
1006	P04	BCU	11	Wound swab	Anterior right thigh	91	R	S	S	S	I	R	I	S	S	S	S	S		
1052	SP22	BCU	9	Water	Shower (Hose)	226														Yes
993	Water sampling	BCU	9	Water	Shower (Hose)	90	R	S	S	S	I	S	S	S	S	S	S	S		
994	Water sampling	BCU	9	Water	Shower (Pre-flush)	90	R	S	S	S	I	S	S	S	S	S	S	S		
992	Water sampling	BCU	9	Water	Tap	90	R	S	S	S	I	S	S	S	S	S	S	S		
985	P04	BCU	11	Water	Tap	89	R	S	S	S	I	S	S	S	S	S	S	S		
986	P04	BCU	11	Water	Tap	89														
967	SP09	BCU	9	Water	Shower (Unknown)	65	R	S	S	S	I	S	S	S	S	S	S	S		Yes
1062	SP29	BCU	6	Environmental	Shower (Rose)	292														
1058	SP25	BCU	6	Environmental	Shower (Rose)	278														
1003	Water sampling	BCU	6	Water	Shower (Hose)	91														
1064	SP29	BCU	6	Water	Shower (Pre-flush)	292														
1036	SP15	BCU	6	Water	Shower (Pre-flush)	177														
1004	Water sampling	BCU	6	Water	Shower (Pre-flush)	91														
1057	SP25	BCU	6	Water	Shower (Pre-flush)	278	R	S	S	S	I	S	S	S	S	S	S	S		
1041	SP16	BCU	6	Water	Shower (Pre-flush)	190														
999	Water sampling	BCU	1	Water	Shower (Hose)	91														
1000	Water sampling	BCU	1	Water	Shower (Pre-flush)	91														
1010	SP11	BCU	8	Environmental	Sink	113	S	S	S	S	I	S	S	S	S	S	S	S	Yes	
989	Water sampling	BCU	12	Water	Shower (Hose)	89	R	S	S	S	I	S	S	S	S	S	S	S		
990	Water sampling	BCU	12	Water	Shower (Pre-flush)	89	R	S	S	S	I	S	S	S	S	S	S	S		
1040	SP16	BCU	6	Environmental	Shower (Rose)	190	R	S	S	S	I	S	S	S	S	S	S	S		
1031	SP13	BCU	8	Water	Shower (Hose)	155													Yes	
1001	Water sampling	BCU	5	Water	Tap	91														
996	Water sampling	BCU	7	Water	Shower (Pre-flush)	90	R	S	S	S	I	S	S	S	S	S	S	S		
1029	SP12	BCU	8	Environmental	Drain	146													Yes	
933	P02	BCU	8	Environmental	EGC	49	R	S	S	S	I	S	S	S	S	S	S	S		
930	P02	BCU	8	Environmental	Shower (Rose)	49	R	S	S	S	I	R	S	S	S	S	S	S		
1028	SP12	BCU	8	Environmental	Shower (Rose)	146													Yes	
942	P02	BCU	8	Environmental	Shower (Rose)	55	R	S	S	S	I	S	S	S	S	S	S	S		
941	P02	BCU	8	Environmental	Shower (Rose)	55	R	S	S	S	I	S	S	S	S	S	S	S		
917	P02	BCU	8	Environmental	Shower (Rose)	40	S	S	S	S	I	S	S	S	S	S	S	S	Yes	
916	P02	BCU	8	Environmental	Shower (Rose)	40	S	S	S	S	I	S	S	S	S	S	S	S	Yes	
934	P02	BCU	8	Water	Shower (Post-flush)	49	S	S	S	S	I	S	S	S	S	S	S	S	Yes	
1030	SP12	BCU	8	Water	Shower (post-flush)	146													Yes	
1035	SP14	BCU	8	Water	Shower (Post-flush)	176													Yes	
910	P02	BCU	8	Water	Shower (Post-flush)	40	S	S	S	S	I	S	S	S	S	S	S	S	Yes	
935	P02	BCU	8	Water	Shower (Pre-flush)	49	R	S	S	S	I	S	S	S	S	S	S	S	Yes	
1046	SP19	BCU	8	Water	Shower (Pre-flush)	203														
943	P02	BCU	8	Water	Shower (Pre-flush)	55	S	S	S	S	I	S	S	S	S	S	S	S	Yes	



Sample Number	Patient	Ward	Bed	Type of Specimen	Site	Days	TIM	PIP	TAZ	CAZ	AZT	IMI	MER	AMI	GEN	TOB	CIP	COL
1069	Water sampling	WCCB	11	Water	Tap	301	R	S	S	S	I	R	I	S	S	S	S	S
1045	SP20	WCCB	11	Environmental	Tap handle	204												
1047	SP20	WCCB	11	Water	Tap	204	R	S	S	S	I	R	I	I	R	S	S	S
1049	P05	WCCB	11	Sputum	Sputum	221		I	I	S	I	I	I	I	R	S	S	S
1053	P05	WCCB	11	Water	Tap	231	R	I	I	S	I	R	S	S	S	S	S	S
1054	P05	WCCB	11	Sputum	Sputum	229	R	R	R	R	R	R	R	I	R	S	I	S
1056	P05	WCCB	11	Sputum	Sputum	242		R	R	R	S	R	R		S	S	R	S



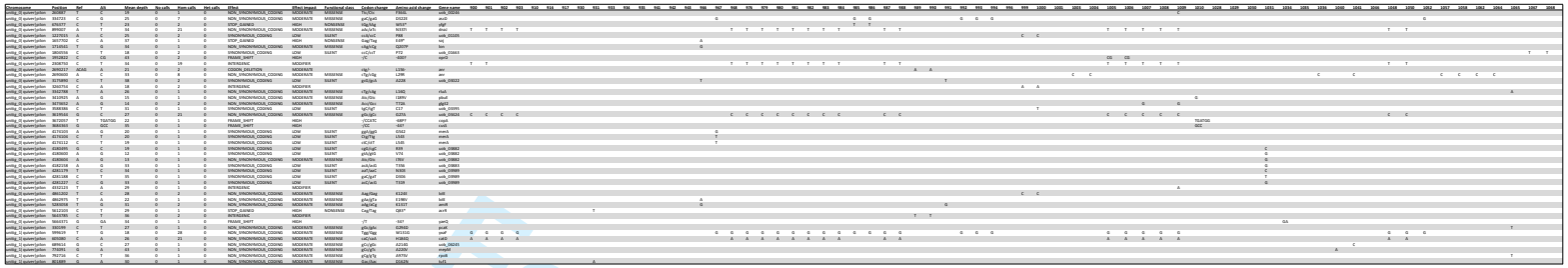


Bed	Type of Specimen	Alignment
11	Shower	TCCTACTCC-CACAGACCTAACCT-----G--ACCGAAATCCTTTTCCG-CGAGGCG
11	Shower	TCCTACTCC-TACAGACCTAACCT-----G--ACCGAAATCCTTTTCCG-CGAGGCG
9	Tap	TGCAACTCC-CACAGACCTAACGT-----G--ACCGAAATCCTTTTCCG-CGCGGCG
9	TMV	TGCA-CT-C-CACAGA---AACGT-----G--ACCG--A-CCTTTTC-G-C--GGCG
11	Tap	TGTAACTCC-CACAGACCTAACGT-----G--ACCGAAATCCTTTTCCG-CGCGGCG
8	Shower	TCCAACTCC-CACAGACCTAACGT-----G--ACCGAAATCCTTTTCCG-CTCGGCG
6	Shower	TCCAACTCC-CACAGCCTAACGT-----G--ACCGAAATCCTTTTCCG-CTCGGCG

Chromosome	Position	Ref	Alt	Mean depth	No calls	Hom calls	Het calls	Effect	Effect impact	Functional class	Codon change	Amino acid change	Gene name	904	905	906	907	908	909	911	912	913	914	915	918	919	920	921	922	923	925	926	927	928	932
NC_002516	1558800	CCATATG	C	40	0	1	0	CODON_DELETION	MODERATE		catatg/-	HM211-	lasR																C						
NC_002516	2806409	G	A	23	0	8	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	Cac/Tac	H321Y	PA2491					A	A							A					A	A	A	A	A
NC_002516	3804666	GCTTGC	G	19	0	1	0	FRAME_SHIFT	HIGH		-	-72	PA3399															G							
NC_002516	4148397	A	T	21	0	1	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	aTt/aAt	I181N	wspD									T													

Online appendix 8

Chromosome	Position	Ref	Alt	Mean depth	No calls	Hom calls	Het calls	Effect	Effect impact	Functional class	Codon change	Amino acid change	Gene name	929	936	937	938	939	940	944	945	946	947	948	949	950	951	953	954	955	956	957	958	959	960	961	962	963	964	969	970	971	972	973	974	975
NC_002516	574989	CCGCAGTTG	C	20	0	9	0	FRAME_SHIFT	HIGH			-37	PA0512				C	C								C		C			C			C					C			C				
NC_002516	856001	TGGCCTGG	T	31	0	1	0	FRAME_SHIFT	HIGH			-242	putP																								T									
NC_002516	1558520	C	T	41	0	5	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	cCg/cTg	P117L	lasR											T														T	T	T						T		
NC_002516	1558668	C	CTT	36	0	2	0	FRAME_SHIFT	HIGH			-/TT	-167?	lasR																	CTT																CTT
NC_002516	2844173	A	C	19	0	26	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	gAc/gCc	D119A	PA2523	C	C	C	C	C	C	C	C	C	C			C	C	C	C	C		C	C	C	C	C	C			C	C	C	C	C			
NC_002516	2926290	G	A	25	0	7	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	cCg/cTg	P48L	gacA											A							A						A	A	A						A	A	
NC_002516	3486673	A	T	25	0	1	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	cTg/cAg	L171Q	PA3106																																			T
NC_002516	3584085	G	A	23	0	2	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	gCc/gTc	A87V	gltR		A	A																															
NC_002516	5796903	C	T	14	0	1	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	gCc/gTc	A317V	mutY																															T				
NC_002516	5853846	A	C	26	0	1	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	Acc/Ccc	T88P	PA5201																												C								



Chromosome	Position	Ref	Alt	Mean depth	No calls	Hom calls	Het calls	Effect	Effect impact	Functional class	Codon change	Amino acid change	Gene name	1045	1047	1049	1053	1054	1056	1069
NC_002516	3558951	G	A	35	0	1	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	aCc/aTc	T83I	gyrA						A	
NC_002516	4166773	A	G	22	0	2	0	NON_SYNONYMOUS_CODING	MODERATE	MISSENSE	Acg/Gcg	T86A	nalC					G	G	

Patient	Antibiotic/antifungal	Start	End	Start	End
2	Ciprofloxacin	16/10/2012	06/11/2012		
	Nitrofurantoin	16/10/2012	11/11/2012		
	Vancomycin	29/10/2012	03/11/2012		
4	Flucloxacillin	23/11/2012	30/11/2012		
	Piperacillin/tazobactam	30/11/2012	03/12/2012		
	Meropenem	03/12/2012	08/12/2012		
	Colistin	15/12/2012	21/12/2012		
5	Gentamycin	12/04/2013	12/04/2013		
	Co-amoxiclav	13/04/2013	14/04/2013		
	Erthromycin	14/04/2013	21/04/2013	11/05/2013	19/05/2013
	Piperacillin/tazobactam	14/04/2013	18/04/2013		
	Meropenem	20/04/2013	08/05/2013	19/05/2013	20/05/2013
	Caspofungin	26/04/2013	14/05/2013	20/05/2013	21/05/2013
	Linezolid	01/05/2013	12/05/2013		
	Ciprofloxacin	06/05/2013	16/05/2013	20/05/2013	21/05/2013
	Colistin	20/05/2013	23/05/2013		

Gene	Mutation type	Effect	AA substitution	Resistance phenotype	Samples
gyrA	SNP	non-synonymous	T83I	Ciprofloxacin	1056
nalC	indel	-	-	Meropenem	1054, 1056
mexS	SNP	non-synonymous	H321Y	Ciprofloxacin	908, 909, 919, 925-928, 932
oprD	indel	frame shift	-400?	Imipenem/meropenem	1005, 1006

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cohort studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	3
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	3-4
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	6
Objectives	3	State specific objectives, including any prespecified hypotheses	7
Methods			
Study design	4	Present key elements of study design early in the paper	8-9
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	8
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	8
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	n/a
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	8-9
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	8-9
Bias	9	Describe any efforts to address potential sources of bias	n/a
Study size	10	Explain how the study size was arrived at	9
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	n/a
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	11-13
		(*b*) Describe any methods used to examine subgroups and interactions	n/a
		(*c*) Explain how missing data were addressed	n/a
		(*d*) If applicable, explain how loss to follow-up was addressed	n/a
		(*e*) Describe any sensitivity analyses	n/a
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	10
		(b) Give reasons for non-participation at each stage	10
		(c) Consider use of a flow diagram	-
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	10-11
		(b) Indicate number of participants with missing data for each variable of interest	10
		(c) Summarise follow-up time (eg, average and total amount)	n/a
Outcome data	15*	Report numbers of outcome events or summary measures over time	10-11
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	n/a
		(*b*) Report category boundaries when continuous variables were categorized	n/a
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	n/a
Discussion			
Key results	18	Summarise key results with reference to study objectives	16
Limitations			
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	16-18
Generalisability	21	Discuss the generalisability (external validity) of the study results	18
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	19

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.