Author/s:
Sherry, NL;Gorrie, CL;Kwong, JC;Higgs, C;Stuart, RL;Marshall, C;Ballard, SA;Sait, M;Korman, TM;Slavin, MA;Lee, RS;Graham, M;Leroi, M;Worth, LJ;Chan, HT;Seemann, T;Grayson, ML;Howden, BP

Title:
Multi-site implementation of whole genome sequencing for hospital infection control: A prospective genomic epidemiological analysis

Date:
2022-06-01

Citation:
Sherry, N. L., Gorrie, C. L., Kwong, J. C., Higgs, C., Stuart, R. L., Marshall, C., Ballard, S. A., Sait, M., Korman, T. M., Slavin, M. A., Lee, R. S., Graham, M., Leroi, M., Worth, L. J., Chan, H. T., Seemann, T., Grayson, M. L. & Howden, B. P. (2022). Multi-site implementation of whole genome sequencing for hospital infection control: A prospective genomic epidemiological analysis. Lancet Regional Health Western Pacific, 23, https://doi.org/10.1016/j.lanwpc.2022.100446.

Persistent Link:
https://hdl.handle.net/11343/306918

The Lancet Regional Health - Western Pacific

Multi-site implementation of whole genome sequencing for hospital infection control: A prospective genomic epidemiological analysis

--Manuscript Draft--

Manuscript Number:	TLRHWESTERNPACIFIC-D-21-00618R2
Article Type:	Article (Original Research)
Keywords:	Antimicrobial resistance; Whole genome sequencing; infection prevention and control; hospital epidemiology
Corresponding Author:	Benjamin Howden University of Melbourne Melbourne, Not Listed or Not Applicable Australia
First Author:	Norelle L Sherry
Order of Authors:	Norelle L Sherry
	Claire L Gorrie
	Jason C Kwong
	Charlie Higgs
	Rhonda L Stuart
	Caroline Marshall
	Susan A Ballard
	Michelle Sait
	Tony Korman
	Monica A Slavin
	Robyn SC Lee
	Maryza Graham
	Marcel Leroi
	Leon J Worth
	Hiu Tat Chan
	Torsten Seemann
	M Lindsay Grayson
	Benjamin Howden
Manuscript Region of Origin:	AUSTRALIA
Abstract:	Background Current microbiological methods lack the resolution to accurately identify multidrug-resistant organism (MDRO) transmission, however, whole genome sequencing can identify highly-related patient isolates providing opportunities for precision infection control interventions. We investigated the feasibility and potential impact of a prospective multi-centre genomics workflow for hospital infection control. Methods We conducted a prospective genomics implementation study across eight Australian hospitals over 15 months (2017-2018), collecting all clinical and screening isolates from inpatients with vanA VRE, MRSA, ESBL Escherichia coli (ESBL-Ec), or ESBL Klebsiella pneumoniae (ESBL-Kp). Genomic and epidemiologic data were integrated to assess MDRO transmission. Findings In total, 2275 isolates were included from 1970 patients, predominantly ESBL-Ec (40·8%) followed by MRSA (35·6%), vanA VRE (15·2%), and ESBL-Kp (8·3%). Overall, hospital and genomic epidemiology showed 607 patients (30·8%) acquired their MDRO in hospital, including the majority of vanA VRE (266 patients, 86·4%), with lower proportions of ESBL-Ec (186 patients, 23·0%), ESBL-Kp (42 patients,

	26·3%), and MRSA (113 patients, 16·3%). Complex patient movements meant the majority of MDRO transmissions would remain undetected without genomic data. The genomics implementation had major impacts, identifying unexpected MDRO transmissions prompting new infection control interventions, and contributing to vanA VRE becoming a notifiable condition. We identified barriers to implementation and recommend strategies for mitigation. Interpretation Implementation of a multi-centre genomics-informed infection control workflow is feasible and identifies many unrecognised MDRO transmissions. This provides critical opportunities for interventions to improve patient safety in hospitals.

Reviewer Comment	Author's response
REVIEWER 3 COMMENTS	
In their rebuttal letter, the authors "acknowledge that the structure of this manuscript is a departure from that of a traditional scientific paper or clinical trial", and they view their paper as "a description of the design and implementation of a new genomics-informed workflow". My concern is that the authors state that they have performed statistical analysis (using Stata). They also describe methods used for statistical inference (e.g. the chi-squared test) and several "significant" findings. However, the purpose of the statistical analysis is very unclear. It is easy to get the impression that the authors confuse inference and description, and readers not familiar with clinical microbiology or hospital infection control practices can be misled about the findings' level of evidence. Therefore, my recommendation to the authors is to explain the statistical analysis' purpose and the statistical approach. The authors should also improve the presentation of significant findings by specifying whether the word significant refers to inferential uncertainty (statistical significance) or practical relevance (clinical significance). Several ambiguities currently exist in the manuscript, for example the claim that "significant impacts on infection control and public health were demonstrated".	Thank you for raising this important point. We address this point in the following way: Comparative statistics are important for some aspects of this study and are applied as follows: - To compare the numbers of patients with probable or possible transmission between different species, to identify whether these differences were statistically significant or more likely observed by chance; and - To compare the MDRO transmission rates between different hospital networks, to identify whether these differences were statistically significant or more likely observed by chance. The finding of differences in MDRO transmission rates for VREfm between networks (both statistically significant and clinically important) allowed the IPC team to approach the hospital executive committee with evidence of increased VREfm transmission in their network. This led to a network-wide change in cleaning practices (a very large undertaking), with attendant increases in costs borne by the hospital. In our view, this demonstrates the potential power of genomics to inform hospital IPC practices. Whilst the statistical analysis plays a small part in this, it is important to maintain scientific rigor (particularly if sample numbers are small) to ensure that chance differences are not being misconstrued as true differences in transmission rates. It is therefore important to preserve this statistical analysis in the manuscript. We do acknowledge that the use of the word 'significant' can be ambiguous, and could lead to some confusion when applied to statistical tests and clinical scenarios. We have therefore modified the manuscript to use alternative terms and remove the word "significant", except when referring to the statistically significant differences in MDRO transmission rates. (Modifications: Lines 63, 280, 327, 374, and Table 5)

Editor's Comments	Author's response
Please follow our "Information for Authors" carefully when revising the manuscript. I will email you the PDF in a separate email	
Please ensure your revised paper conforms to any applicable reporting standards as detailed in the "Information for Authors". A checklist of the relevant reporting standard is required. STROBE for Observational studies CONSORT - for RCTs PRISMA - For meta-analyses and systematic reviews For other study types, please refer to: http://www.equator-network.org/	STROME-ID checklist completed and included in submission
Please use British English throughout the manuscript	Complete
Please include a "Research in context" section, with the following subheadings: • Evidence before this study • Added value of this study • Implications of all the available evidence	Complete
The Article title should include a study descriptor—eg, randomised trial, case-control study, prospective analysis, population-based study. Titles should be non-declamatory (ie, not state the findings of the paper).	Complete
Please provide full names for all authors in the authorship of the manuscript. For example, Michael A. Morpurgo, NOT M.A. Morpurgo or Morpurgo M.A.	Complete
Please provide a structured "Abstract", with the following subheadings: Background Methods Findings Interpretation Funding	Complete

Please use the following main headings for your manuscript, in this order: • Introduction • Methods (with "Role of the funding source" at the end, stating whether the funders had any role in study design, data collection, data analysis, interpretation, writing of the report) • Results • Discussion • Contributors: detailing which author did what in this paper • Declaration of interests: the statements should match those in the ICMJE COI form. Note that you do not have to provide new forms if you have already provided them. • Acknowledgements : if there are people to be acknowledged, please ask them to provide a signed agreement (word file or email), along the line of "I agree to be included in the Acknowledgements of this article [ARTICLE TITLE]. Signed, dated, full name") • Data sharing statement: Kindly add this statement to your manuscript. Please add a Data Availability statement to the end of the manuscript that explains where readers can access the data and provides any necessary accession codes. If there are restrictions to the availability of any materials or data, these must be disclosed here. For more information on Data Sharing, please see our Information for Authors here. • References (Vancouver numbering style)	Section headings complete in this order Acknowledgements collated and included in submission
Please upload Figures as individual files (one Figure per file), in either EPS, PDF, TIFF or JPEG format. Note that the figures should be at least 300dpi. Images that have been published previously should be accompanied by a statement indicating permission to reproduce the image. If all the figures are your own and have not been published before, this requirement does not apply.	All original figures Figures submitted in separate files
Tables should be supplied in a separate Word file (not Excel or fdf/pdf). Each row of data should be in a separate line. Please ensure that rows and columns are not tabbed; data should be entered in cell form.	Table submitted in separate files
If you intend to file a patent related to the findings presented in this manuscript, please ensure your patent application is submitted before submitting your revised manuscript	N/A
Please complete the "Author Statements Form". I will email you the PDF in a separate email. The corresponding author should complete the form along with their signature/initials on the last page (lower left side), and all authors should sign the last page. Note that you do not have to provide new forms if you have already provided them.	Author statement forms collated and included in submission **Please note**: additional author form included for Dr Suraya Hanim Abdullah Hashim (study group member), who was unable to return her forms prior to the deadline for the first revision due to severe COVID.
Each individual author should complete one "ICMJE COI Form". I will also email you the PDF in a separate email. The disclosure in the ICMJE COI form should match EXACTLY the "Declarations of interests" statement in the manuscript.	COI forms included in submission

TECHNICAL INFORMATION: When you submit the revised paper, please provide the following: 1. One "clean" copy of your manuscript 2. One copy where your changes are highlighted (tracked changes). 3. A separate, point by point response to the editorial and referee comments, clearly indicating where in the manuscript changes have been made. 4. Any figures, tables, supplementary files (even if no revisions have been made). 5. Kindly provide a word file named “caption for supplementary material” with the list of captions for each Supplementary file as they should appear online. 6. One completed "Author Declaration Form" signed by all authors on the last page. 7. One completed "ICMJE COI Form" for each author. Please do NOT include a copy of your original manuscript. All text files should be supplied as MS Word files. Please also supply the word count for the body of your paper and your abstract (word count for the body of your paper should not include abstract, references, figures or tables).	Complete
Please also check whether you need to provide the following: • Signed copyright permissions for previously published material • Signed consent from individuals cited in the Acknowledgements • Signed consent for use of cited personal communications • Signed patient's consent and permission to publish (if not already submitted)	 N/A Included N/A N/A

Research in Context

Evidence before the study

We searched PubMed for studies using the search terms “genomics”, “antimicrobial resistance”, “hospital”, infection control”, and “transmission” without date or language restrictions, and identified 262 results. We considered all studies from this search that used whole genome sequencing to assess likely transmission of multidrug-resistant organisms (MDROs) in the hospital environment, published before 5th September 2021. We excluded reviews, case reports and retrospective cohort analyses and outbreak investigations. We also excluded studies of fungi, mycobacteria, viruses, resistance to specific antimicrobial agents/classes not included in our study (e.g. carbapenems, colistin, linezolid), and studies not employing whole genome sequencing. Of the remaining nine studies with prospective sample collection, five focused on single species or resistance mechanisms (MRSA, *E. coli*, *Enterococcus faecium*/VRE, *Klebsiella pneumoniae*), one focused on VRE and multi-resistant Gram negatives, and two focussed on four different species. Study durations varied from three to 13 months; five studies included 1-2 wards, two studies included one hospital, and one study included five hospitals (single organism only). Only one study discussed results being reported to hospital infection control teams, or any changes made in response to genomic data; this was a pre- and post-intervention surveillance study to assess the effects of removing isolation requirements for patients colonised with MDR Gram negatives. No studies reported prospective genomic epidemiological investigations across multiple species across multiple hospital sites.

Added value of this study

This study is the first to implement and evaluate prospective genomic sequencing of more than one target MDRO across multiple institutions. We designed and implemented a study using prospective genomic surveillance for four MDROs, *vanA* VRE, MRSA, ESBL *Escherichia coli* (ESBL-Ec), and ESBL *Klebsiella pneumoniae* (ESBL-Kp), from all hospital inpatients across eight healthcare facilities over a period of 15 months. We combined genomic and epidemiologic data to identify the number of MDRO transmissions that occurred within and between hospitals, including up to 12 months prior to first MDRO identification. We demonstrated that *vanA* VRE was most likely to be acquired due to transmissions in hospitals. Hospital transmission rates for other MDROs (ESBL-Ec, ESBL-Kp and MRSA) were also significantly underestimated prior to the study, however colonisation with these pathogens is relatively common in our community. We identified key components of a prospective genomics workflow for successful implementation, and recommend mitigation strategies for challenges identified for implementation.

Implications of all available evidence

This study addresses the lack of evidence for the prospective implementation of whole genome sequencing for uncovering and responding to MDRO transmissions across multiple healthcare sites. Taken together, studies of whole genome sequencing of MDROs in hospitals demonstrate the enormous potential for precision infection control interventions at the hospital level to be guided by this data. The next steps in transitioning from research to implementation require close engagement between hospital infection control teams, diagnostic microbiology laboratories and laboratories with expertise in genome sequencing and analysis, to create workflows that are efficient and sustainable in different settings, and accurately and rapidly communicate results to maximise the benefits to hospitals and patients. Our study also demonstrates that integrated data across multiple sites can inform higher level public health decision making regarding the notification and systematic surveillance of MDRO threats. Future implementation of whole genome sequencing for hospital infection control should consider and integrate a multi-layered approach (hospital and network/state level) to ensure that the data can inform precision infection control interventions at the hospital level as well as inform system-wide decision making.

Title: Multi-site implementation of whole genome sequencing for hospital infection control: A prospective genomic epidemiological analysis

Brief title: Prospective genomics for hospital infection control

Authors: Norelle L Sherry MBBS[1,2,3], Claire L Gorrie PhD[1,3], Jason C Kwong PhD[2,3,4], Charlie Higgs BSc[3], Rhonda L Stuart PhD[5,6,7], Caroline Marshall PhD[8,9], Susan A Ballard PhD[1], Michelle Sait PhD[1], Tony M Korman MBBS[5,10], Monica A Slavin MD[11,12], Robyn S Lee PhD[3†], Maryza Graham MBBS[5,10], Marcel Leroi MBBS[2,13], Leon J Worth PhD[11.12], Hiu Tat Chan MBBS[14‡], Torsten Seemann PhD[1], M Lindsay Grayson PhD[2,4,13]* & Benjamin P Howden PhD[1,2,3]* on behalf of the Controlling Superbugs Study Group

*authors contributed equally

[†] Current affiliation: Epidemiology Division, Dalla Lana School of Public Health, University of Toronto, Toronto, Canada

[‡] Current affiliation: Department of Molecular Microbiology, VCS Pathology, Australian Centre for Prevention of Cervical Cancer, Carlton, Victoria, Australia

[1] Microbiological Diagnostic Unit (MDU) Public Health Laboratory, Department of Microbiology & Immunology at the Peter Doherty Institute for Infection & Immunity, University of Melbourne, Melbourne, Victoria, Australia

[2]Department of Infectious Diseases, Austin Health, Heidelberg, Victoria, Australia

[3]Department of Microbiology & Immunology at the Peter Doherty Institute for Infection & Immunity, University of Melbourne, Melbourne, Victoria, Australia

[4]Department of Medicine, Austin Health, University of Melbourne, Heidelberg, Victoria, Australia

[5]Monash Infectious Diseases, Monash Health, Clayton, Victoria, Australia

[6]Monash University, Clayton, Victoria, Australia

[7]South East Public Health Unit, Monash Health, Clayton, Victoria, Australia

[8]Infection Prevention & Surveillance, Victorian Infectious Diseases Service, Melbourne Health, Parkville, Victoria, Australia

[9]Department of Infectious Diseases, The University of Melbourne at the Peter Doherty Institute for Infection & Immunity, Melbourne, Victoria, Australia

[10]Department of Microbiology, Monash Health, Clayton, Victoria, Australia

[11]Department of Infectious Diseases, Peter MacCallum Cancer Centre, Parkville, Victoria, Australia

[112]National Centre for Infections in Cancer, Sir Peter MacCallum Department of Oncology, University of Melbourne, Melbourne, Victoria, Australia

[13]Department of Microbiology, Austin Health, University of Melbourne, Heidelberg, Victoria, Australia

[14]Department of Microbiology, Melbourne Health, Parkville, Victoria, Australia

Corresponding author: Prof. Benjamin Howden, MDU Public Health Laboratory, Department of Microbiology & Immunology at the Peter Doherty Institute for Infection and Immunity, University of Melbourne. Email: bhowden@unimelb.edu.au.

Manuscript word count: 4504 words (excluding abstract, references, figures and tables)

Abstract word count: 250 words

Abstract (250 words)

Background

Current microbiological methods lack the resolution to accurately identify multidrug-resistant organism (MDRO) transmission, however, whole genome sequencing can identify highly-related patient isolates providing opportunities for precision infection control interventions. We investigated the feasibility and potential impact of a prospective multi-centre genomics workflow for hospital infection control.

Methods

We conducted a prospective genomics implementation study across eight Australian hospitals over 15 months (2017-2018), collecting all clinical and screening isolates from inpatients with *vanA* VRE, MRSA, ESBL *Escherichia coli* (ESBL-Ec), or ESBL *Klebsiella pneumoniae* (ESBL-Kp). Genomic and epidemiologic data were integrated to assess MDRO transmission.

Findings

In total, 2275 isolates were included from 1970 patients, predominantly ESBL-Ec (40·8%) followed by MRSA (35·6%), *vanA* VRE (15·2%), and ESBL-Kp (8·3%).

Overall, hospital and genomic epidemiology showed 607 patients (30·8%) acquired their MDRO in hospital, including the majority of *vanA* VRE (266 patients, 86·4%), with lower proportions of ESBL-Ec (186 patients, 23·0%), ESBL-Kp (42 patients, 26·3%), and MRSA (113 patients, 16·3%). Complex patient movements meant the majority of MDRO transmissions would remain undetected without genomic data.

The genomics implementation had major impacts, identifying unexpected MDRO transmissions prompting new infection control interventions, and contributing to *vanA* VRE becoming a notifiable condition. We identified barriers to implementation and recommend strategies for mitigation.

Interpretation

Implementation of a multi-centre genomics-informed infection control workflow is feasible and identifies many unrecognised MDRO transmissions. This provides critical opportunities for interventions to improve patient safety in hospitals.

Funding

Melbourne Genomics Health Alliance (supported by State Government of Victoria, Australia), and National Health and Medical Research Council (Australia).

Introduction

Antimicrobial resistance is a serious threat to human health, and disproportionately affects hospitalised patients, with many multidrug-resistant organisms (MDROs) acquired directly from other hospital patients or indirectly from the hospital environment.[1-3] Whilst these MDROs can usually be adequately detected by routine diagnostic microbiology methods, these methods are insufficient to identify transmission chains, as multiple strains of each MDRO may be circulating in the hospital and the community.[4,5]

Whole genome sequencing (WGS, or "genomics") of MDROs provides high-resolution typing data to identify transmission networks between patients and/or the environment.[6-8] Genomics has increasingly been used in the research setting to identify hospital MDRO outbreaks, most often in a retrospective manner after a possible outbreak has already been identified by epidemiologic surveillance.[9-11] However, the greatest potential for genomics for hospital infection control is to *prospectively* identify likely transmission of MDROs between a small number patients, allowing early infection control interventions before an outbreak has become established,[12,13] therefore limiting patient morbidity, mortality, unnecessary antibiotic use, and waste of hospital resources.[14-16] Although WGS costs have been steadily decreasing over the last decade, implementation should be directed to the most cost-effective targets with the greatest benefits to health and resources.[17,18]

Translation of genomics from the research setting to the real-world infection control setting requires consideration of the elements that are likely to be critical for successful implementation.[5,17] Factors to consider include an optimised genomic workflow model to achieve clinically meaningful turnaround times, which patient populations and MDROs are likely to yield the greatest benefit from genomics, and how to effectively communicate complex genomic analyses to infection control teams at the coalface, to enable implementation of infection control interventions for maximal benefit to the patients and healthcare system.[4,5]

Here we present a study implementing a prospective genomics workflow to detect transmission of MDROs across multiple hospital networks over 15 months, with all MDRO samples from hospital inpatients being sequenced at a central laboratory, and results returned to the hospital infection control team for detailed epidemiological analysis and further management. We aim to establish the feasibility of such a workflow and potential impacts, and identify key components and potential barriers to implementation.

Methods

Study scoping and design

As a scoping exercise to design a genomics workflow for infection control, we first listed the *components* to be included, identified the *key considerations* for each component, then identified a *strategy for implementation* for each component, taking into account existing local infrastructure, MDRO patterns and resources (staffing, consumables, and financial resources), before settling on a study design for local implementation to establish the feasibility and potential impacts of infection control genomics.

We subsequently performed a prospective "real-world" multi-centre genomics implementation study of MDRO transmission in eight hospitals from four hospital networks, including approximately 2800 acute and subacute inpatient beds, in Melbourne, Australia (population 4·9 million in 2018); hospital characteristics are described in Table 1. MDRO isolates from four defined species (see below) were included from patient samples (clinical or screening samples) collected routinely from hospital inpatients (>24h) at any time during their admission. Duplicate screening isolates were excluded, and duplicate clinical isolates were only included if collected more than 14 days after the previous sample (see Supplementary Data for details).

The study was conducted in two phases: a pilot phase (24th April – 18th June 2017), and an implementation phase (30th October 2017 – 30th November 2018), totalling 15 months. Results of transmission analyses were only reported to hospitals during the implementation phase. Pilot phase results have previously been reported,[19] but are included here as isolates from these patients may be implicated in future transmission events detected in the implementation phase. Outbreak investigation support (including sequencing of environmental and relevant historical isolates) was also offered to participating sites during the study.

MDRO definitions

Four principal MDROs were included in the pilot phase, chosen as AMR pathogens of global significance (WHO global priority pathogen list[20]) with known transmission in hospitals; the selected MDROs were *vanA* vancomycin-resistant *Enterococcus faecium* (*vanA* VREfm), methicillin-resistant *Staphylococcus aureus* (MRSA), and extended-spectrum beta-lactamase phenotype *Escherichia coli* and *Klebsiella pneumoniae* (ESBL-Ec and ESBL-Kp) (see Supplementary data for details). Carbapenem-resistant *Acinetobacter baumannii* complex (CRAb) and carbapenem-resistant *Pseudomonas aeruginosa* (CRPa) were also included in the pilot phase,[21] but excluded here due to low prevalence. Carbapenem-resistant Enterobacterales (CPE) were excluded, as these were already

under genomic surveillance as part of the state-wide CPE surveillance program.[22] As the pathogens included in the study were not notifiable by legislation, no specific public health interventions or reporting was in place that the time, and infection control units managed these independently.

Due to the large volume of ESBL-Ec included during the pilot phase, we modified inclusion criteria for the implementation phase to only include ESBL-Ec that were also resistant to fluoroquinolones, as this was more consistent with existing infection control practices at participating sites.

Infection control, clinical and epidemiologic data

Hospital infection control practices and MDROs screening protocols at each site were assessed regularly during the study, including any changes made in response to genomic data (Table 1 and Supplementary data). Basic demographic and isolate data were collected for all patients; epidemiologic data were collected for a subset of patients where genomics suggested potential MDRO transmission to assess the clinical likelihood of such cross-transmission having occurred. Transmission rates were estimated by calculating the number of putative transmissions per 100,000 occupied bed days (OBDs, inpatient beds occupied on daily bed count, excluding mental health).

Laboratory workflows

MDROs were isolated, worked up and reported by the hospital laboratories as per their usual protocols. For patients and isolates meeting inclusion criteria, a pure subculture was sent to the central laboratory for sequencing and isolate storage.

The sequencing laboratory is an ISO-accredited laboratory with high-throughput genomic sequencing capacity, located centrally between the participating hospitals. Once samples arrived, a single colony was subcultured onto horse blood agar, incubated overnight, then 1-2 colonies selected and placed into lysis buffer. DNA extraction, NexteraXT library preparation and quality control (QC) were performed as previous described,[19] and isolates sequenced on Illumina NextSeq (San Diego, CA, USA).

Bioinformatics and transmission analysis

An overview of the combined genomic and epidemiologic analysis workflow to identify MDRO transmission can be found in Figure 1, with details included in Supplementary Data. Briefly, all sequences and assemblies underwent routine quality control checks, and underwent *de novo* genome assembly, species identification, multi-locus sequence typing (MLST), and identification of antimicrobial resistance (AMR) genes.

Transmission analyses were performed within sequence type (ST) for all STs with two or more patients. Reads were aligned to a complete reference genome of the same ST, and recombinant sites were masked (except for *S. aureus*). Variant sites were detected from short reads for each ST, and pairwise single nucleotide polymorphism (SNP) differences were calculated for all isolate pairs.

To screen for potential MDRO transmission, we defined isolate pairs with ≤25 core SNP differences (*E. coli, K. pneumoniae,* and *E. faecium*) or ≤15 core SNP differences (MRSA) as 'genomically related'. Isolate pairs with different key AMR genes were excluded from further analysis (i.e. *mec* type for MRSA, and ESBL or AmpC gene (allele) for ESBL-Ec and ESBL-Kp; *vanA* alleles were excluded as this was a requirement for study inclusion). Patients with isolates that were 'genomically related' then had further epidemiologic data collected, detailing their admission history (dates, hospitals and wards) for 12 months before the earliest isolate until the end of the study period.

Epidemiologic data were then compared for patients above the screening thresholds ('genomically related') within ST (or ST131 *E. coli* subclades). Patient pairs were classified according to a modification of previously published definitions[23] as 'probable transmission' if on the same ward at the same time, 'possible transmission' if admitted to the same ward at a different time (within 60 days), or admitted to the same hospital at the same time; all other patients were classified as 'unlikely transmission'. Statistical analyses were performed using Stata v16.

Reporting to infection control teams and report design

During the implementation phase of the study, genomic results for each hospital network were compiled and reported in person to infection control teams, with epidemiologic data where available; written reports were then supplied after each presentation. As this was the first time that most infection control teams had been exposed to genomic data, the reporting presentations included an informal education component to familiarise teams with genomic analyses and interpretation of hospital's genomic results. The study team also sought feedback from infection control teams during and after each reporting session, and made themselves available for any questions from infection control teams between reports or presentations.

Using the feedback from the reporting sessions, two focus group sessions were held at the end of the study with infectious diseases physicians, infection control practitioners, and medical microbiologists to collect feedback to design a genomics report for infection control use (see Supplementary Methods for details). A final report was designed based on feedback from the focus group sessions.

Ethics approval

This study was approved by the Melbourne Health Human Research Ethics Committee (HREC) and endorsed by the corresponding HREC at each participating site.

Results

Scoping and designing a workflow for implementation of genomics-informed infection control

To design a genomics workflow for infection control, we considered the critical elements for implementation: target population, target MDROs, MDRO screening strategy, sample types, clinical data collection, sequencing and analysis strategy, bioinformatics analysis, and communication of results (Table 2). The final workflow, as implemented in this study, is shown in Figure 1.

Isolate and patient characteristics

Overall, 2641 bacterial isolates were submitted by hospital laboratories for WGS; 175 were excluded due to patients not meeting inclusion criteria, 116 isolates were excluded as duplicates, and a further 64 isolates did not meet microbiological inclusion criteria (see Supplementary Figure S1 for details).

Ultimately, 2275 isolates from 1970 patients were included for analysis, predominantly ESBL-Ec (929 isolates, 47·2%), followed by MRSA (811 isolates, 41·2%), *vanA* VRE (346 isolates, 17·6%), and ESBL-Kp (189 isolates, 9·4%). Most isolates were collected for clinical reasons (62 ·8% overall), although this varied according to species and screening programs at each hospital network (Figure 2). After screening swabs (37·1%), urine (20·4%), and swabs from non-sterile sites (20·4%) were the most common sample types, with only 6·9% of isolates being from blood cultures (Supplementary Figure S2). Median patient age was 65 (IQR 47-77), and 56·8% of patients were male. 297 patients (15 ·1%) had more than one isolate included (range 1-7 isolates)(Table 3). For patients with one screening and at least one clinical isolate with the same MDRO, median time between screening and first clinical isolate was 13 days (range 0-365 days, IQR 4-55 days)(Supplementary Table S2).

vanA *VREfm was transmitted more frequently in hospital than other MDROs in this study*

Initial genomic analysis identified that 844 (42·8%) of isolates were genomically-related to at least one other isolate from the study (i.e. below screening pairwise SNP threshold, excluding isolates from same patient, and excluding key AMR gene mismatch) (Supplementary Table S3 and Supplementary Figure S3).

Aligning the patient admission history of each pair of patients with genomically-related isolates (see workflow Figure 1), MDRO isolates from 354 (18·0%) patients were identified as probable

transmission in hospital, 253 patients (12·8%) as possible transmission in hospital, and 236 (12.0%) as unlikely acquired in hospital (Figure 3a, and pairwise SNP distributions shown Supplementary Figure S4). This varied significantly by species: 266 *vanA* VREfm patients (86·4%) were identified as probable or possible transmission in hospital, whereas lower proportions of patients with ESBL-Ec, ESBL-Kp, and MRSA were identified as probable or possible in-hospital transmission (186 ESBL-Ec patients (23·0%), 113 MRSA patients (16·3%), 42 ESBL-Kp patients (26·3%); χ^2=488 ·32 (VREfm vs other species), p<0 ·001). Based on combined genomic and epidemiologic analysis, 44·8% of transmissions were thought to have occurred on a previous admission during the study period (up to 12 months prior to first sample collection), whilst another 17·5% of transmissions were likely to have occurred on a previous ward during the same admission as sample collection; only 34·7% of transmissions were likely to be from the same ward and admission when the MDRO sample was collected.

Hospital networks varied in the MDRO species with the highest transmission rates

Transmission rates per 100,000 OBDs varied widely between networks and species (Figure 3b). The greatest difference between networks was observed in ESBL-Ec, with network A having 47·5 transmissions per 100,000 OBDs, compared with the other two large networks (network B, 4·9 and network C, 5·4 transmissions per 100,000 OBDs) and zero transmissions for network D (incidence difference between network A and networks B and C, 43·14 transmissions per 100,000 OBDs (95% CI 37·54-48·75, p<0·001)). This may partly reflect differences in screening practices (see Table 1), especially the network-wide point-prevalence surveys for Network A which contributed 27·2% of their transmissions (18·5% of ESBL-Ec, 62·9% of *vanA* VREfm and 11·1% of ESBL-Kp).

Conversely, network B registered a higher rate of *vanA* VRE transmission than other networks (transmissions per 100,000 OBDs: network B, 34·2; network A, 22·3; network C, 14·3; network D, 2·4; incidence difference between network B and networks A and C, 15·69 transmissions per 100,000 OBDs (95% CI 9·54-21·84, p<0·001)). MRSA transmission was also higher for network B, but these data likely underestimate the true MRSA transmission rates for all networks due to low rates of routine MRSA screening.

Successful vanA *VREfm clones spread widely within and between hospital networks*

Eight genomic clusters of *vanA* VREfm were detected in more than one hospital network (Figure 4); some were spread across networks from the commencement of the study (e.g. ST1421 cluster 1), whilst others emerged during the study, initially at one site, then other sites (e.g. ST1424 cluster 2). Whilst this study was not designed to capture transmissions between hospitals, and the state health

information systems are decentralised (limiting tracking of patient movement between hospitals), this strongly suggests that patient movement between networks also contributes to the dissemination of *vanA* VRE in study sites, and likely more widely in Victoria.

Impacts of genomics program

The implementation of genomic surveillance for hospital MDROs led to a wide range of impacts, falling into two main groups: identifying unexpected MDRO transmissions requiring targeted infection control interventions, or supporting the management of existing outbreaks through the provision of high-resolution genomic data to clearly define transmission networks (Table 4). The genomics program was particularly useful in identifying MDRO outbreaks occurring over a longer period of time, which had not been detected by conventional surveillance methods. Outbreak investigations were supported by including genomic analysis of historical and environmental isolates (including shared patient equipment), allowing infection control teams to trace the likely transmission pathways, and implement targeted measures to prevent further spread. This was especially useful where epidemiological links were complex, such as when a patient may have been exposed to *vanA* VREfm in more than one location, making attribution of MDRO acquisition impossible with epidemiologic data alone. Furthermore, genomics allowed the cessation of screening in one *vanA* VREfm outbreak, by demonstrating that new cases outside this ward were not related to the original outbreak, thus saving considerable resources and allowing the ward to re-open to new admissions.

This study also had the unanticipated effect of encouraging data sharing between study sites, resulting in identification of one network with a higher rate of *vanA* VREfm transmissions than the others. This led to a review of cleaning practices, which were subsequently changed for the whole network towards the end of the study period. Additionally, data from this study also informed public health actions, with *vanA* VREfm newly added to the list of notifiable conditions by the state health department[24].

Workflow implementation issues

A number of challenges for program implementation were identified throughout the study, primarily those causing delays in turnaround times or difficulty accessing and integrating electronic patient data (Table 5). Many of these issues relate specifically to this study, and would not be expected to limit the future implementation of a genomics workflow for infection control.

Whilst several factors contributed to delays in turnaround times (TAT), two factors dominated; i) delays in hospital laboratory sending isolates to the sequencing laboratory (mostly due to unusual

workflow disruptions in the hospital laboratories), and ii) delays in reporting results to hospitals, as we elected to report results in person (to encourage engagement and education), rather than releasing results cumulatively (Supplementary Figure S5). Established systems for rapid referral, sequencing and reporting of notifiable pathogens in our jurisdiction, such as carbapenemase-producing Enterobacterales (CPE) and *Salmonella,* are usually complete in 7 to 10 days. Once MDRO infection control sequencing is fully embedded using routine referral processes, it is anticipated that these delays experienced for some samples in this this study would be unlikely to occur.

Communication

To design an effective written report format for infection control genomics, feedback from focus group participants at the conclusion of the study was used to design a template for reporting aggregated hospital/network-level genomics data (e.g. monthly report for each network)(Figure 5 and Supplementary Figure S6). Recurrent themes included a desire for simplified results, strong preference for use of a 'traffic light' system to classify likelihood of transmission, general preference to avoid inclusion of phylogenetic trees (although not amongst more genomics-experienced participants), and specific comments directing ICPs to which patients to investigate for possible transmission.

Discussion

Here we have presented a large, multicentre study implementing a new workflow for prospective genomic surveillance of MDROs to inform infection control interventions in hospitals. We have demonstrated wide-ranging impacts of prospective genomic surveillance, identifying MDRO transmission where it was not suspected, providing support for outbreak investigations, and data to inform public health coordination of hospital-acquired infections. We believe that this is a critical step towards changing how we think about MDROs in hospitals, no longer relying on low-resolution traditional microbiological and epidemiologic surveillance, but proactively detecting MDRO transmission at a genomic level. This then enables early and tailored infection control interventions, potentially limiting the number of patients infected, morbidity and mortality, and saving limited hospital resources.[4,18,25,26] This is most relevant with more common MDROs found in both hospitals and the community, as hospital transmission may be under-estimated without precision genomic data.[4] Importantly, most of the MDRO transmission detected here would not have been identified by existing infection control surveillance, offering huge opportunities to detect silent MDRO transmission, and intervene.

Implementing a new workflow represents a major change for hospital laboratories and infection control teams, and identifies challenges to be overcome for successful implementation. One of the pressure points for successful implementation is minimising turnaround times, critical for timely and effective infection control responses. COVID-19 has demonstrated what TATs for sequencing may be achievable, with 5-7 days TATs routinely attained in our state (1-3 days for urgent samples).[27] With adequate resourcing and buy-in from hospital laboratories, and capacity in sequencing laboratories, similar TATs could be achieved for MDRO surveillance. Alternatively, a decentralised genomics model (local WGS and analysis, or hybrid model with centralised analysis) may also optimise TATs; the choice of model ultimately depends on the available resources in each setting.[28] Other key challenges include the optimisation of bioinformatics processes[29] and integration of epidemiologic data, without which the genomic analyses are not interpretable.[30,31] In our study, the linkage of genomic and epidemiologic data was impaired by poor integration with electronic medical records; this integration would be an important component for successful implementation in the future.[32]

Communication of complex genomic results to a new clinical setting is another challenge for implementation. Ideally, results should be easy to comprehend, and easily integrated with epidemiologic data to facilitate rapid interpretation, enabling rapid infection control intervention where required.[33] As an early implementation study in a group of hospitals unfamiliar with genomics, we chose to engage infection control teams by presenting results in-person to educate and answer questions interactively, with the trade-off of delayed reporting times. Whilst this approach was very useful for early implementation and engagement, alternative reporting methods would be more suitable going forward, such as regular written reports designed for infection control (such as the example provided here). Further innovation to allow infection control teams rapid access to genomics results and interpretation would be welcome in future implementation studies.

Identifying strong epidemiologic evidence of transmission is difficult, particularly with common MDROs, but necessary to validate genomic inferences of MDRO transmission. In this study, we used genomic links (SNP thresholds) and epidemiologic data (overlapping admissions in space and time) to call probable or possible transmission, similar to other studies.[25,34] SNP thresholds are a relatively blunt instrument to suggest potential genomic links, but core genome SNP analyses are accepted to be far higher resolution than existing typing methods (such as MLST), and at least equivalent to core genome MLST.[35,36] Importantly, SNP thresholds are specific to the methods used to derive them, hence other studies using lower thresholds are not directly comparable to this study.[32,34] In a pragmatic sense, if high-resolution core genome SNP methods are used for large-scale prospective genomics, then SNP thresholds are likely to be required; however, future studies in this area should be open to exploring alternative methods and thresholds.[29,34,35,37]

Our study has a number of limitations. Firstly, screening practices were not standardised between hospitals, and included one site performing a biannual hospital-wide point-prevalence survey. However, many transmissions were still identified at hospitals performing less screening; this suggests the potential advantages of additional screening to uncover unsuspected transmissions. Secondly, bed movement data would ideally have been acquired for all study patients; this was addressed in the pilot study and did not greatly change the interpretation of the transmission data.[19] Third, these methods are not designed to detect transmission of AMR plasmids mechanisms, as short-read sequencing methods are limited in this ability.

Limited access to data (clinical data and centralised state admissions data) was also a limitation of the study, and largely prevented identification of inter-hospital transmission; with the increasing implementation of electronic medical records, we hope that these issues could be more comprehensively addressed in future studies.[30] Turnaround times were also impacted by extraneous factors (laboratory delays); despite this, major impacts on infection control and public health were demonstrated, and future changes in resource management and system design have now been proposed to improve TATs. Lastly, the cost-effectiveness of this intervention has not been addressed in this study, but demonstrated elsewhere,[18,25] and will be the subject of future work.

This study represents a framework considering the critical components and challenges for implementation of prospective genomics for infection control, and may be tailored to local needs and conditions, encompassing local MDROs of interest, different hospital and network structures, laboratory set-up, and screening practices. Further implementation studies should focus on reducing TATs for more real-time results, streamlined data integration and reporting systems, as well as furthering the education of hospital infection control, infectious diseases and laboratory teams to improve patient safety, and care through genomics-informed infection control in the near future.

Contributors

NLS, JCK, MLG & BPH conceived and designed the study. CG, RL, TS & CH designed and performed bioinformatic analyses and created figures. SB & MLS coordinated sequencing laboratory workflows. TK, MG, ML & HC arranged for isolates to be sent to the sequencing laboratory and provided microbiological data. RS, CM, MAS, LW & JCK coordinated collection of clinical and epidemiologic data. NLS, CG, JCK, RS, CM, TK, MAS, RL, MG, ML, LW, HC, MLG & BPH reviewed the data during the study, had input into the ongoing performance of the study and provided feedback on the manuscript. NLS and CLG wrote the manuscript; BPH, MLG and JCK edited the manuscript.

Declaration of Interests

The authors declare no conflicts of interest.

Acknowledgements

Financial support. This work was supported by the Melbourne Genomics Health Alliance (funded by the State Government of Victoria, Department of Health and Human Services, and the 10 member organizations); and individual grants from National Health and Medical Research Council (Australia) to NLS (GNT1093468), JCK (GNT1008549), BPH (GNT1105905), and MAS (GNT1116576).

Other members of the Controlling Superbugs Study Group include: Elizabeth Grabsch (Microbiology, Austin Health), Joanna Price and Carolyn Mason (Infection Control, Austin Health), Despina Kotsanas (Monash Infectious Diseases), Louise Wright (Infection Control, Monash Health), Dr Suraya Hanim Abdullah Hashim and Jennifer Mitchell (Infectious Diseases, Melbourne Health), Olivia Smibert and Victoria Madigan (Infectious Diseases, Peter MacCallum Cancer Centre) and Peter Pham (MDU Public Health Laboratory, University of Melbourne).

The authors would like to gratefully acknowledge the assistance of the following people: MDU PHL staff including Kerrie Stevens, Nicole Orlando, Samantha Tawil, Savitra Rambocus and Diane Daniel; John Greenough (infection control queries and patient movement data), Ben Rogers and Jean Lee (focus groups).

Data sharing

Raw sequence data for all included study samples (Supplementary Table S1), are available in NCBI Sequence Read Archive (SRA), BioProject PRJNA565795. All software and code used is publicly available using links provided in the Methods section.

References

1. World Health Organization. Antimicrobial resistance: global report on surveillance. Geneva, Switzerland: WHO Press, 2014.
2. Chia PY, Sengupta S, Kukreja A, S.L. Ponnampalavanar S, Ng OT, Marimuthu K. The role of hospital environment in transmissions of multidrug-resistant gram-negative organisms. *Antimicrob Resist Infect Control* 2020; **9**(1).
3. Kritsotakis E, Kontopidou F, Astrinaki E, Roumbelaki M, Ioannidou E, Gikas A. Prevalence, incidence burden, and clinical impact of healthcare-associated infections and antimicrobial resistance: a national prevalent cohort study in acute care hospitals in Greece. *Infect Drug Resist* 2017; **10**: 317-28.
4. Popovich KJ, Snitkin ES. Whole genome sequencing-implications for infection prevention and outbreak investigations. *Curr Infect Dis Rep* 2017; **19**(4): 15.
5. Parcell BJ, Gillespie SH, Pettigrew KA, Holden MTG. Clinical perspectives in integrating whole-genome sequencing into the investigation of healthcare and public health outbreaks – hype or help? *J Hosp Infect* 2021; **109**: 1-9.
6. Ruppe E, Olearo F, Pires D, et al. Clonal or not clonal? Investigating hospital outbreaks of KPC-producing *Klebsiella pneumoniae* with whole-genome sequencing. *Clin Microbiol Infect* 2017; **23**(7): 470-5.
7. Willems S, Kampmeier S, Bletz S, et al. Whole-genome sequencing elucidates epidemiology of nosocomial clusters of *Acinetobacter baumannii*. *J Clin Microbiol* 2016; **54**(9): 2391-4.
8. Tang P, Croxen MA, Hasan MR, Hsiao WW, Hoang LM. Infection control in the new age of genomic epidemiology. *Am J Infect Control* 2017; **45**(2): 170-9.
9. Conlan S, Thomas PJ, Deming C, et al. Single-molecule sequencing to track plasmid diversity of hospital-associated carbepenemase-producing *Enterobacteriaceae. Sci Transl Med* 2014; **6**: 254ra126.
10. 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**(2): 130-6.
11. De Man TJB, Yaffee AQ, Zhu W, et al. Multispecies outbreak of Verona integron-encoded metallo-ß-lactamase producing multidrug-resistant bacteria driven by a promiscuous incompatibility group A/C2 plasmid. *Clin Infect Dis* 2021; **72**(3): 414-20.
12. Peacock SJ, Parkhill J, Brown NM. Changing the paradigm for hospital outbreak detection by leading with genomic surveillance of nosocomial pathogens. *Microbiology* 2018; **164**(10): 1213-9.

13. Reuter S, Ellington MJ, Cartwright EJ, et al. Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology. *JAMA Intern Med* 2013; **173**(15): 1397-404.
14. Barrasa-Villar JI, Aibar-Remón C, Prieto-Andrés P, Mareca-Doñate R, Moliner-Lahoz J. Impact on morbidity, mortality, and length of stay of hospital-acquired infections by resistant microorganisms. *Clin Infect Dis* 2017; **65**(4): 644-52.
15. Mollers M, Lutgens SP, Schoffelen AF, Schneeberger PM, Suijkerbuijk AWM. Cost of nosocomial outbreak caused by NDM-1–containing *Klebsiella pneumoniae* in the Netherlands, October 2015–January 2016. *Emerg Infect Dis* 2017; **23**(9): 1574-6.
16. Otter JA, Burgess P, Davies F, et al. Counting the cost of an outbreak of carbapenemase-producing Enterobacteriaceae: an economic evaluation from a hospital perspective. *Clin Microbiol Infect* 2017; **23**(3): 188-96.
17. Balloux F, Bronstad Brynildsrud O, van Dorp L, et al. From theory to practice: translating whole-genome sequencing (WGS) into the clinic. *Trends Microbiol* 2018; **26**(12): 1035-48.
18. Gordon LG, Elliott TM, Forde B, et al. Budget impact analysis of routinely using whole-genomic sequencing of six multidrug-resistant bacterial pathogens in Queensland, Australia. *BMJ Open* 2021; **11**(2): e041968.
19. Sherry NL, Lee RS, Gorrie CL, et al. Pilot study of a combined genomic and epidemiologic surveillance program for hospital-acquired multidrug-resistant pathogens across multiple hospital networks in Australia. *Infect Control Hosp Epidemiol* 2020; **42**(5): 573-81.
20. World Health Organization. Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics. Geneva: WHO; 2017.
21. Sherry NL, Lane CR, Kwong JC, et al. Genomics for molecular epidemiology and detecting transmission of carbapenemase-producing Enterobacterales in Victoria, Australia, 2012 to 2016. J Clin Microbiol 2019; **57**(9).
22. Department of Health and Human Services Victoria. Victorian guideline on carbapenemase-producing Enterobacteriaceae for health services (version 2.1). 2018. https://www2.health.vic.gov.au/public-health/infectious-diseases/infection-control-guidelines/carbapenemase-producing-enterobacteriaceae-management/cpe-for-health-services.
23. Voor In 't Holt AF, Wattel AA, Boers SA, et al. Detection of healthcare-related extended-spectrum beta-lactamase-producing *Escherichia coli* transmission events using combined genetic and phenotypic epidemiology. *PLoS One* 2016; **11**(7): e0160156.
24. State Government of Victoria. Public Health and Wellbeing Regulations 2019, Statutory Rule no. 139/2019 v004. 2019.

25. Mellmann A, Bletz S, Boking T, et al. Real-time genome sequencing of resistant bacteria provides precision infection control in an institutional setting. *J Clin Microbiol* 2016; **54**(12): 2874-81.
26. Clermont O, Christenson JK, Denamur E, Gordon DM. The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ Microbiol Rep* 2013; **5**(1): 58-65.
27. Lane CR, Sherry NL, Porter AF, et al. Genomics-informed responses in the elimination of COVID-19 in Victoria, Australia: an observational, genomic epidemiological study. *Lancet Pub Health* 2021; doi:10.1016/s2468-2667(21)00133-x.
28. Mintzer V, Moran-Gilad J, Simon-Tuval T. Operational models and criteria for incorporating microbial whole genome sequencing in hospital microbiology - A systematic literature review. *Clin Microbiol Infect* 2019; **25**(9): 1086-95.
29. Gorrie CL, Da Silva AG, Ingle DJ, et al. Key parameters for genomics-based real-time detection and tracking of multidrug-resistant bacteria: a systematic analysis. *Lancet Microbe* 2021: doi:10.1016/S2666-5247(21)00149-X.
30. Pak TR, Kasarskis A. How next-generation sequencing and multiscale data analysis will transform infectious disease management. *Clin Infect Dis* 2015; **61**(11): 1695-702.
31. Armstrong GL, Maccannell DR, Taylor J, et al. Pathogen genomics in public health. *New Engl J Med* 2019; **381**(26): 2569-80.
32. Ward DV, Hoss AG, Kolde R, et al. Integration of genomic and clinical data augments surveillance of healthcare-acquired infections. *Infect Control Hosp Epidemiol* 2019; **40**(6): 649-55.
33. Crisan A, McKee G, Munzner T, Gardy JL. Evidence-based design and evaluation of a whole genome sequencing clinical report for the reference microbiology laboratory. *PeerJ* 2018; **6**: e4218.
34. Gouliouris T, Coll F, Ludden C, et al. Quantifying acquisition and transmission of *Enterococcus faecium* using genomic surveillance. *Nat Microbiol* 2021; **6**(1): 103-11.
35. Neumann B, Bender JK, Maier BF, et al. Comprehensive integrated NGS-based surveillance and contact-network modeling unravels transmission dynamics of vancomycin-resistant enterococci in a high-risk population within a tertiary care hospital. *PLoS One* 2020; **15**(6): e0235160.
36. Pinholt M, Larner-Svensson H, Littauer P, et al. Multiple hospital outbreaks of *vanA Enterococcus faecium* in Denmark, 2012-13, investigated by WGS, MLST and PFGE. *J Antimicrob Chemother* 2015; **70**(9): 2474-82.
37. Schurch AC, Arredondo-Alonso S, Willems RJL, Goering RV. Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene-based approaches. *Clin Microbiol Infect* 2018; **24**(4): 350-4.

38. Attwood LO, Sherry NL, Gorrie C, et al. Active surveillance to predict superbug infections (conference presentation). Australian Society for Infectious Diseases (ASID) conference. Melbourne, Victoria; 2020.

Table 1. Characteristics and infection control practices of hospital sites included in study

Hospital network	Hospital code	Hospital description	No. of inpatient beds[a]	High-risk wards	MDRO screening practices during study period and changes during study	Management of patients colonised or infection with MDRO
A	**A1**	Tertiary referral center, including ICU, solid-organ and bone marrow transplant	560	ICU Hematology/BMT and Oncology Renal Transplant Liver Transplant (Spinal ward and respiratory ward added)	*vanA* VRE and MRGN screening: ICU, on admission and twice weekly; haematology/oncology, renal and liver transplant wards screened on admission and weekly Additional MRSA screening in ICU (on admission and twice weekly) Biannual point-prevalence survey for *vanA* VRE and MRGN MRSA screening before critical surgeries (prosthetic joint, spinal and cardiac) *Change during study:* Added spinal ward and respiratory ward (ventilator support service) to high-risk wards for regular MDRO screening (October 2018)	MRSA: patients only isolated in certain circumstances, e.g. highly exudative wound; single room, disposable apron, gloves *vanA* VRE: single room with own bathroom, full gown, gloves; sometimes cohorted[b] with other *vanA* VRE-colonised patients ESBL-Ec and ESBL-Kp (multidrug-resistant)[c]: single room, own bathroom, disposable apron
	A2	Subacute hospital, aged care and rehabilitation services	150	None	Biannual point-prevalence survey for *vanA* VRE and MRGN	
	A3	Subacute hospital, rehabilitation services	60	None	Biannual point-prevalence survey for *vanA* VRE and MRGN	
B	**B1**	Tertiary referral center, including ICU and solid-organ transplant and specialist pediatric hospital (including neonatal ICU)	640	ICU Renal Transplant	ICU and renal ward screened for *vanA* VRE and carbapenem-resistant Gram negatives (CRGN) on admission and weekly MRSA screening before cardiac surgery	MRSA: single room, gloves and short-sleeved gown *vanA* VRE: single room, own bathroom, occasional cohorting[b] with other *vanA* VRE-colonised

					Change during study: Stopped routine screening of renal ward for VRE (June 2018). Network-wide changes in cleaning practices to from microfibre/steam cleaning to bleach cleaning (September 2018)	patients; gloves and short-sleeved gown. Additional measures if diarrhoea (full gown) ESBL-Ec: no specific management measures ESBL-Kp: single room, own bathroom, gloves and short-sleeve gown
	B2	Tertiary referral center, including ICU, trauma and some aged care & rehabilitation services	573	ICU	ICU patients screened for *vanA* VRE and carbapenem-resistant Gram negatives (CRGN) on admission and weekly *Changes during study:* Cleaning protocol changes as above	
C	**C1**	Tertiary referral center, including ICU, solid-organ and bone marrow transplant	571	ICU Haematology/BMT	ICU and haematology ward screened on admission and weekly for *vanA* VRE and MRGN	MRSA: single room only in certain situations, e.g. highly exudative wound *vanA* VRE: single room, own bathroom, gloves and full gown ESBL-Ec (multidrug-resistant)[c] and ESBL-Kp: single room, own bathroom, gloves and full gown
	C2	Subacute hospital, aged care and rehabilitation services	150	None	None	
D	**D1**	Specialized cancer care center. Located adjacent to Hospital C1 (ICU patients cared for at C1 before transfer back to hospital D1)	96	Hematology	Haematology ward patients screened on admission and weekly for *vanA* VRE and MRGN	MRSA: single room when possible; if open wounds, gown and gloves; if respiratory, mask and gloves *vanA* VRE: single room, own bathroom, gloves and full gown ESBL-Ec and ESBL-Kp: single room, own bathroom, no extra PPE but additional cleaning if incontinent or discharging wounds

ICU, intensive care unit; MRGN, multi-resistant Gram negatives (includes ESBL and carbapenem-resistant phenotypes); BMT, bone marrow transplant (allogeneic); PPE, personal protective equipment. [a]Inpatient beds, excludes day cases, hospital-in-the-home and mental health; [b]Cohorting, practice of patients colonised with the same MDRO sharing a room; [c]Multidrug-resistant ESBL, ESBL isolates with additional resistance aminoglycosides, fluoroquinolones and trimethoprim-sulfamethoxazole (at least two classes).

Table 2. Considerations for implementation of genomics-informed infection control workflow

Component of genomics workflow	Key considerations	Implementation strategy selected for this study
Target population	Consider targeting high-risk populations (e.g. ICU, haematology/oncology, transplant) vs acute inpatients vs all inpatients (acute and subacute)	Included all inpatients due to frequent movement between wards, and unknown burden of disease and transmission in subacute care
Target MDRO/s	Consider AMR spectrum/high-risk organisms; MDROs not covered by existing surveillance programs; organisms where healthcare acquisition more likely than community acquisition; MDROs of high prevalence in local population	Selected *van A* VRE (high prevalence, recently emerged), ESBL-Ec and ESBL-Kp (unclear healthcare contribution) and MRSA (changing epidemiology, unclear healthcare contribution) Note CPE already covered by existing statewide surveillance system; CRPa and CRAb low prevalence
MDRO screening strategy	Consider using existing MDRO screening strategies vs harmonised screening approaches. Could target high-risk populations or all inpatients (e.g. ward-based or hospital-wide point-prevalence surveys)	Elected for pragmatic approach, with each hospital network continuing existing screening strategies, noting differences between networks in analysis
Sample types	All samples vs clinical samples only (latter would underestimate transmission)	All samples included to maximise likelihood of identifying transmission
Clinical data collection	Consider data required to identify likely transmission in hospital, including admissions at other sites; consider ability to accurately identify admissions at different networks by centralised electronic systems	Patient movement data (admissions, wards and beds) critical to identify likely transmissions in conjunction with genomic data. Minimal clinical data collected in implementation phase, as not required to infer transmission.
Sequencing and analysis strategy	(i) Sequencing location: centralised (sequencing laboratory) vs decentralised (hospital laboratories) (ii) Bioinformatics analysis location: centralised vs decentralised, based on resources and expertise	(i) Sequencing in central laboratory, as sequencing not established at study sites (ii) Centralised analysis at sequencing laboratory as not established at study sites

	(iii) Analysis of combined genomic and epidemiologic data, depending on IPC resources and training, and established systems for electronic data analysis	(iii) Genomic and epidemiologic data integrated centrally by study team, presented to IPC teams for further interpretation and action
Bioinformatics analysis	Consider methods for transmission analysis. Examples: core genome SNP analysis, within species or within ST vs cgMLST approach, SNP thresholds, masking recombination Other considerations: frequency of analysis, inclusion of all data vs rolling window (e.g. last 6 or 12 months)	Elected for core genome SNP analysis within ST based on existing accredited bioinformatics pipelines; recombination masked except for MRSA; SNP thresholds selected based on existing data at the time. All samples included in each analysis (approximately monthly), noting computational expense and personnel requirements
Communication of results	Consider audience (IPC/ID/microbiology teams) and understanding of genomics Written reports (frequency) vs discussions in person vs interactive data portal with effective data visualisation	Chose to engage with IPC/ID/microbiology teams using in-person presentations as part of initial implementation, and development of written report format via focus group

MDRO, multidrug-resistant organism; VRE, vancomycin-resistant *Enterococcus*; ESBL-Ec, extended-spectrum beta-lactamase-phenotype *Escherichia coli*; ESBL-Kp, extended-spectrum beta-lactamase-phenotype *Klebsiella pneumoniae*; MRSA, methicillin-resistance *Staphylococcus aureus*; ST, sequence type; cgMLST, core genome multi-locus sequence type; SNP, single nucleotide polymorphisms; IPC, infection prevention and control; ID, infectious diseases.

Table 3. Patient and isolate characteristics

Total no. of isolates included	2275		
Total no. of patients included	1970		
Age (median, IQR)	65 years (47-77yrs)		
Male sex (no., %)	1063 (56·8%)		
Sample collected for clinical purposes (suspected infection)	1428 (62·8%)		
Sample collected for MDRO screening purposes	847 (37·2%)		
Time from admission to isolate collection (median, IQR)	3 days (1-12 days)		
Ward where sample was collected	**All samples (% total samples)**	**Clinical samples (% per ward)**	**Screening samples (% per ward)**
Intensive care unit	373 (16·4%)	117 (31·4%)	256 (68·6%)
Other high-risk ward [a]	434 (19·1%)	121 (27·9%)	313 (72·1%)
General acute ward [b]	1090 (47·9%)	865 (79·4%)	225 (20·6%)
Emergency department	209 (9·2%)	204 (97·6%)	5 (2·4%)
Subacute care ward [c]	169 (7·4%)	121 (71·6%)	48 (28·4%)

MDRO, multidrug-resistant organism; IQR, interquartile range. [a] High-risk wards, includes hematology, oncology, renal ward (including renal transplant), and liver transplant wards.

[b] General acute wards, include all wards not designated as ICU, high-risk or subacute.

[c] Subacute care, includes aged care, rehabilitation, palliative care and spinal wards.

Table 4. Impacts of genomics program for hospital MDRO surveillance

Unexpected MDRO transmission identified, resulting in specific ward interventions	• *vanA* VREfm cluster identified in haematology ward • ESBL-Kp 'super-spreader' patient identified in one ward[a] • Intensive ward cleaning instituted ('super-clean')[b] • Review of patient placement/movements in affected wards • Increased surveillance in affected wards[c]
Outbreak investigation support	• Supported patient-to-patient transmission and implication of shared patient equipment in outbreaks[d] • Identified transmission where epidemiological links were complex • Able to show that new cases in other hospital wards were not linked to an outbreak, demonstrating that the outbreak management had been successful and allowing infection control interventions to be stepped down
Comparison of MDRO transmission rates between networks, resulting in network-wide changes	• Identified higher rate of *vanA* VRE transmission in one network than others, resulting in network-wide change in cleaning practices[e]
Increased awareness of MDRO transmission patterns	• Increased awareness of MRSA and ESBL-Ec transmissions in hospital (previously thought to all be community-acquired) • Confirmed that MDRO identified through screening predicted the MDRO that subsequently caused clinical infection. Useful for advising empiric therapy e.g. in neutropaenic sepsis[38]
Public health impacts	• Large burden of hospital-acquired *vanA* VREfm demonstrated, contributing to new legislative requirement[24] for all *vanA* VREfm cases to be notified to the state health department • Low burden of carbapenem-resistant *Pseudomonas aeruginosa* and *Acinetobacter baumannii* established as baseline before inclusion in state notifiable diseases list

[a] 'Super-spreader' indicating source patient with unusually high number of transmissions to other patients

[b] Ward 'super-cleans' included thorough terminal/discharge-type cleaning for all areas of the ward (including communal areas, corridors, offices, and storerooms) using a bleach-based detergent where possible.

[c] Increased surveillance included additional environmental screening, additional patient screening (added to roster of high-risk wards with weekly patient screening for VREfm and MRGN, and additional focused hand hygiene surveillance

[d] Shared patient equipment includes any equipment that is not restricted to a single patient, e.g. blood glucometers, blood pressure cuffs, weigh chairs

[e] Cleaning practices changed from steam wands and microfibre cloth cleaning to bleach cleaning

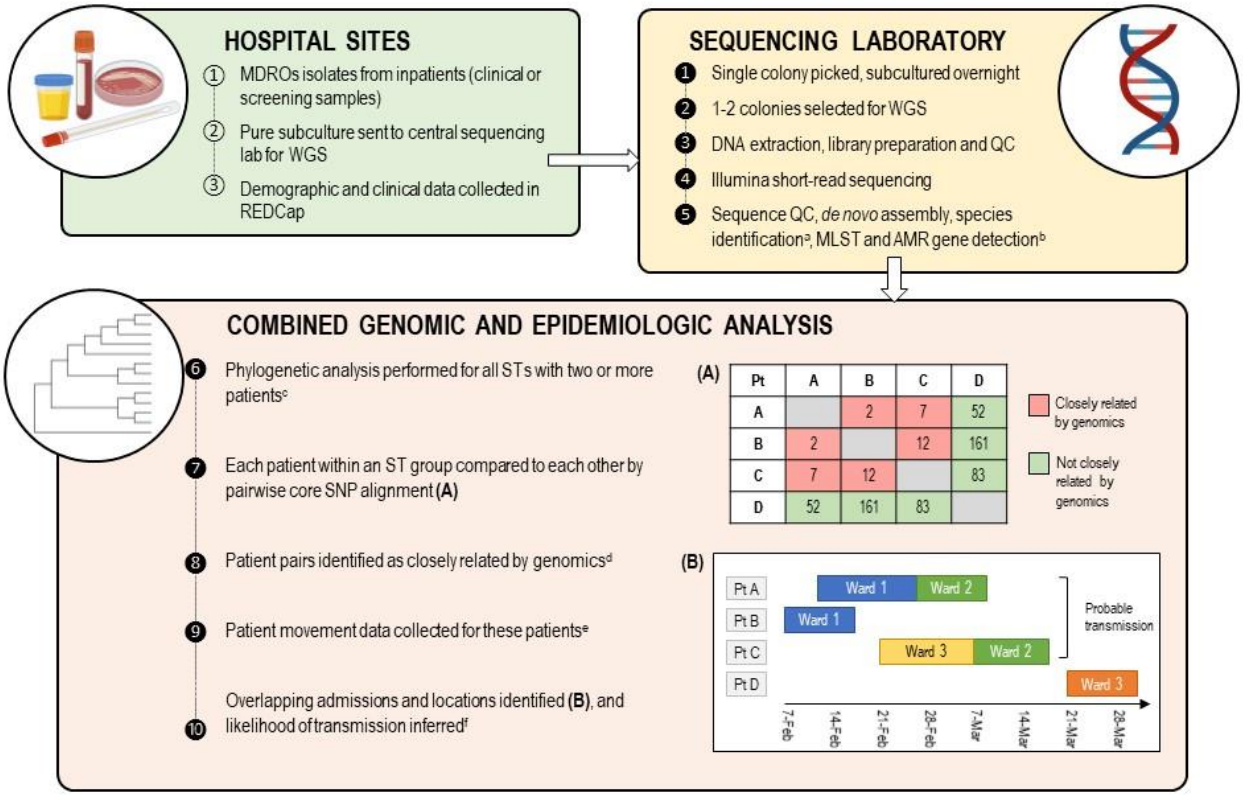
Table 5. Areas for further development for implementation of infection control genomics workflow

Issue and domain	Description	Potential solutions
Delays in time to results		
a. Hospital laboratory	Delayed time to send isolates to sequencing laboratory • Inadequate staffing • Batch processing of isolates (weekly or longer intervals) • Major disruptions from introduction of new LIS at two laboratories (inability to perform data extractions for study, affecting three sites)	Dedicated staff (supervising microbiologist and technician) More frequent processing of isolates for transit to sequencing lab Ensure robust data extraction capabilities available when implementing new LIS
b. Sequencing laboratory	Minor delays, due to high volume of samples and lower relative priority	Increase in sequencing capacity, redundancy to account for surges in sample volume
c. Bioinformatic analysis	Delays in analysis due to inadequate staffing Increasing complexity of analysis and computational requirements towards end of study due to large isolate numbers (all study samples of same MLST included in each transmission analysis)	Increase dedicated computing resources for complex bioinformatic analyses Work towards partial automation of transmission analysis Develop alternative bioinformatics methods to deal with genomics 'big data' Explore effect of rolling windows (e.g. only including last 6 or 12 months of isolates) on detecting transmission
d. Reporting process	In-person presentations to each network used in this study (to engage and educate staff at study sites) did not allow for results to be conveyed in a timely manner	Issue individual reports for each patient isolate, including MLST result (can be used by Infection Control teams to rapidly exclude or include patients in potential transmissions) New report designed for infection control genomics results (aggregated hospital or network-level data) from focus groups Develop innovative approaches for reporting aggregated results on a hospital/network level e.g. interactive dashboards with advanced data visualisation capabilities
Impaired access to patient epidemiologic data	Each hospital network using different patient information management systems; no readily available way to extract patient data; data had to be collected manually	Develop informatics systems to reliably extract patient movement data, and integrate with results of transmission analyses to inform Infection Control interventions

(including admission, ward and bed data)	(transcription errors, variable formatting requiring manual correction)	
Inability to track patient movements between hospital networks	No readily-available centralised system available for tracking admissions of patients across different hospital networks to identify inter-network MDRO transmission	Engage with health department to advocate for development of centralised admissions database, available in real-time
Perceived cost issues	Sequencing costs perceived to be prohibitive for widespread application currently	Optimise utility of prospective genomics by targeting to local MDROs or locations of concern; demonstrate cost-effectiveness of prospective genomics

LIS, laboratory information system; MLST, multi-locus sequence type (a typing method to subset isolates of the same species into smaller, closely-related groups for targeted transmission analysis); MDRO, multidrug-resistant organism

Figure 1. Infection control genomics workflow implemented in this study



[a] Species identification by *k*-mer identification (*kraken*); *Klebsiella* further subspeciated using *kleborate* tool.

[b] Sequences analysed for presence of complete *vanA* operon, ESBL/AmpC genes and *mecA/mecC*; if these were absent, isolates underwent further phenotypic testing to ensure they met inclusion criteria (phenotypic antimicrobial resistance).

[c] Aligning isolates to reference genome of same ST, or *de novo* assembly of earliest isolate if this was not available; ST 131 *E. coli* analysed in two subclades due to large number of isolates. Recombinant sites were masked with *gubbins* for species other than *S. aureus.*

[d] Each isolate sequence within an ST was compared to all others, and closely-related isolates were determined by core SNP differences; ≤15 SNPs for *Staphylococcus aureus,* ≤25 SNPs for other species

[e] Data collected for 12 months prior to first study sample until the end of study

[f] Likelihood of transmission inferred from combined genomic and epidemiologic data, categorised as 'Probable' (same ward at same time), 'Possible' (same ward at different time but within 60 days, or same hospital at the same time), 'Unlikely' (neither of the above), or 'Above screening threshold' (not closely related by genomics).

(A) Example of pairwise SNP distance matrix for sequences from four patient isolates; patient pairs with pairwise SNP distances at or below the screening threshold (15 SNPs for MRSA, 25 SNPs for other MDROs) are classified as closely related by genomics (red), and pairs above these thresholds are designated as not closely related by genomics. **(B)** Example of Gantt chart demonstrating bed movements (ward admissions) of same four patients from (A) over time; patients A, B and C had closely related MDROs by genomics, and had overlapping admissions (same ward at the same time), therefore constituting probable MDRO transmission. Patient D's sequence was not related to the other patients' sequences by genomics (above screening threshold), and not considered to be involved in MDRO transmission to or from these other patients, despite having an MDRO of the same sequence type (ST). Abbreviations: MDRO, multidrug-resistant organism; WGS, whole genome sequencing; QC, quality control; MLST, multi-locus sequence typing; AMR, antimicrobial resistance; ST, sequence type; SNP, single-nucleotide polymorphism (single base difference between two or more isolates); Pt, patient.

Figure 2. MDRO isolates by network, species, and reason for sample collection. Bars are shaded by reason for sample collection with lightest shades representing clinical samples, middle shades representing screening samples, and darkest shades (only for Network A) representing screening samples performed as part of network-wide biannual point-prevalence surveys (PPS)(note, MRSA not included in PPS).

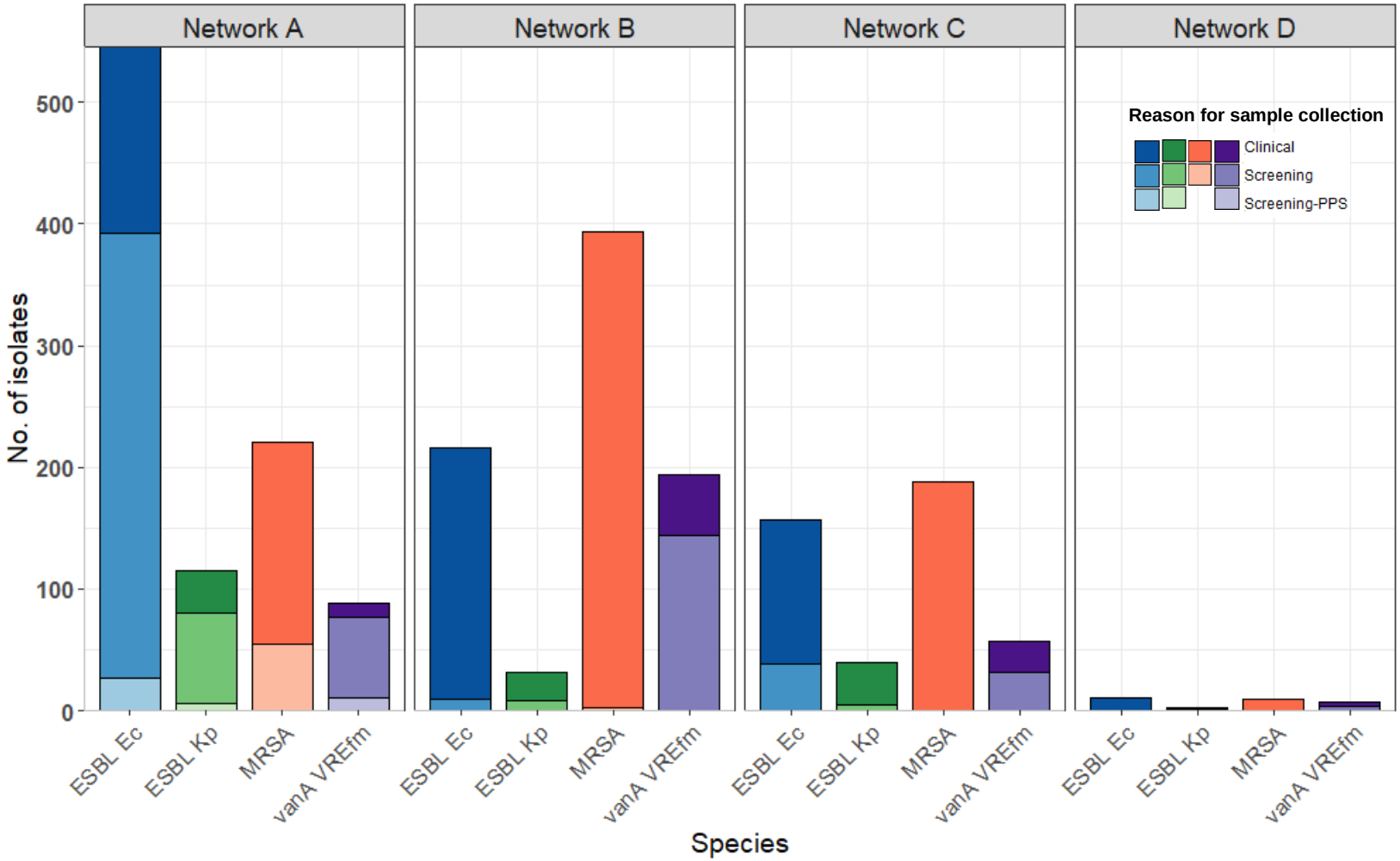
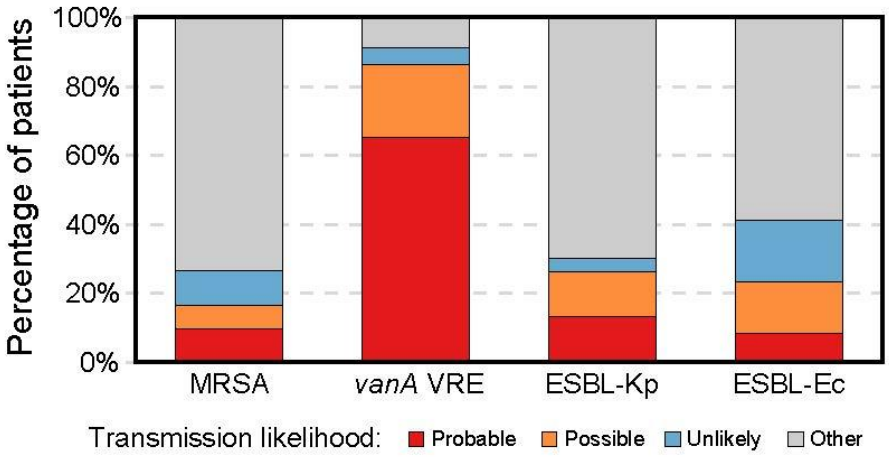


Figure 3. MDRO transmission as assessed by combined genomics and epidemiologic assessment

(A) Likelihood of transmission using combined genomic and epidemiologic data, by species. 'Other' (grey) includes patients with samples with no genomic links to other study samples (singleton STs, or above SNP screening thresholds). (**B)** Putative transmission rates per 100,000 occupied bed days (OBDs) by species, coloured by hospital network. Horizontal dotted line represents the mean transmission rate for each species across all hospital networks.



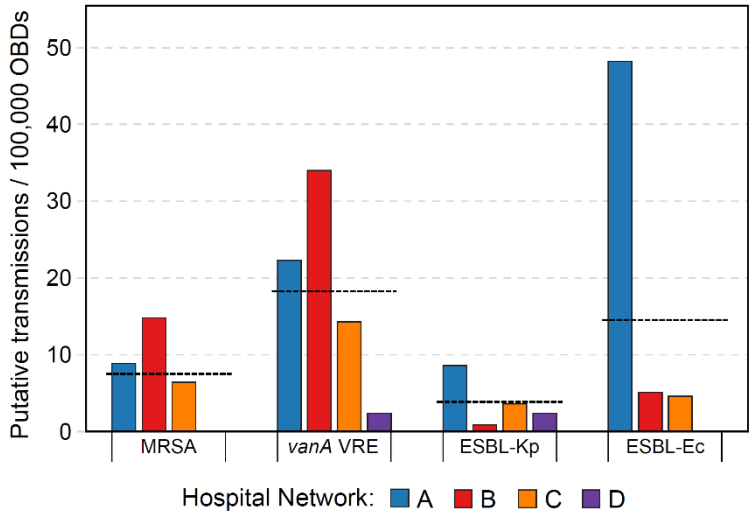
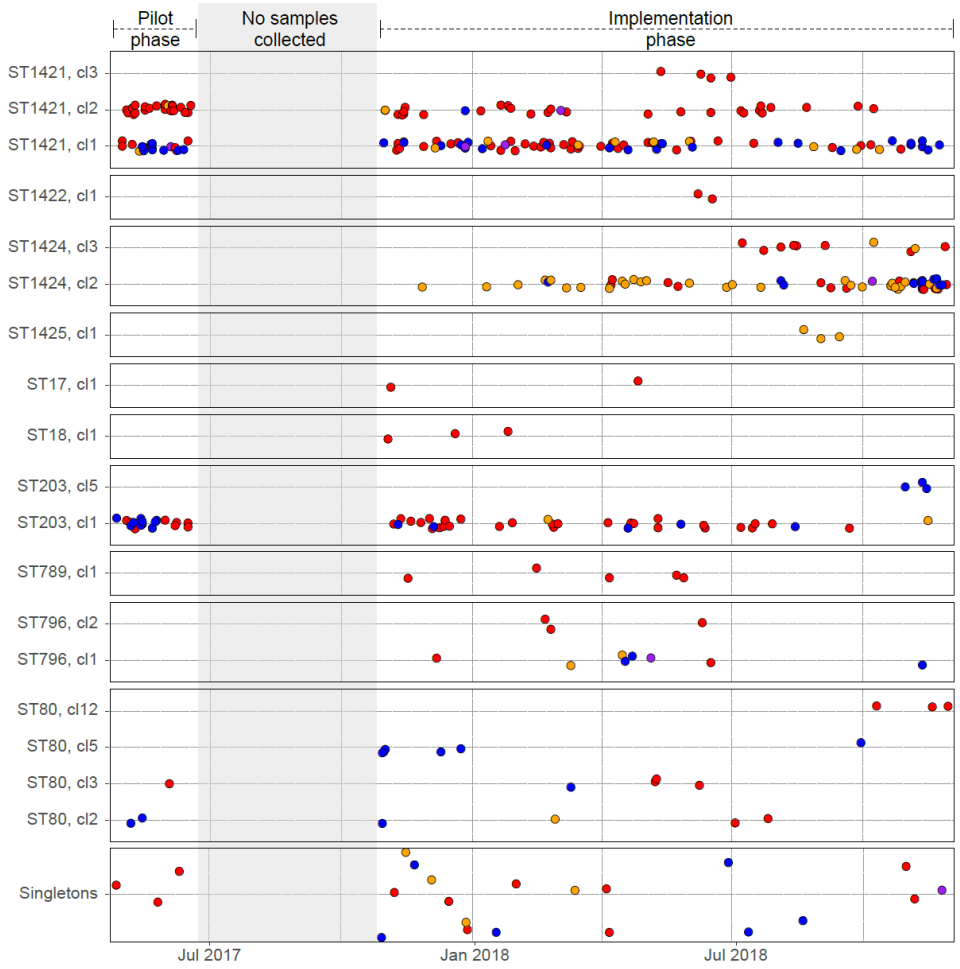
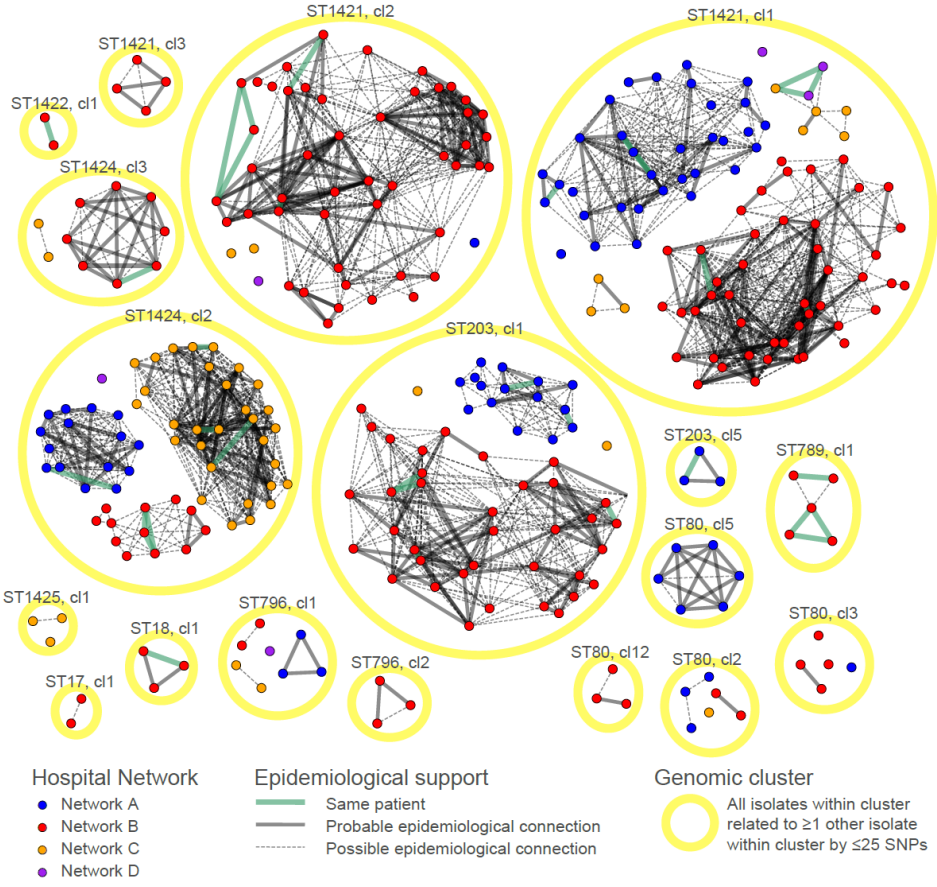


Figure 4. *vanA* VREfm cluster timeline and genomic transmission networks

A.

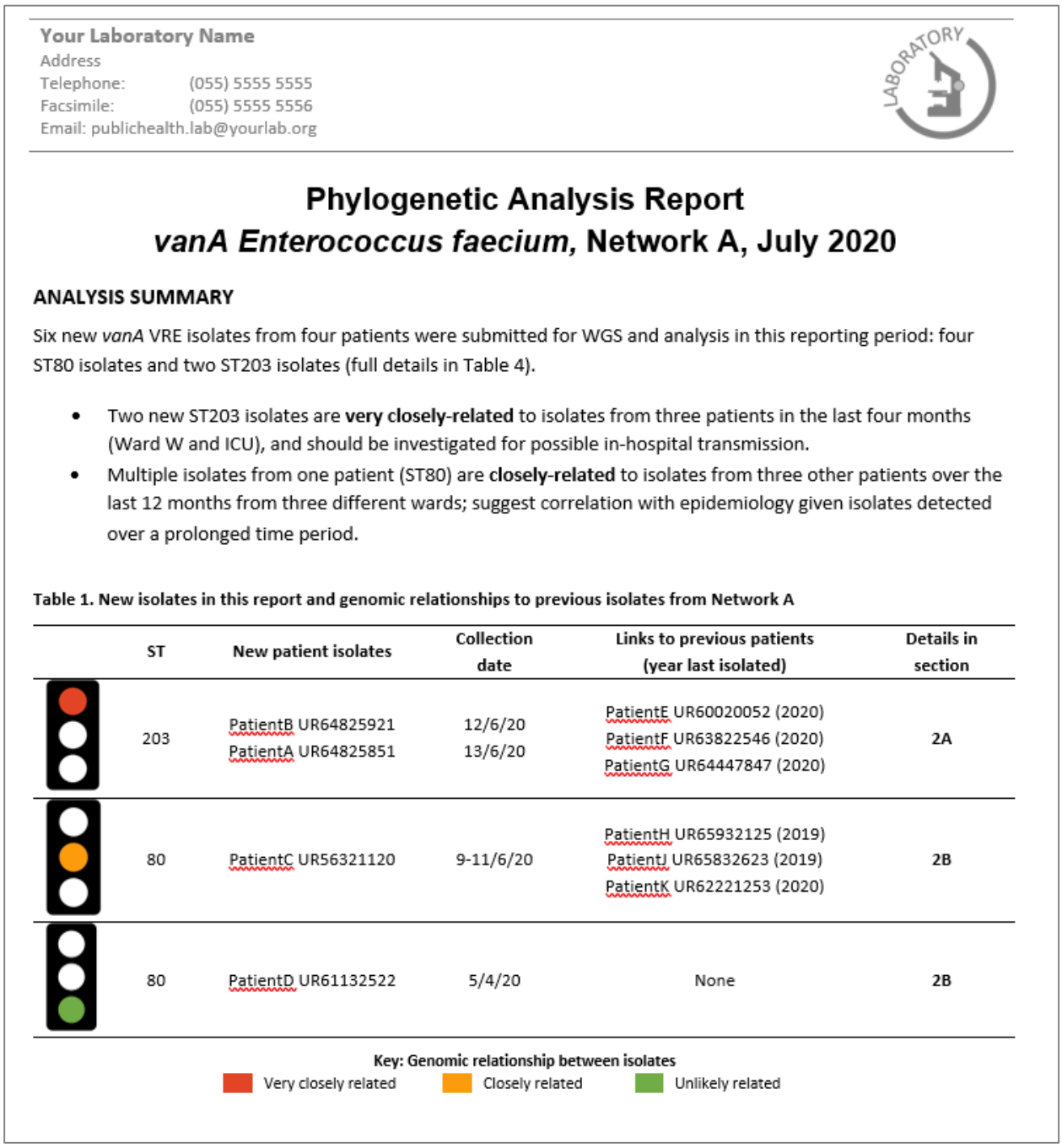


B.



A. Timeline of *vanA* VREfm cases over the study periods, separated by ST and sub-cluster (cl) along the Y axis, and coloured by hospital network. Timeline is separated into pilot phase (left) and implementation phase (right); shaded timeframe shows the time between phases without sampling. Each point represents one case; only the first sample for each case in the study is included. Points have been separated vertically to allow for visualisation of closely-spaced cases. Sample sequences that did not cluster with any other study cases are shown in the 'Singletons' panel.
B. Network diagram of genomic and epidemiologic links between *vanA* VREfm cases. Each point represents one case; only the first sample for each case in the study is included, and is coloured by hospital network. The yellow circles outline genomic clusters, and lines between points represent epidemiologic links between cases (thick line, probable transmission; dotted line, possible transmission; green line, same patient. Absence of line indicates unlikely transmission by epidemiology, i.e. no known overlap between patients in space or time).
Abbreviations: VREfm, vancomycin-resistant *Enterococcus* faecium; ST, sequence type.

Figure 5. Suggested format for reporting prospective genomics results to an infection control team



Your Laboratory Name
Address
Telephone: (055) 5555 5555
Facsimile: (055) 5555 5556
Email: publichealth.lab@yourlab.org

Phylogenetic Analysis Report
vanA Enterococcus faecium, Network A, July 2020

ANALYSIS SUMMARY

Six new *vanA* VRE isolates from four patients were submitted for WGS and analysis in this reporting period: four ST80 isolates and two ST203 isolates (full details in Table 4).

- Two new ST203 isolates are **very closely-related** to isolates from three patients in the last four months (Ward W and ICU), and should be investigated for possible in-hospital transmission.
- Multiple isolates from one patient (ST80) are **closely-related** to isolates from three other patients over the last 12 months from three different wards; suggest correlation with epidemiology given isolates detected over a prolonged time period.

Table 1. New isolates in this report and genomic relationships to previous isolates from Network A

	ST	New patient isolates	Collection date	Links to previous patients (year last isolated)	Details in section
	203	PatientB UR64825921 PatientA UR64825851	12/6/20 13/6/20	PatientE UR60020052 (2020) PatientF UR63822546 (2020) PatientG UR64447847 (2020)	2A
	80	PatientC UR56321120	9-11/6/20	PatientH UR65932125 (2019) PatientJ UR65832623 (2019) PatientK UR62221253 (2020)	2B
	80	PatientD UR61132522	5/4/20	None	2B

Key: Genomic relationship between isolates
Very closely related Closely related Unlikely related

Excerpt of a suggested report format for intermittently reporting results from prospective genomic surveillance to an infection control team. This report was designed after feedback from focus groups of infection control, infectious diseases and microbiology staff, including some participants with experience in genomics, and some without. All data are fictitious. Full report included in Supplementary Data (Figure S6).

Title: Multi-site implementation of whole genome sequencing for hospital infection control: A prospective genomic epidemiological analysis

Brief title: Prospective genomics for hospital infection control

Authors: Norelle L Sherry MBBS[1,2,3], Claire L Gorrie PhD[1,3], Jason C Kwong PhD[2,3,4], Charlie Higgs BSc[3], Rhonda L Stuart PhD[5,6,7], Caroline Marshall PhD[8,9], Susan A Ballard PhD[1], Michelle Sait PhD[1], Tony M Korman MBBS[5,10], Monica A Slavin MD[11,12], Robyn S Lee PhD[3†], Maryza Graham MBBS[5,10], Marcel Leroi MBBS[2,13], Leon J Worth PhD[11.12], Hiu Tat Chan MBBS[14‡], Torsten Seemann PhD[1], M Lindsay Grayson PhD[2,4,13]* & Benjamin P Howden PhD[1,2,3]* on behalf of the Controlling Superbugs Study Group

*authors contributed equally

[†] Current affiliation: Epidemiology Division, Dalla Lana School of Public Health, University of Toronto, Toronto, Canada

[‡] Current affiliation: Department of Molecular Microbiology, VCS Pathology, Australian Centre for Prevention of Cervical Cancer, Carlton, Victoria, Australia

[1] Microbiological Diagnostic Unit (MDU) Public Health Laboratory, Department of Microbiology & Immunology at the Peter Doherty Institute for Infection & Immunity, University of Melbourne, Melbourne, Victoria, Australia

[2]Department of Infectious Diseases, Austin Health, Heidelberg, Victoria, Australia

[3]Department of Microbiology & Immunology at the Peter Doherty Institute for Infection & Immunity, University of Melbourne, Melbourne, Victoria, Australia

[4]Department of Medicine, Austin Health, University of Melbourne, Heidelberg, Victoria, Australia

[5]Monash Infectious Diseases, Monash Health, Clayton, Victoria, Australia

[6]Monash University, Clayton, Victoria, Australia

[7]South East Public Health Unit, Monash Health, Clayton, Victoria, Australia

[8]Infection Prevention & Surveillance, Victorian Infectious Diseases Service, Melbourne Health, Parkville, Victoria, Australia

[9]Department of Infectious Diseases, The University of Melbourne at the Peter Doherty Institute for Infection & Immunity, Melbourne, Victoria, Australia

[10]Department of Microbiology, Monash Health, Clayton, Victoria, Australia

[11]Department of Infectious Diseases, Peter MacCallum Cancer Centre, Parkville, Victoria, Australia

[112]National Centre for Infections in Cancer, Sir Peter MacCallum Department of Oncology, University of Melbourne, Melbourne, Victoria, Australia

[13]Department of Microbiology, Austin Health, University of Melbourne, Heidelberg, Victoria, Australia

[14]Department of Microbiology, Melbourne Health, Parkville, Victoria, Australia

Corresponding author: Prof. Benjamin Howden, MDU Public Health Laboratory, Department of Microbiology & Immunology at the Peter Doherty Institute for Infection and Immunity, University of Melbourne. Email: bhowden@unimelb.edu.au.

Manuscript word count: 4504 words (excluding abstract, references, figures and tables)

Abstract word count: 250 words

Abstract (250 words)

Background

Current microbiological methods lack the resolution to accurately identify multidrug-resistant organism (MDRO) transmission, however, whole genome sequencing can identify highly-related patient isolates providing opportunities for precision infection control interventions. We investigated the feasibility and potential impact of a prospective multi-centre genomics workflow for hospital infection control.

Methods

We conducted a prospective genomics implementation study across eight Australian hospitals over 15 months (2017-2018), collecting all clinical and screening isolates from inpatients with *vanA* VRE, MRSA, ESBL *Escherichia coli* (ESBL-Ec), or ESBL *Klebsiella pneumoniae* (ESBL-Kp). Genomic and epidemiologic data were integrated to assess MDRO transmission.

Findings

In total, 2275 isolates were included from 1970 patients, predominantly ESBL-Ec (40·8%) followed by MRSA (35·6%), *vanA* VRE (15·2%), and ESBL-Kp (8·3%).

Overall, hospital and genomic epidemiology showed 607 patients (30·8%) acquired their MDRO in hospital, including the majority of *vanA* VRE (266 patients, 86·4%), with lower proportions of ESBL-Ec (186 patients, 23·0%), ESBL-Kp (42 patients, 26·3%), and MRSA (113 patients, 16·3%). Complex patient movements meant the majority of MDRO transmissions would remain undetected without genomic data.

The genomics implementation had ~~significant~~ major impacts, identifying unexpected MDRO transmissions prompting new infection control interventions, and contributing to *vanA* VRE becoming a notifiable condition. We identified barriers to implementation and recommend strategies for mitigation.

Interpretation

Implementation of a multi-centre genomics-informed infection control workflow is feasible and identifies many unrecognised MDRO transmissions. This provides critical opportunities for interventions to improve patient safety in hospitals.

Funding

72 Melbourne Genomics Health Alliance (supported by State Government of Victoria, Australia), and
73 National Health and Medical Research Council (Australia).

Introduction

Antimicrobial resistance is a serious threat to human health, and disproportionately affects hospitalised patients, with many multidrug-resistant organisms (MDROs) acquired directly from other hospital patients or indirectly from the hospital environment.[1-3] Whilst these MDROs can usually be adequately detected by routine diagnostic microbiology methods, these methods are insufficient to identify transmission chains, as multiple strains of each MDRO may be circulating in the hospital and the community.[4,5]

Whole genome sequencing (WGS, or "genomics") of MDROs provides high-resolution typing data to identify transmission networks between patients and/or the environment.[6-8] Genomics has increasingly been used in the research setting to identify hospital MDRO outbreaks, most often in a retrospective manner after a possible outbreak has already been identified by epidemiologic surveillance.[9-11] However, the greatest potential for genomics for hospital infection control is to *prospectively* identify likely transmission of MDROs between a small number patients, allowing early infection control interventions before an outbreak has become established,[12,13] therefore limiting patient morbidity, mortality, unnecessary antibiotic use, and waste of hospital resources.[14-16] Although WGS costs have been steadily decreasing over the last decade, implementation should be directed to the most cost-effective targets with the greatest benefits to health and resources.[17,18]

Translation of genomics from the research setting to the real-world infection control setting requires consideration of the elements that are likely to be critical for successful implementation.[5,17] Factors to consider include an optimised genomic workflow model to achieve clinically meaningful turnaround times, which patient populations and MDROs are likely to yield the greatest benefit from genomics, and how to effectively communicate complex genomic analyses to infection control teams at the coalface, to enable implementation of infection control interventions for maximal benefit to the patients and healthcare system.[4,5]

Here we present a study implementing a prospective genomics workflow to detect transmission of MDROs across multiple hospital networks over 15 months, with all MDRO samples from hospital inpatients being sequenced at a central laboratory, and results returned to the hospital infection control team for detailed epidemiological analysis and further management. We aim to establish the feasibility of such a workflow and potential impacts, and identify key components and potential barriers to implementation.

Methods

Study scoping and design

As a scoping exercise to design a genomics workflow for infection control, we first listed the *components* to be included, identified the *key considerations* for each component, then identified a *strategy for implementation* for each component, taking into account existing local infrastructure, MDRO patterns and resources (staffing, consumables, and financial resources), before settling on a study design for local implementation to establish the feasibility and potential impacts of infection control genomics.

We subsequently performed a prospective "real-world" multi-centre genomics implementation study of MDRO transmission in eight hospitals from four hospital networks, including approximately 2800 acute and subacute inpatient beds, in Melbourne, Australia (population 4·9 million in 2018); hospital characteristics are described in Table 1. MDRO isolates from four defined species (see below) were included from patient samples (clinical or screening samples) collected routinely from hospital inpatients (>24h) at any time during their admission. Duplicate screening isolates were excluded, and duplicate clinical isolates were only included if collected more than 14 days after the previous sample (see Supplementary Data for details).

The study was conducted in two phases: a pilot phase (24th April – 18th June 2017), and an implementation phase (30th October 2017 – 30th November 2018), totalling 15 months. Results of transmission analyses were only reported to hospitals during the implementation phase. Pilot phase results have previously been reported,[19] but are included here as isolates from these patients may be implicated in future transmission events detected in the implementation phase. Outbreak investigation support (including sequencing of environmental and relevant historical isolates) was also offered to participating sites during the study.

MDRO definitions

Four principal MDROs were included in the pilot phase, chosen as AMR pathogens of global significance (WHO global priority pathogen list[20]) with known transmission in hospitals; the selected MDROs were *vanA* vancomycin-resistant *Enterococcus faecium* (*vanA* VREfm), methicillin-resistant *Staphylococcus aureus* (MRSA), and extended-spectrum beta-lactamase phenotype *Escherichia coli* and *Klebsiella pneumoniae* (ESBL-Ec and ESBL-Kp) (see Supplementary data for details). Carbapenem-resistant *Acinetobacter baumannii* complex (CRAb) and carbapenem-resistant *Pseudomonas aeruginosa* (CRPa) were also included in the pilot phase,[21] but excluded here due to low prevalence. Carbapenem-resistant Enterobacterales (CPE) were excluded, as these were already

under genomic surveillance as part of the state-wide CPE surveillance program.[22] As the pathogens included in the study were not notifiable by legislation, no specific public health interventions or reporting was in place that the time, and infection control units managed these independently.

Due to the large volume of ESBL-Ec included during the pilot phase, we modified inclusion criteria for the implementation phase to only include ESBL-Ec that were also resistant to fluoroquinolones, as this was more consistent with existing infection control practices at participating sites.

Infection control, clinical and epidemiologic data

Hospital infection control practices and MDROs screening protocols at each site were assessed regularly during the study, including any changes made in response to genomic data (Table 1 and Supplementary data). Basic demographic and isolate data were collected for all patients; epidemiologic data were collected for a subset of patients where genomics suggested potential MDRO transmission to assess the clinical likelihood of such cross-transmission having occurred. Transmission rates were estimated by calculating the number of putative transmissions per 100,000 occupied bed days (OBDs, inpatient beds occupied on daily bed count, excluding mental health).

Laboratory workflows

MDROs were isolated, worked up and reported by the hospital laboratories as per their usual protocols. For patients and isolates meeting inclusion criteria, a pure subculture was sent to the central laboratory for sequencing and isolate storage.

The sequencing laboratory is an ISO-accredited laboratory with high-throughput genomic sequencing capacity, located centrally between the participating hospitals. Once samples arrived, a single colony was subcultured onto horse blood agar, incubated overnight, then 1-2 colonies selected and placed into lysis buffer. DNA extraction, NexteraXT library preparation and quality control (QC) were performed as previous described,[19] and isolates sequenced on Illumina NextSeq (San Diego, CA, USA).

Bioinformatics and transmission analysis

An overview of the combined genomic and epidemiologic analysis workflow to identify MDRO transmission can be found in Figure 1, with details included in Supplementary Data. Briefly, all sequences and assemblies underwent routine quality control checks, and underwent *de novo* genome assembly, species identification, multi-locus sequence typing (MLST), and identification of antimicrobial resistance (AMR) genes.

Transmission analyses were performed within sequence type (ST) for all STs with two or more patients. Reads were aligned to a complete reference genome of the same ST, and recombinant sites were masked (except for *S. aureus*). Variant sites were detected from short reads for each ST, and pairwise single nucleotide polymorphism (SNP) differences were calculated for all isolate pairs.

To screen for potential MDRO transmission, we defined isolate pairs with ≤25 core SNP differences (*E. coli, K. pneumoniae,* and *E. faecium*) or ≤15 core SNP differences (MRSA) as 'genomically related'. Isolate pairs with different key AMR genes were excluded from further analysis (i.e. *mec* type for MRSA, and ESBL or AmpC gene (allele) for ESBL-Ec and ESBL-Kp; *vanA* alleles were excluded as this was a requirement for study inclusion). Patients with isolates that were 'genomically related' then had further epidemiologic data collected, detailing their admission history (dates, hospitals and wards) for 12 months before the earliest isolate until the end of the study period.

Epidemiologic data were then compared for patients above the screening thresholds ('genomically related') within ST (or ST131 *E. coli* subclades). Patient pairs were classified according to a modification of previously published definitions[23] as 'probable transmission' if on the same ward at the same time, 'possible transmission' if admitted to the same ward at a different time (within 60 days), or admitted to the same hospital at the same time; all other patients were classified as 'unlikely transmission'. Statistical analyses were performed using Stata v16.

Reporting to infection control teams and report design

During the implementation phase of the study, genomic results for each hospital network were compiled and reported in person to infection control teams, with epidemiologic data where available; written reports were then supplied after each presentation. As this was the first time that most infection control teams had been exposed to genomic data, the reporting presentations included an informal education component to familiarise teams with genomic analyses and interpretation of hospital's genomic results. The study team also sought feedback from infection control teams during and after each reporting session, and made themselves available for any questions from infection control teams between reports or presentations.

Using the feedback from the reporting sessions, two focus group sessions were held at the end of the study with infectious diseases physicians, infection control practitioners, and medical microbiologists to collect feedback to design a genomics report for infection control use (see Supplementary Methods for details). A final report was designed based on feedback from the focus group sessions.

Ethics approval

This study was approved by the Melbourne Health Human Research Ethics Committee (HREC) and endorsed by the corresponding HREC at each participating site.

Results

Scoping and designing a workflow for implementation of genomics-informed infection control

To design a genomics workflow for infection control, we considered the critical elements for implementation: target population, target MDROs, MDRO screening strategy, sample types, clinical data collection, sequencing and analysis strategy, bioinformatics analysis, and communication of results (Table 2). The final workflow, as implemented in this study, is shown in Figure 1.

Isolate and patient characteristics

Overall, 2641 bacterial isolates were submitted by hospital laboratories for WGS; 175 were excluded due to patients not meeting inclusion criteria, 116 isolates were excluded as duplicates, and a further 64 isolates did not meet microbiological inclusion criteria (see Supplementary Figure S1 for details).

Ultimately, 2275 isolates from 1970 patients were included for analysis, predominantly ESBL-Ec (929 isolates, 47·2%), followed by MRSA (811 isolates, 41·2%), *vanA* VRE (346 isolates, 17·6%), and ESBL-Kp (189 isolates, 9·4%). Most isolates were collected for clinical reasons (62 ·8% overall), although this varied according to species and screening programs at each hospital network (Figure 2). After screening swabs (37·1%), urine (20·4%), and swabs from non-sterile sites (20·4%) were the most common sample types, with only 6·9% of isolates being from blood cultures (Supplementary Figure S2). Median patient age was 65 (IQR 47-77), and 56·8% of patients were male. 297 patients (15 ·1%) had more than one isolate included (range 1-7 isolates)(Table 3). For patients with one screening and at least one clinical isolate with the same MDRO, median time between screening and first clinical isolate was 13 days (range 0-365 days, IQR 4-55 days)(Supplementary Table S2).

vanA *VREfm was transmitted more frequently in hospital than other MDROs in this study*

Initial genomic analysis identified that 844 (42·8%) of isolates were genomically-related to at least one other isolate from the study (i.e. below screening pairwise SNP threshold, excluding isolates from same patient, and excluding key AMR gene mismatch) (Supplementary Table S3 and Supplementary Figure S3).

Aligning the patient admission history of each pair of patients with genomically-related isolates (see workflow Figure 1), MDRO isolates from 354 (18·0%) patients were identified as probable

transmission in hospital, 253 patients (12·8%) as possible transmission in hospital, and 236 (12.0%) as unlikely acquired in hospital (Figure 3a, and pairwise SNP distributions shown Supplementary Figure S4). This varied significantly by species: 266 *vanA* VREfm patients (86·4%) were identified as probable or possible transmission in hospital, whereas lower proportions of patients with ESBL-Ec, ESBL-Kp, and MRSA were identified as probable or possible in-hospital transmission (186 ESBL-Ec patients (23·0%), 113 MRSA patients (16·3%), 42 ESBL-Kp patients (26·3%); χ^2=488 ·32 (VREfm vs other species), p<0 ·001). Based on combined genomic and epidemiologic analysis, 44·8% of transmissions were thought to have occurred on a previous admission during the study period (up to 12 months prior to first sample collection), whilst another 17·5% of transmissions were likely to have occurred on a previous ward during the same admission as sample collection; only 34·7% of transmissions were likely to be from the same ward and admission when the MDRO sample was collected.

Hospital networks varied in the MDRO species with the highest transmission rates

Transmission rates per 100,000 OBDs varied widely between networks and species (Figure 3b). The greatest difference between networks was observed in ESBL-Ec, with network A having 47·5 transmissions per 100,000 OBDs, compared with the other two large networks (network B, 4·9 and network C, 5·4 transmissions per 100,000 OBDs) and zero transmissions for network D (incidence difference between network A and networks B and C, 43·14 transmissions per 100,000 OBDs (95% CI 37·54-48·75, p<0·001)). This may partly reflect differences in screening practices (see Table 1), especially the network-wide point-prevalence surveys for Network A which contributed 27·2% of their transmissions (18·5% of ESBL-Ec, 62·9% of *vanA* VREfm and 11·1% of ESBL-Kp).

Conversely, network B registered a higher rate of *vanA* VRE transmission than other networks (transmissions per 100,000 OBDs: network B, 34·2; network A, 22·3; network C, 14·3; network D, 2·4; incidence difference between network B and networks A and C, 15·69 transmissions per 100,000 OBDs (95% CI 9·54-21·84, p<0·001)). MRSA transmission was also higher for network B, but these data likely underestimate the true MRSA transmission rates for all networks due to low rates of routine MRSA screening.

Successful vanA *VREfm clones spread widely within and between hospital networks*

Eight genomic clusters of *vanA* VREfm were detected in more than one hospital network (Figure 4); some were spread across networks from the commencement of the study (e.g. ST1421 cluster 1), whilst others emerged during the study, initially at one site, then other sites (e.g. ST1424 cluster 2). Whilst this study was not designed to capture transmissions between hospitals, and the state health

information systems are decentralised (limiting tracking of patient movement between hospitals), this strongly suggests that patient movement between networks also contributes to the dissemination of *vanA* VRE in study sites, and likely more widely in Victoria.

Impacts of genomics program

The implementation of genomic surveillance for hospital MDROs led to a wide range of impacts, falling into two main groups: identifying unexpected MDRO transmissions requiring targeted infection control interventions, or supporting the management of existing outbreaks through the provision of high-resolution genomic data to clearly define transmission networks (Table 4). The genomics program was particularly useful in identifying MDRO outbreaks occurring over a longer period of time, which had not been detected by conventional surveillance methods. Outbreak investigations were supported by including genomic analysis of historical and environmental isolates (including shared patient equipment), allowing infection control teams to trace the likely transmission pathways, and implement targeted measures to prevent further spread. This was especially useful where epidemiological links were complex, such as when a patient may have been exposed to *vanA* VREfm in more than one location, making attribution of MDRO acquisition impossible with epidemiologic data alone. Furthermore, genomics allowed the cessation of screening in one *vanA* VREfm outbreak, by demonstrating that new cases outside this ward were not related to the original outbreak, thus saving ~~significant~~ considerable resources and allowing the ward to re-open to new admissions.

This study also had the unanticipated effect of encouraging data sharing between study sites, resulting in identification of one network with a higher rate of *vanA* VREfm transmissions than the others. This led to a review of cleaning practices, which were subsequently changed for the whole network towards the end of the study period. Additionally, data from this study also informed public health actions, with *vanA* VREfm newly added to the list of notifiable conditions by the state health department[24].

Workflow implementation issues

A number of challenges for program implementation were identified throughout the study, primarily those causing delays in turnaround times or difficulty accessing and integrating electronic patient data (Table 5). Many of these issues relate specifically to this study, and would not be expected to limit the future implementation of a genomics workflow for infection control.

Whilst several factors contributed to delays in turnaround times (TAT), two factors dominated; i) delays in hospital laboratory sending isolates to the sequencing laboratory (mostly due to unusual

workflow disruptions in the hospital laboratories), and ii) delays in reporting results to hospitals, as we elected to report results in person (to encourage engagement and education), rather than releasing results cumulatively (Supplementary Figure S5). Established systems for rapid referral, sequencing and reporting of notifiable pathogens in our jurisdiction, such as carbapenemase-producing Enterobacterales (CPE) and *Salmonella,* are usually complete in 7 to 10 days. Once MDRO infection control sequencing is fully embedded using routine referral processes, it is anticipated that these delays experienced for some samples in this this study would be unlikely to occur.

Communication

To design an effective written report format for infection control genomics, feedback from focus group participants at the conclusion of the study was used to design a template for reporting aggregated hospital/network-level genomics data (e.g. monthly report for each network)(Figure 5 and Supplementary Figure S6). Recurrent themes included a desire for simplified results, strong preference for use of a 'traffic light' system to classify likelihood of transmission, general preference to avoid inclusion of phylogenetic trees (although not amongst more genomics-experienced participants), and specific comments directing ICPs to which patients to investigate for possible transmission.

Discussion

Here we have presented a large, multicentre study implementing a new workflow for prospective genomic surveillance of MDROs to inform infection control interventions in hospitals. We have demonstrated wide-ranging impacts of prospective genomic surveillance, identifying MDRO transmission where it was not suspected, providing support for outbreak investigations, and data to inform public health coordination of hospital-acquired infections. We believe that this is a critical step towards changing how we think about MDROs in hospitals, no longer relying on low-resolution traditional microbiological and epidemiologic surveillance, but proactively detecting MDRO transmission at a genomic level. This then enables early and tailored infection control interventions, potentially limiting the number of patients infected, morbidity and mortality, and saving limited hospital resources.[4,18,25,26] This is most relevant with more common MDROs found in both hospitals and the community, as hospital transmission may be under-estimated without precision genomic data.[4] Importantly, most of the MDRO transmission detected here would not have been identified by existing infection control surveillance, offering huge opportunities to detect silent MDRO transmission, and intervene.

Implementing a new workflow represents a ~~significant~~ major change for hospital laboratories and infection control teams, and identifies challenges to be overcome for successful implementation. One of the pressure points for successful implementation is minimising turnaround times, critical for timely and effective infection control responses. COVID-19 has demonstrated what TATs for sequencing may be achievable, with 5-7 days TATs routinely attained in our state (1-3 days for urgent samples).[27] With adequate resourcing and buy-in from hospital laboratories, and capacity in sequencing laboratories, similar TATs could be achieved for MDRO surveillance. Alternatively, a decentralised genomics model (local WGS and analysis, or hybrid model with centralised analysis) may also optimise TATs; the choice of model ultimately depends on the available resources in each setting.[28] Other key challenges include the optimisation of bioinformatics processes[29] and integration of epidemiologic data, without which the genomic analyses are not interpretable.[30,31] In our study, the linkage of genomic and epidemiologic data was impaired by poor integration with electronic medical records; this integration would be an important component for successful implementation in the future.[32]

Communication of complex genomic results to a new clinical setting is another challenge for implementation. Ideally, results should be easy to comprehend, and easily integrated with epidemiologic data to facilitate rapid interpretation, enabling rapid infection control intervention where required.[33] As an early implementation study in a group of hospitals unfamiliar with genomics, we chose to engage infection control teams by presenting results in-person to educate and answer questions interactively, with the trade-off of delayed reporting times. Whilst this approach was very useful for early implementation and engagement, alternative reporting methods would be more suitable going forward, such as regular written reports designed for infection control (such as the example provided here). Further innovation to allow infection control teams rapid access to genomics results and interpretation would be welcome in future implementation studies.

Identifying strong epidemiologic evidence of transmission is difficult, particularly with common MDROs, but necessary to validate genomic inferences of MDRO transmission. In this study, we used genomic links (SNP thresholds) and epidemiologic data (overlapping admissions in space and time) to call probable or possible transmission, similar to other studies.[25,34] SNP thresholds are a relatively blunt instrument to suggest potential genomic links, but core genome SNP analyses are accepted to be far higher resolution than existing typing methods (such as MLST), and at least equivalent to core genome MLST.[35,36] Importantly, SNP thresholds are specific to the methods used to derive them, hence other studies using lower thresholds are not directly comparable to this study.[32,34] In a pragmatic sense, if high-resolution core genome SNP methods are used for large-scale prospective

genomics, then SNP thresholds are likely to be required; however, future studies in this area should be open to exploring alternative methods and thresholds.[29,34,35,37]

Our study has a number of limitations. Firstly, screening practices were not standardised between hospitals, and included one site performing a biannual hospital-wide point-prevalence survey. However, many transmissions were still identified at hospitals performing less screening; this suggests the potential advantages of additional screening to uncover unsuspected transmissions. Secondly, bed movement data would ideally have been acquired for all study patients; this was addressed in the pilot study and did not greatly change the interpretation of the transmission data.[19] Third, these methods are not designed to detect transmission of AMR plasmids mechanisms, as short-read sequencing methods are limited in this ability.

Limited access to data (clinical data and centralised state admissions data) was also a limitation of the study, and largely prevented identification of inter-hospital transmission; with the increasing implementation of electronic medical records, we hope that these issues could be more comprehensively addressed in future studies.[30] Turnaround times were also impacted by extraneous factors (laboratory delays); despite this, ~~significant~~ major impacts on infection control and public health were demonstrated, and future changes in resource management and system design have now been proposed to improve TATs. Lastly, the cost-effectiveness of this intervention has not been addressed in this study, but demonstrated elsewhere,[18,25] and will be the subject of future work.

This study represents a framework considering the critical components and challenges for implementation of prospective genomics for infection control, and may be tailored to local needs and conditions, encompassing local MDROs of interest, different hospital and network structures, laboratory set-up, and screening practices. Further implementation studies should focus on reducing TATs for more real-time results, streamlined data integration and reporting systems, as well as furthering the education of hospital infection control, infectious diseases and laboratory teams to improve patient safety, and care through genomics-informed infection control in the near future.

Contributors

NLS, JCK, MLG & BPH conceived and designed the study. CG, RL, TS & CH designed and performed bioinformatic analyses and created figures. SB & MLS coordinated sequencing laboratory workflows. TK, MG, ML & HC arranged for isolates to be sent to the sequencing laboratory and provided microbiological data. RS, CM, MAS, LW & JCK coordinated collection of clinical and epidemiologic data. NLS, CG, JCK, RS, CM, TK, MAS, RL, MG, ML, LW, HC, MLG & BPH reviewed the data during the study, had input into the ongoing performance of the study and provided feedback on the manuscript. NLS and CLG wrote the manuscript; BPH, MLG and JCK edited the manuscript.

Declaration of Interests

The authors declare no conflicts of interest.

Acknowledgements

Financial support. This work was supported by the Melbourne Genomics Health Alliance (funded by the State Government of Victoria, Department of Health and Human Services, and the 10 member organizations); and individual grants from National Health and Medical Research Council (Australia) to NLS (GNT1093468), JCK (GNT1008549), BPH (GNT1105905), and MAS (GNT1116576).

Other members of the Controlling Superbugs Study Group include: Elizabeth Grabsch (Microbiology, Austin Health), Joanna Price and Carolyn Mason (Infection Control, Austin Health), Despina Kotsanas (Monash Infectious Diseases), Louise Wright (Infection Control, Monash Health), Dr Suraya Hanim Abdullah Hashim and Jennifer Mitchell (Infectious Diseases, Melbourne Health), Olivia Smibert and Victoria Madigan (Infectious Diseases, Peter MacCallum Cancer Centre) and Peter Pham (MDU Public Health Laboratory, University of Melbourne).

The authors would like to gratefully acknowledge the assistance of the following people: MDU PHL staff including Kerrie Stevens, Nicole Orlando, Samantha Tawil, Savitra Rambocus and Diane Daniel; John Greenough (infection control queries and patient movement data), Ben Rogers and Jean Lee (focus groups).

Data sharing

Raw sequence data for all included study samples (Supplementary Table S1), are available in NCBI Sequence Read Archive (SRA), BioProject PRJNA565795. All software and code used is publicly available using links provided in the Methods section.

References

1. World Health Organization. Antimicrobial resistance: global report on surveillance. Geneva, Switzerland: WHO Press, 2014.
2. Chia PY, Sengupta S, Kukreja A, S.L. Ponnampalavanar S, Ng OT, Marimuthu K. The role of hospital environment in transmissions of multidrug-resistant gram-negative organisms. *Antimicrob Resist Infect Control* 2020; **9**(1).
3. Kritsotakis E, Kontopidou F, Astrinaki E, Roumbelaki M, Ioannidou E, Gikas A. Prevalence, incidence burden, and clinical impact of healthcare-associated infections and antimicrobial resistance: a national prevalent cohort study in acute care hospitals in Greece. *Infect Drug Resist* 2017; **10**: 317-28.
4. Popovich KJ, Snitkin ES. Whole genome sequencing-implications for infection prevention and outbreak investigations. *Curr Infect Dis Rep* 2017; **19**(4): 15.
5. Parcell BJ, Gillespie SH, Pettigrew KA, Holden MTG. Clinical perspectives in integrating whole-genome sequencing into the investigation of healthcare and public health outbreaks – hype or help? *J Hosp Infect* 2021; **109**: 1-9.
6. Ruppe E, Olearo F, Pires D, et al. Clonal or not clonal? Investigating hospital outbreaks of KPC-producing *Klebsiella pneumoniae* with whole-genome sequencing. *Clin Microbiol Infect* 2017; **23**(7): 470-5.
7. Willems S, Kampmeier S, Bletz S, et al. Whole-genome sequencing elucidates epidemiology of nosocomial clusters of *Acinetobacter baumannii*. *J Clin Microbiol* 2016; **54**(9): 2391-4.
8. Tang P, Croxen MA, Hasan MR, Hsiao WW, Hoang LM. Infection control in the new age of genomic epidemiology. *Am J Infect Control* 2017; **45**(2): 170-9.
9. Conlan S, Thomas PJ, Deming C, et al. Single-molecule sequencing to track plasmid diversity of hospital-associated carbepenemase-producing *Enterobacteriaceae. Sci Transl Med* 2014; **6**: 254ra126.
10. 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**(2): 130-6.
11. De Man TJB, Yaffee AQ, Zhu W, et al. Multispecies outbreak of Verona integron-encoded metallo-ß-lactamase producing multidrug-resistant bacteria driven by a promiscuous incompatibility group A/C2 plasmid. *Clin Infect Dis* 2021; **72**(3): 414-20.
12. Peacock SJ, Parkhill J, Brown NM. Changing the paradigm for hospital outbreak detection by leading with genomic surveillance of nosocomial pathogens. *Microbiology* 2018; **164**(10): 1213-9.

13. Reuter S, Ellington MJ, Cartwright EJ, et al. Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology. *JAMA Intern Med* 2013; **173**(15): 1397-404.
14. Barrasa-Villar JI, Aibar-Remón C, Prieto-Andrés P, Mareca-Doñate R, Moliner-Lahoz J. Impact on morbidity, mortality, and length of stay of hospital-acquired infections by resistant microorganisms. *Clin Infect Dis* 2017; **65**(4): 644-52.
15. Mollers M, Lutgens SP, Schoffelen AF, Schneeberger PM, Suijkerbuijk AWM. Cost of nosocomial outbreak caused by NDM-1–containing *Klebsiella pneumoniae* in the Netherlands, October 2015–January 2016. *Emerg Infect Dis* 2017; **23**(9): 1574-6.
16. Otter JA, Burgess P, Davies F, et al. Counting the cost of an outbreak of carbapenemase-producing Enterobacteriaceae: an economic evaluation from a hospital perspective. *Clin Microbiol Infect* 2017; **23**(3): 188-96.
17. Balloux F, Bronstad Brynildsrud O, van Dorp L, et al. From theory to practice: translating whole-genome sequencing (WGS) into the clinic. *Trends Microbiol* 2018; **26**(12): 1035-48.
18. Gordon LG, Elliott TM, Forde B, et al. Budget impact analysis of routinely using whole-genomic sequencing of six multidrug-resistant bacterial pathogens in Queensland, Australia. *BMJ Open* 2021; **11**(2): e041968.
19. Sherry NL, Lee RS, Gorrie CL, et al. Pilot study of a combined genomic and epidemiologic surveillance program for hospital-acquired multidrug-resistant pathogens across multiple hospital networks in Australia. *Infect Control Hosp Epidemiol* 2020; **42**(5): 573-81.
20. World Health Organization. Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics. Geneva: WHO; 2017.
21. Sherry NL, Lane CR, Kwong JC, et al. Genomics for molecular epidemiology and detecting transmission of carbapenemase-producing Enterobacterales in Victoria, Australia, 2012 to 2016. J Clin Microbiol 2019; **57**(9).
22. Department of Health and Human Services Victoria. Victorian guideline on carbapenemase-producing Enterobacteriaceae for health services (version 2.1). 2018. https://www2.health.vic.gov.au/public-health/infectious-diseases/infection-control-guidelines/carbapenemase-producing-enterobacteriaceae-management/cpe-for-health-services.
23. Voor In 't Holt AF, Wattel AA, Boers SA, et al. Detection of healthcare-related extended-spectrum beta-lactamase-producing *Escherichia coli* transmission events using combined genetic and phenotypic epidemiology. *PLoS One* 2016; **11**(7): e0160156.
24. State Government of Victoria. Public Health and Wellbeing Regulations 2019, Statutory Rule no. 139/2019 v004. 2019.

25. Mellmann A, Bletz S, Boking T, et al. Real-time genome sequencing of resistant bacteria provides precision infection control in an institutional setting. *J Clin Microbiol* 2016; **54**(12): 2874-81.
26. Clermont O, Christenson JK, Denamur E, Gordon DM. The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ Microbiol Rep* 2013; **5**(1): 58-65.
27. Lane CR, Sherry NL, Porter AF, et al. Genomics-informed responses in the elimination of COVID-19 in Victoria, Australia: an observational, genomic epidemiological study. *Lancet Pub Health* 2021; doi:10.1016/s2468-2667(21)00133-x.
28. Mintzer V, Moran-Gilad J, Simon-Tuval T. Operational models and criteria for incorporating microbial whole genome sequencing in hospital microbiology - A systematic literature review. *Clin Microbiol Infect* 2019; **25**(9): 1086-95.
29. Gorrie CL, Da Silva AG, Ingle DJ, et al. Key parameters for genomics-based real-time detection and tracking of multidrug-resistant bacteria: a systematic analysis. *Lancet Microbe* 2021: doi:10.1016/S2666-5247(21)00149-X.
30. Pak TR, Kasarskis A. How next-generation sequencing and multiscale data analysis will transform infectious disease management. *Clin Infect Dis* 2015; **61**(11): 1695-702.
31. Armstrong GL, Maccannell DR, Taylor J, et al. Pathogen genomics in public health. *New Engl J Med* 2019; **381**(26): 2569-80.
32. Ward DV, Hoss AG, Kolde R, et al. Integration of genomic and clinical data augments surveillance of healthcare-acquired infections. *Infect Control Hosp Epidemiol* 2019; **40**(6): 649-55.
33. Crisan A, McKee G, Munzner T, Gardy JL. Evidence-based design and evaluation of a whole genome sequencing clinical report for the reference microbiology laboratory. *PeerJ* 2018; **6**: e4218.
34. Gouliouris T, Coll F, Ludden C, et al. Quantifying acquisition and transmission of *Enterococcus faecium* using genomic surveillance. *Nat Microbiol* 2021; **6**(1): 103-11.
35. Neumann B, Bender JK, Maier BF, et al. Comprehensive integrated NGS-based surveillance and contact-network modeling unravels transmission dynamics of vancomycin-resistant enterococci in a high-risk population within a tertiary care hospital. *PLoS One* 2020; **15**(6): e0235160.
36. Pinholt M, Larner-Svensson H, Littauer P, et al. Multiple hospital outbreaks of *vanA Enterococcus faecium* in Denmark, 2012-13, investigated by WGS, MLST and PFGE. *J Antimicrob Chemother* 2015; **70**(9): 2474-82.
37. Schurch AC, Arredondo-Alonso S, Willems RJL, Goering RV. Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene-based approaches. *Clin Microbiol Infect* 2018; **24**(4): 350-4.

38. Attwood LO, Sherry NL, Gorrie C, et al. Active surveillance to predict superbug infections (conference presentation). Australian Society for Infectious Diseases (ASID) conference. Melbourne, Victoria; 2020.

Table 1. Characteristics and infection control practices of hospital sites included in study

Hospital network	Hospital code	Hospital description	No. of inpatient beds[a]	High-risk wards	MDRO screening practices during study period and changes during study	Management of patients colonised or infection with MDRO
A	**A1**	Tertiary referral center, including ICU, solid-organ and bone marrow transplant	560	ICU Hematology/BMT and Oncology Renal Transplant Liver Transplant (Spinal ward and respiratory ward added)	*vanA* VRE and MRGN screening: ICU, on admission and twice weekly; haematology/oncology, renal and liver transplant wards screened on admission and weekly Additional MRSA screening in ICU (on admission and twice weekly) Biannual point-prevalence survey for *vanA* VRE and MRGN MRSA screening before critical surgeries (prosthetic joint, spinal and cardiac) *Change during study:* Added spinal ward and respiratory ward (ventilator support service) to high-risk wards for regular MDRO screening (October 2018)	MRSA: patients only isolated in certain circumstances, e.g. highly exudative wound; single room, disposable apron, gloves *vanA* VRE: single room with own bathroom, full gown, gloves; sometimes cohorted[b] with other *vanA* VRE-colonised patients ESBL-Ec and ESBL-Kp (multidrug-resistant)[c]: single room, own bathroom, disposable apron
	A2	Subacute hospital, aged care and rehabilitation services	150	None	Biannual point-prevalence survey for *vanA* VRE and MRGN	
	A3	Subacute hospital, rehabilitation services	60	None	Biannual point-prevalence survey for *vanA* VRE and MRGN	
B	**B1**	Tertiary referral center, including ICU and solid-organ transplant and specialist pediatric hospital (including neonatal ICU)	640	ICU Renal Transplant	ICU and renal ward screened for *vanA* VRE and carbapenem-resistant Gram negatives (CRGN) on admission and weekly MRSA screening before cardiac surgery	MRSA: single room, gloves and short-sleeved gown *vanA* VRE: single room, own bathroom, occasional cohorting[b] with other *vanA* VRE-colonised

					Change during study: Stopped routine screening of renal ward for VRE (June 2018). Network-wide changes in cleaning practices to from microfibre/steam cleaning to bleach cleaning (September 2018)	patients; gloves and short-sleeved gown. Additional measures if diarrhoea (full gown) ESBL-Ec: no specific management measures ESBL-Kp: single room, own bathroom, gloves and short-sleeve gown
	B2	Tertiary referral center, including ICU, trauma and some aged care & rehabilitation services	573	ICU	ICU patients screened for *vanA* VRE and carbapenem-resistant Gram negatives (CRGN) on admission and weekly *Changes during study:* Cleaning protocol changes as above	
C	**C1**	Tertiary referral center, including ICU, solid-organ and bone marrow transplant	571	ICU Haematology/BMT	ICU and haematology ward screened on admission and weekly for *vanA* VRE and MRGN	MRSA: single room only in certain situations, e.g. highly exudative wound *vanA* VRE: single room, own bathroom, gloves and full gown ESBL-Ec (multidrug-resistant)[c] and ESBL-Kp: single room, own bathroom, gloves and full gown
	C2	Subacute hospital, aged care and rehabilitation services	150	None	None	
D	**D1**	Specialized cancer care center. Located adjacent to Hospital C1 (ICU patients cared for at C1 before transfer back to hospital D1)	96	Hematology	Haematology ward patients screened on admission and weekly for *vanA* VRE and MRGN	MRSA: single room when possible; if open wounds, gown and gloves; if respiratory, mask and gloves *vanA* VRE: single room, own bathroom, gloves and full gown ESBL-Ec and ESBL-Kp: single room, own bathroom, no extra PPE but additional cleaning if incontinent or discharging wounds

ICU, intensive care unit; MRGN, multi-resistant Gram negatives (includes ESBL and carbapenem-resistant phenotypes); BMT, bone marrow transplant (allogeneic); PPE, personal protective equipment. [a]Inpatient beds, excludes day cases, hospital-in-the-home and mental health; [b]Cohorting, practice of patients colonised with the same MDRO sharing a room; [c]Multidrug-resistant ESBL, ESBL isolates with additional resistance aminoglycosides, fluoroquinolones and trimethoprim-sulfamethoxazole (at least two classes).

Table 2. Considerations for implementation of genomics-informed infection control workflow

Component of genomics workflow	Key considerations	Implementation strategy selected for this study
Target population	Consider targeting high-risk populations (e.g. ICU, haematology/oncology, transplant) vs acute inpatients vs all inpatients (acute and subacute)	Included all inpatients due to frequent movement between wards, and unknown burden of disease and transmission in subacute care
Target MDRO/s	Consider AMR spectrum/high-risk organisms; MDROs not covered by existing surveillance programs; organisms where healthcare acquisition more likely than community acquisition; MDROs of high prevalence in local population	Selected *van A* VRE (high prevalence, recently emerged), ESBL-Ec and ESBL-Kp (unclear healthcare contribution) and MRSA (changing epidemiology, unclear healthcare contribution) Note CPE already covered by existing statewide surveillance system; CRPa and CRAb low prevalence
MDRO screening strategy	Consider using existing MDRO screening strategies vs harmonised screening approaches. Could target high-risk populations or all inpatients (e.g. ward-based or hospital-wide point-prevalence surveys)	Elected for pragmatic approach, with each hospital network continuing existing screening strategies, noting differences between networks in analysis
Sample types	All samples vs clinical samples only (latter would underestimate transmission)	All samples included to maximise likelihood of identifying transmission
Clinical data collection	Consider data required to identify likely transmission in hospital, including admissions at other sites; consider ability to accurately identify admissions at different networks by centralised electronic systems	Patient movement data (admissions, wards and beds) critical to identify likely transmissions in conjunction with genomic data. Minimal clinical data collected in implementation phase, as not required to infer transmission.
Sequencing and analysis strategy	(i) Sequencing location: centralised (sequencing laboratory) vs decentralised (hospital laboratories) (ii) Bioinformatics analysis location: centralised vs decentralised, based on resources and expertise	(i) Sequencing in central laboratory, as sequencing not established at study sites (ii) Centralised analysis at sequencing laboratory as not established at study sites

	(iii) Analysis of combined genomic and epidemiologic data, depending on IPC resources and training, and established systems for electronic data analysis	(iii) Genomic and epidemiologic data integrated centrally by study team, presented to IPC teams for further interpretation and action
Bioinformatics analysis	Consider methods for transmission analysis. Examples: core genome SNP analysis, within species or within ST vs cgMLST approach, SNP thresholds, masking recombination Other considerations: frequency of analysis, inclusion of all data vs rolling window (e.g. last 6 or 12 months)	Elected for core genome SNP analysis within ST based on existing accredited bioinformatics pipelines; recombination masked except for MRSA; SNP thresholds selected based on existing data at the time. All samples included in each analysis (approximately monthly), noting computational expense and personnel requirements
Communication of results	Consider audience (IPC/ID/microbiology teams) and understanding of genomics Written reports (frequency) vs discussions in person vs interactive data portal with effective data visualisation	Chose to engage with IPC/ID/microbiology teams using in-person presentations as part of initial implementation, and development of written report format via focus group

MDRO, multidrug-resistant organism; VRE, vancomycin-resistant *Enterococcus*; ESBL-Ec, extended-spectrum beta-lactamase-phenotype *Escherichia coli*; ESBL-Kp, extended-spectrum beta-lactamase-phenotype *Klebsiella pneumoniae*; MRSA, methicillin-resistance *Staphylococcus aureus*; ST, sequence type; cgMLST, core genome multi-locus sequence type; SNP, single nucleotide polymorphisms; IPC, infection prevention and control; ID, infectious diseases.

Table 3. Patient and isolate characteristics

Total no. of isolates included	2275		
Total no. of patients included	1970		
Age (median, IQR)	65 years (47-77yrs)		
Male sex (no., %)	1063 (56·8%)		
Sample collected for clinical purposes (suspected infection)	1428 (62·8%)		
Sample collected for MDRO screening purposes	847 (37·2%)		
Time from admission to isolate collection (median, IQR)	3 days (1-12 days)		
Ward where sample was collected	**All samples (% total samples)**	**Clinical samples (% per ward)**	**Screening samples (% per ward)**
Intensive care unit	373 (16·4%)	117 (31·4%)	256 (68·6%)
Other high-risk ward [a]	434 (19·1%)	121 (27·9%)	313 (72·1%)
General acute ward [b]	1090 (47·9%)	865 (79·4%)	225 (20·6%)
Emergency department	209 (9·2%)	204 (97·6%)	5 (2·4%)
Subacute care ward [c]	169 (7·4%)	121 (71·6%)	48 (28·4%)

MDRO, multidrug-resistant organism; IQR, interquartile range. [a] High-risk wards, includes hematology, oncology, renal ward (including renal transplant), and liver transplant wards.

[b] General acute wards, include all wards not designated as ICU, high-risk or subacute.

[c] Subacute care, includes aged care, rehabilitation, palliative care and spinal wards.

Table 4. Impacts of genomics program for hospital MDRO surveillance

Unexpected MDRO transmission identified, resulting in specific ward interventions	• *vanA* VREfm cluster identified in haematology ward • ESBL-Kp 'super-spreader' patient identified in one ward[a] • Intensive ward cleaning instituted ('super-clean')[b] • Review of patient placement/movements in affected wards • Increased surveillance in affected wards[c]
Outbreak investigation support	• Supported patient-to-patient transmission and implication of shared patient equipment in outbreaks[d] • Identified transmission where epidemiological links were complex • Able to show that new cases in other hospital wards were not linked to an outbreak, demonstrating that the outbreak management had been successful and allowing infection control interventions to be stepped down
Comparison of MDRO transmission rates between networks, resulting in network-wide changes	• Identified higher rate of *vanA* VRE transmission in one network than others, resulting in network-wide change in cleaning practices[e]
Increased awareness of MDRO transmission patterns	• Increased awareness of MRSA and ESBL-Ec transmissions in hospital (previously thought to all be community-acquired) • Confirmed that MDRO identified through screening predicted the MDRO that subsequently caused clinical infection. Useful for advising empiric therapy e.g. in neutropaenic sepsis[38]
Public health impacts	• Large burden of hospital-acquired *vanA* VREfm demonstrated, contributing to new legislative requirement[24] for all *vanA* VREfm cases to be notified to the state health department • Low burden of carbapenem-resistant *Pseudomonas aeruginosa* and *Acinetobacter baumannii* established as baseline before inclusion in state notifiable diseases list

[a] 'Super-spreader' indicating source patient with unusually high number of transmissions to other patients

[b] Ward 'super-cleans' included thorough terminal/discharge-type cleaning for all areas of the ward (including communal areas, corridors, offices, and storerooms) using a bleach-based detergent where possible.

[c] Increased surveillance included additional environmental screening, additional patient screening (added to roster of high-risk wards with weekly patient screening for VREfm and MRGN, and additional focused hand hygiene surveillance

[d] Shared patient equipment includes any equipment that is not restricted to a single patient, e.g. blood glucometers, blood pressure cuffs, weigh chairs

[e] Cleaning practices changed from steam wands and microfibre cloth cleaning to bleach cleaning

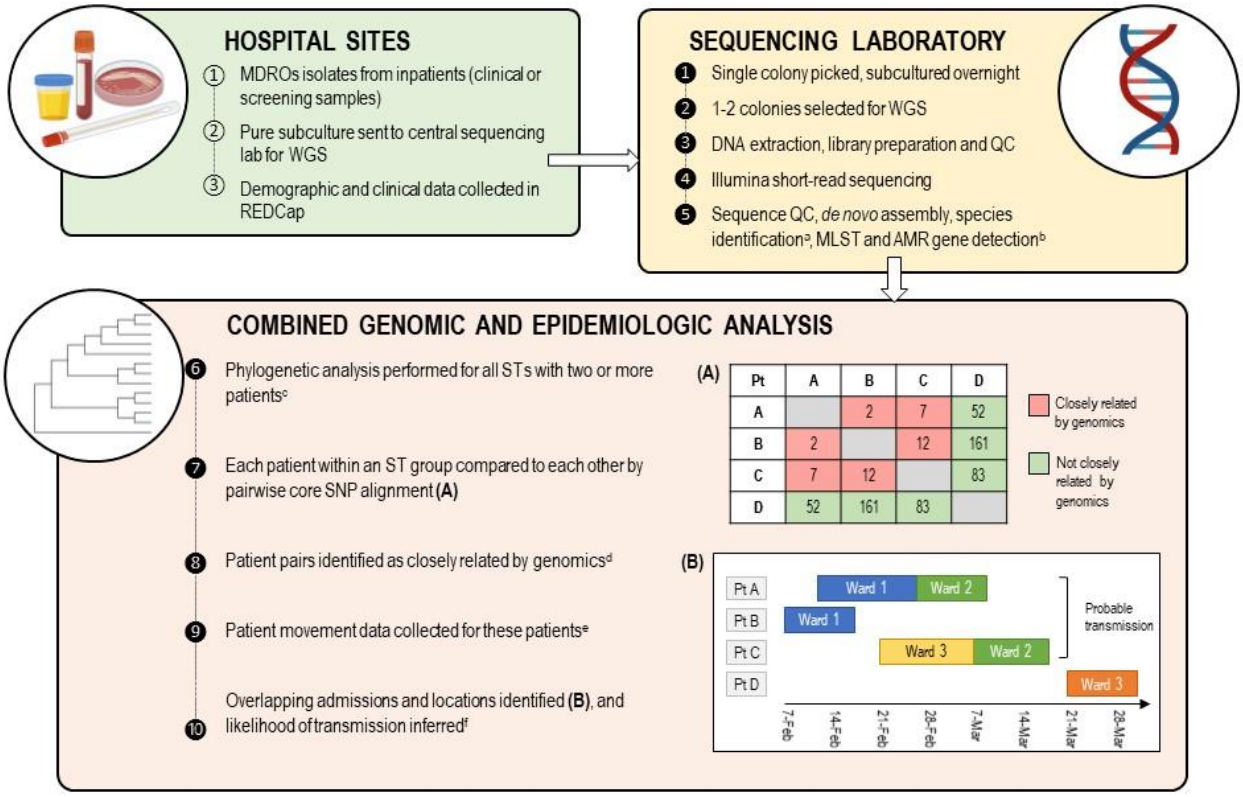
Table 5. Areas for further development for implementation of infection control genomics workflow

Issue and domain	Description	Potential solutions
Delays in time to results		
a. Hospital laboratory	Delayed time to send isolates to sequencing laboratory • Inadequate staffing • Batch processing of isolates (weekly or longer intervals) • ~~Significant~~ Major disruptions from introduction of new LIS at two laboratories (inability to perform data extractions for study, affecting three sites)	Dedicated staff (supervising microbiologist and technician) More frequent processing of isolates for transit to sequencing lab Ensure robust data extraction capabilities available when implementing new LIS
b. Sequencing laboratory	Minor delays, due to high volume of samples and lower relative priority	Increase in sequencing capacity, redundancy to account for surges in sample volume
c. Bioinformatic analysis	Delays in analysis due to inadequate staffing Increasing complexity of analysis and computational requirements towards end of study due to large isolate numbers (all study samples of same MLST included in each transmission analysis)	Increase dedicated computing resources for complex bioinformatic analyses Work towards partial automation of transmission analysis Develop alternative bioinformatics methods to deal with genomics 'big data' Explore effect of rolling windows (e.g. only including last 6 or 12 months of isolates) on detecting transmission
d. Reporting process	In-person presentations to each network used in this study (to engage and educate staff at study sites) did not allow for results to be conveyed in a timely manner	Issue individual reports for each patient isolate, including MLST result (can be used by Infection Control teams to rapidly exclude or include patients in potential transmissions) New report designed for infection control genomics results (aggregated hospital or network-level data) from focus groups Develop innovative approaches for reporting aggregated results on a hospital/network level e.g. interactive dashboards with advanced data visualisation capabilities
Impaired access to patient epidemiologic data	Each hospital network using different patient information management systems; no readily available way to extract patient data; data had to be collected manually	Develop informatics systems to reliably extract patient movement data, and integrate with results of transmission analyses to inform Infection Control interventions

(including admission, ward and bed data)	(transcription errors, variable formatting requiring manual correction)	
Inability to track patient movements between hospital networks	No readily-available centralised system available for tracking admissions of patients across different hospital networks to identify inter-network MDRO transmission	Engage with health department to advocate for development of centralised admissions database, available in real-time
Perceived cost issues	Sequencing costs perceived to be prohibitive for widespread application currently	Optimise utility of prospective genomics by targeting to local MDROs or locations of concern; demonstrate cost-effectiveness of prospective genomics

LIS, laboratory information system; MLST, multi-locus sequence type (a typing method to subset isolates of the same species into smaller, closely-related groups for targeted transmission analysis); MDRO, multidrug-resistant organism

Figure 1. Infection control genomics workflow implemented in this study



[a] Species identification by *k*-mer identification (*kraken*); *Klebsiella* further subspeciated using *kleborate* tool.

[b] Sequences analysed for presence of complete *vanA* operon, ESBL/AmpC genes and *mecA/mecC*; if these were absent, isolates underwent further phenotypic testing to ensure they met inclusion criteria (phenotypic antimicrobial resistance).

[c] Aligning isolates to reference genome of same ST, or *de novo* assembly of earliest isolate if this was not available; ST 131 *E. coli* analysed in two subclades due to large number of isolates. Recombinant sites were masked with *gubbins* for species other than *S. aureus.*

[d] Each isolate sequence within an ST was compared to all others, and closely-related isolates were determined by core SNP differences; ≤15 SNPs for *Staphylococcus aureus,* ≤25 SNPs for other species

[e] Data collected for 12 months prior to first study sample until the end of study

[f] Likelihood of transmission inferred from combined genomic and epidemiologic data, categorised as 'Probable' (same ward at same time), 'Possible' (same ward at different time but within 60 days, or same hospital at the same time), 'Unlikely' (neither of the above), or 'Above screening threshold' (not closely related by genomics).

(A) Example of pairwise SNP distance matrix for sequences from four patient isolates; patient pairs with pairwise SNP distances at or below the screening threshold (15 SNPs for MRSA, 25 SNPs for other MDROs) are classified as closely related by genomics (red), and pairs above these thresholds are designated as not closely related by genomics. **(B)** Example of Gantt chart demonstrating bed movements (ward admissions) of same four patients from (A) over time; patients A, B and C had closely related MDROs by genomics, and had overlapping admissions (same ward at the same time), therefore constituting probable MDRO transmission. Patient D's sequence was not related to the other patients' sequences by genomics (above screening threshold), and not considered to be involved in MDRO transmission to or from these other patients, despite having an MDRO of the same sequence type (ST). Abbreviations: MDRO, multidrug-resistant organism; WGS, whole genome sequencing; QC, quality control; MLST, multi-locus sequence typing; AMR, antimicrobial resistance; ST, sequence type; SNP, single-nucleotide polymorphism (single base difference between two or more isolates); Pt, patient.

Figure 2. MDRO isolates by network, species, and reason for sample collection. Bars are shaded by reason for sample collection with lightest shades representing clinical samples, middle shades representing screening samples, and darkest shades (only for Network A) representing screening samples performed as part of network-wide biannual point-prevalence surveys (PPS)(note, MRSA not included in PPS).

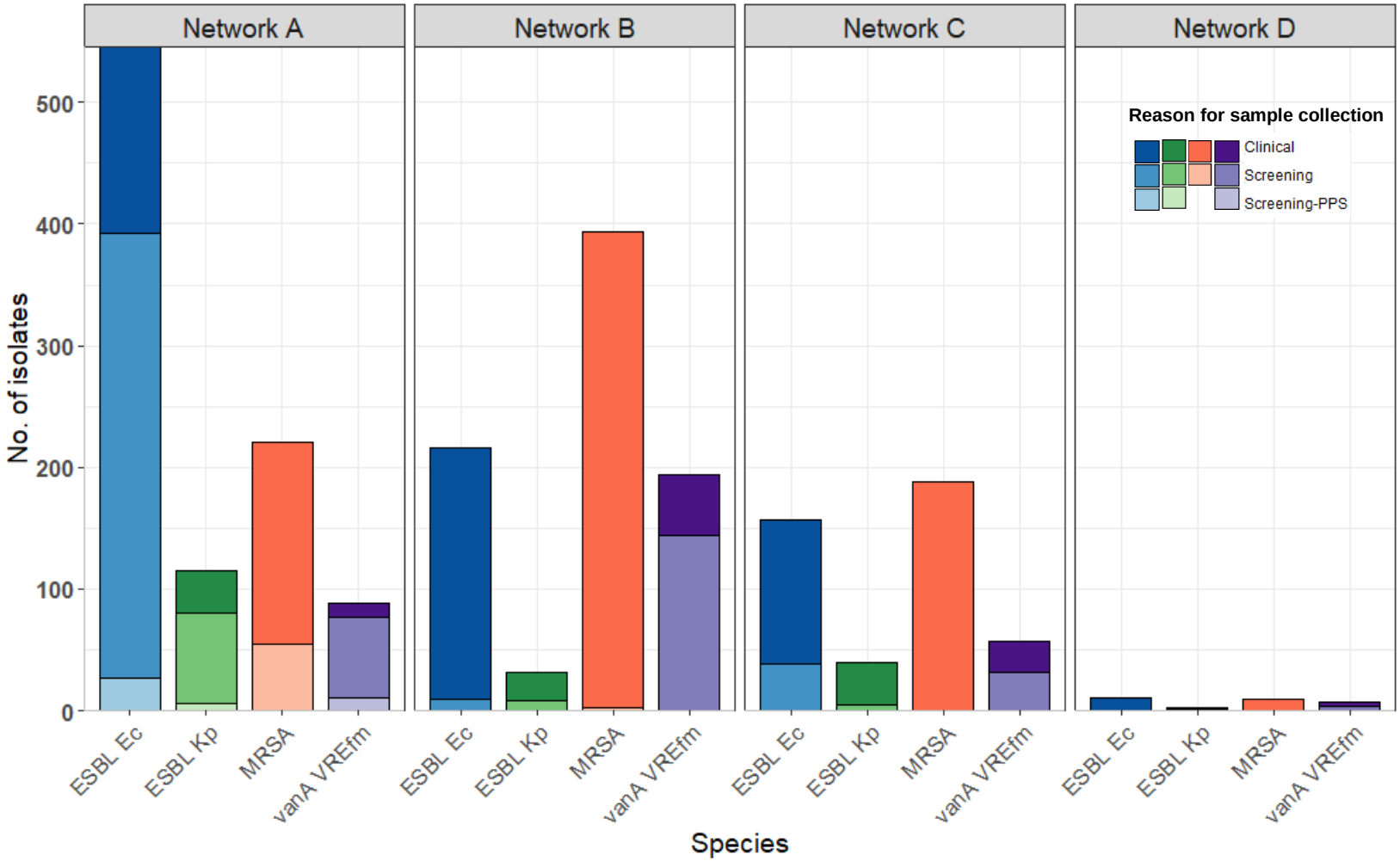
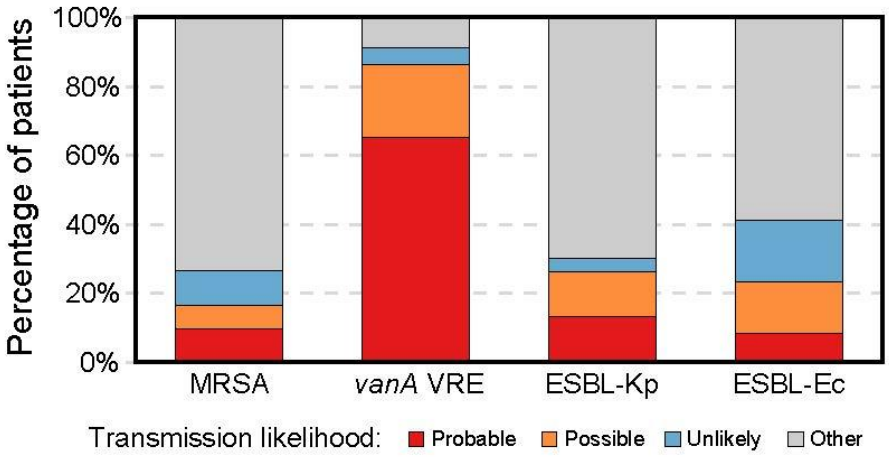


Figure 3. MDRO transmission as assessed by combined genomics and epidemiologic assessment

(A) Likelihood of transmission using combined genomic and epidemiologic data, by species. 'Other' (grey) includes patients with samples with no genomic links to other study samples (singleton STs, or above SNP screening thresholds). (**B)** Putative transmission rates per 100,000 occupied bed days (OBDs) by species, coloured by hospital network. Horizontal dotted line represents the mean transmission rate for each species across all hospital networks.



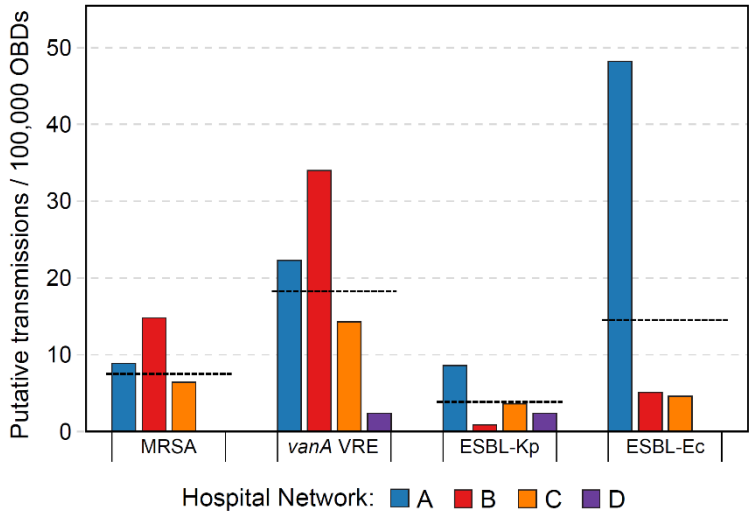
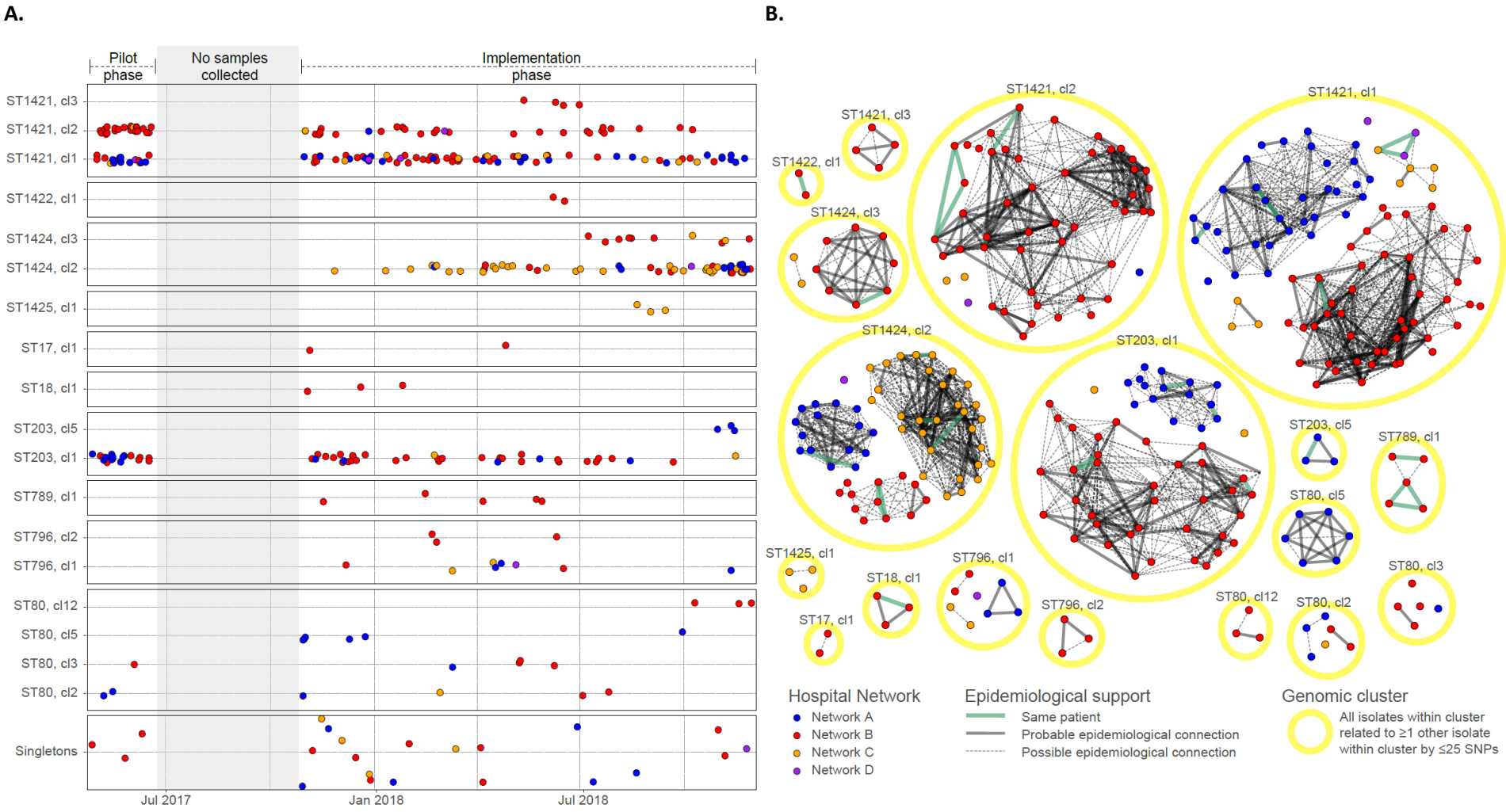
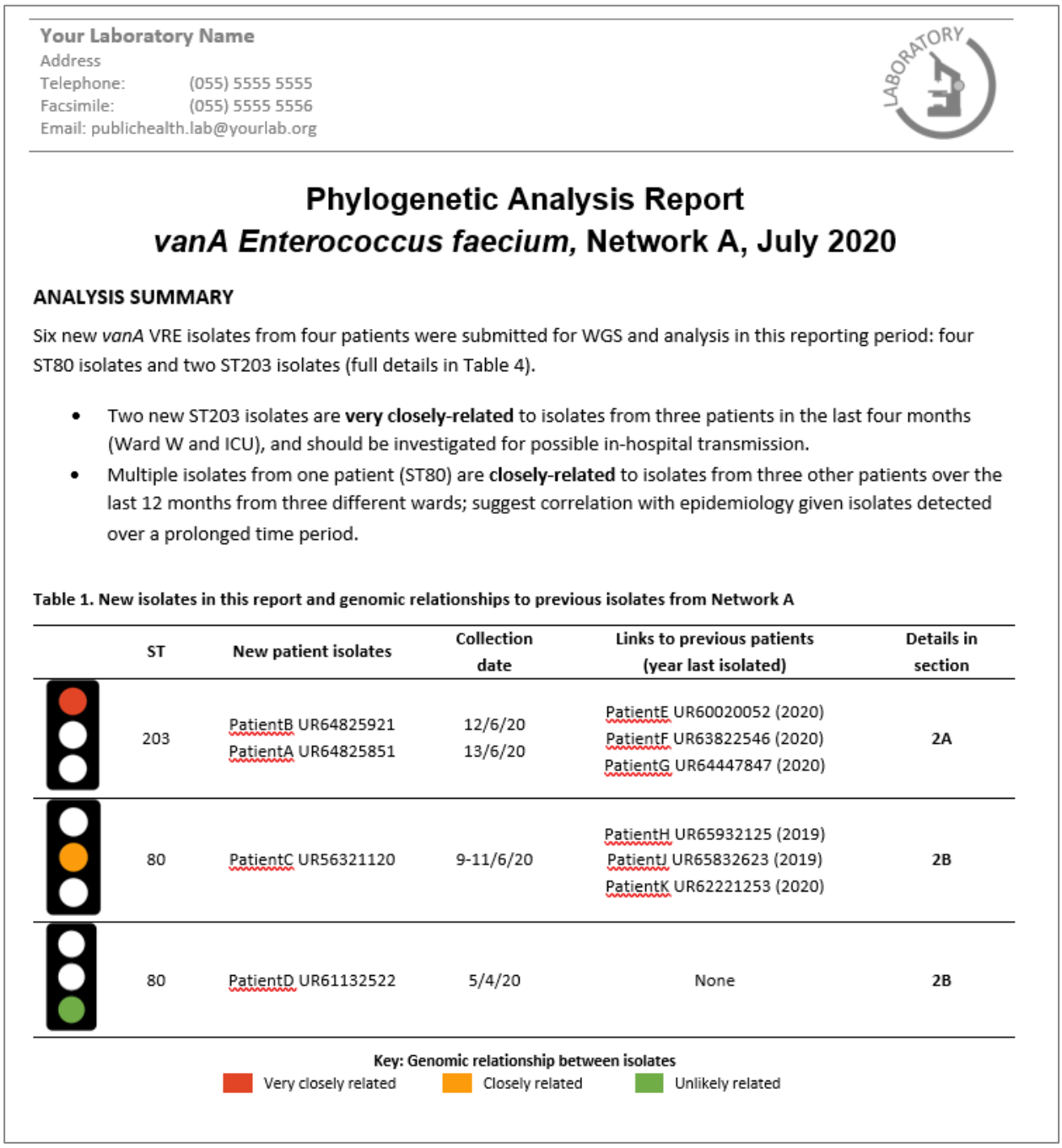


Figure 4. *vanA* VREfm cluster timeline and genomic transmission networks



A. Timeline of *vanA* VREfm cases over the study periods, separated by ST and sub-cluster (cl) along the Y axis, and coloured by hospital network. Timeline is separated into pilot phase (left) and implementation phase (right); shaded timeframe shows the time between phases without sampling. Each point represents one case; only the first sample for each case in the study is included. Points have been separated vertically to allow for visualisation of closely-spaced cases. Sample sequences that did not cluster with any other study cases are shown in the 'Singletons' panel.
B. Network diagram of genomic and epidemiologic links between *vanA* VREfm cases. Each point represents one case; only the first sample for each case in the study is included, and is coloured by hospital network. The yellow circles outline genomic clusters, and lines between points represent epidemiologic links between cases (thick line, probable transmission; dotted line, possible transmission; green line, same patient. Absence of line indicates unlikely transmission by epidemiology, i.e. no known overlap between patients in space or time).
Abbreviations: VREfm, vancomycin-resistant *Enterococcus* faecium; ST, sequence type.

Figure 5. Suggested format for reporting prospective genomics results to an infection control team



Your Laboratory Name
Address
Telephone: (055) 5555 5555
Facsimile: (055) 5555 5556
Email: publichealth.lab@yourlab.org

Phylogenetic Analysis Report
vanA Enterococcus faecium, Network A, July 2020

ANALYSIS SUMMARY

Six new *vanA* VRE isolates from four patients were submitted for WGS and analysis in this reporting period: four ST80 isolates and two ST203 isolates (full details in Table 4).

- Two new ST203 isolates are **very closely-related** to isolates from three patients in the last four months (Ward W and ICU), and should be investigated for possible in-hospital transmission.
- Multiple isolates from one patient (ST80) are **closely-related** to isolates from three other patients over the last 12 months from three different wards; suggest correlation with epidemiology given isolates detected over a prolonged time period.

Table 1. New isolates in this report and genomic relationships to previous isolates from Network A

	ST	New patient isolates	Collection date	Links to previous patients (year last isolated)	Details in section
	203	PatientB UR64825921 PatientA UR64825851	12/6/20 13/6/20	PatientE UR60020052 (2020) PatientF UR63822546 (2020) PatientG UR64447847 (2020)	2A
	80	PatientC UR56321120	9-11/6/20	PatientH UR65932125 (2019) PatientJ UR65832623 (2019) PatientK UR62221253 (2020)	2B
	80	PatientD UR61132522	5/4/20	None	2B

Key: Genomic relationship between isolates
Very closely related Closely related Unlikely related

Excerpt of a suggested report format for intermittently reporting results from prospective genomic surveillance to an infection control team. This report was designed after feedback from focus groups of infection control, infectious diseases and microbiology staff, including some participants with experience in genomics, and some without. All data are fictitious. Full report included in Supplementary Data (Figure S6).

Table 1. Characteristics and infection control practices of hospital sites included in study

Hospital network	Hospital code	Hospital description	No. of inpatient beds[a]	High-risk wards	MDRO screening practices during study period and changes during study	Management of patients colonised or infection with MDRO
A	**A1**	Tertiary referral center, including ICU, solid-organ and bone marrow transplant	560	ICU Hematology/BMT and Oncology Renal Transplant Liver Transplant (Spinal ward and respiratory ward added)	*vanA* VRE and MRGN screening: ICU, on admission and twice weekly; haematology/oncology, renal and liver transplant wards screened on admission and ~~twice~~ weekly ~~for *vanA* VRE and MRGN~~ Additional MRSA screening in ICU (on admission and twice weekly) ~~Quarterly~~ Biannual point-prevalence survey for *vanA* VRE and MRGN MRSA screening before critical surgeries (prosthetic joint, spinal and cardiac) *Change during study:* Added spinal ward and respiratory ward (ventilator support service) to high-risk wards for regular MDRO screening (October 2018)	MRSA: patients only isolated in certain circumstances, e.g. highly exudative wound; single room, disposable apron, gloves *vanA* VRE: single room with own bathroom, full gown, gloves; sometimes cohorted[b] with other *vanA* VRE-colonised patients ESBL-Ec and ESBL-Kp (multidrug-resistant)[c]: single room, own bathroom, disposable apron
	A2	Subacute hospital, aged care and rehabilitation services	150	None	Biannual point-prevalence survey for *vanA* VRE and MRGN	
	A3	Subacute hospital, rehabilitation services	60	None	Biannual point-prevalence survey for *vanA* VRE and MRGN	
B	**B1**	Tertiary referral center, including ICU and solid-organ transplant and	640	ICU Renal Transplant	ICU and renal ward screened for *vanA* VRE and carbapenem-resistant Gram negatives (CRGN) on admission and weekly MRSA screening before cardiac surgery	MRSA: single room, gloves and short-sleeved gown

		specialist pediatric hospital (including neonatal ICU)			*Change during study:* Stopped routine screening of renal ward for VRE (June 2018). Network-wide changes in cleaning practices to from microfibre/steam cleaning to bleach cleaning (September 2018)	*vanA* VRE: single room, own bathroom, occasional cohorting[b] with other *vanA* VRE-colonised patients; gloves and short-sleeved gown. Additional measures if diarrhoea (full gown) ESBL-Ec: no specific management measures ESBL-Kp: single room, own bathroom, gloves and short-sleeve gown
	B2	Tertiary referral center, including ICU, trauma and some aged care & rehabilitation services	573	ICU	ICU patients screened for *vanA* VRE and carbapenem-resistant Gram negatives (CRGN) on admission and weekly *Changes during study:* Cleaning protocol changes as above	
C	**C1**	Tertiary referral center, including ICU, solid-organ and bone marrow transplant	571	ICU Haematology/BMT	ICU and haematology ward screened on admission and weekly for *vanA* VRE and MRGN	MRSA: single room only in certain situations, e.g. highly exudative wound *vanA* VRE: single room, own bathroom, gloves and full gown ESBL-Ec (multidrug-resistant)[c] and ESBL-Kp: single room, own bathroom, gloves and full gown
	C2	Subacute hospital, aged care and rehabilitation services	150	None	None	
D	**D1**	Specialized cancer care center. Located adjacent to Hospital C1 (ICU patients cared for at C1 before transfer back to hospital D1)	96	Hematology	Haematology ward patients screened on admission and weekly for *vanA* VRE and MRGN	MRSA: single room when possible; if open wounds, gown and gloves; if respiratory, mask and gloves *vanA* VRE: single room, own bathroom, gloves and full gown ESBL-Ec and ESBL-Kp: single room, own bathroom, no extra PPE but additional cleaning if incontinent or discharging wounds

ICU, intensive care unit; MRGN, multi-resistant Gram negatives (includes ESBL and carbapenem-resistant phenotypes); BMT, bone marrow transplant (allogeneic); PPE, personal protective equipment. [a]Inpatient beds, excludes day cases, hospital-in-the-home and mental health; [b]Cohorting, practice of patients colonised with the same MDRO sharing a room; [c]Multidrug-resistant ESBL, ESBL isolates with additional resistance aminoglycosides, fluoroquinolones and trimethoprim-sulfamethoxazole (at least two classes).

Table 2. Considerations for implementation of genomics-informed infection control workflow

Component of genomics workflow	Key considerations	Implementation strategy selected for this study
Target population	Consider targeting high-risk populations (e.g. ICU, haematology/oncology, transplant) vs acute inpatients vs all inpatients (acute and subacute)	Included all inpatients due to frequent movement between wards, and unknown burden of disease and transmission in subacute care
Target MDRO/s	Consider AMR spectrum/high-risk organisms; MDROs not covered by existing surveillance programs; organisms where healthcare acquisition more likely than community acquisition; MDROs of high prevalence in local population	Selected *van A* VRE (high prevalence, recently emerged), ESBL-Ec and ESBL-Kp (unclear healthcare contribution) and MRSA (changing epidemiology, unclear healthcare contribution) Note CPE already covered by existing statewide surveillance system; CRPa and CRAb low prevalence
MDRO screening strategy	Consider using existing MDRO screening strategies vs harmonised screening approaches. Could target high-risk populations or all inpatients (e.g. ward-based or hospital-wide point-prevalence surveys)	Elected for pragmatic approach, with each hospital network continuing existing screening strategies, noting differences between networks in analysis
Sample types	All samples vs clinical samples only (latter would underestimate transmission)	All samples included to maximise likelihood of identifying transmission
Clinical data collection	Consider data required to identify likely transmission in hospital, including admissions at other sites; consider ability to accurately identify admissions at different networks by centralised electronic systems	Patient movement data (admissions, wards and beds) critical to identify likely transmissions in conjunction with genomic data. Minimal clinical data collected in implementation phase, as not required to infer transmission.
Sequencing and analysis strategy	(i) Sequencing location: centralised (sequencing laboratory) vs decentralised (hospital laboratories) (ii) Bioinformatics analysis location: centralised vs decentralised, based on resources and expertise (iii) Analysis of combined genomic and epidemiologic data, depending on IPC resources and training, and established systems for electronic data analysis	(i) Sequencing in central laboratory, as sequencing not established at study sites (ii) Centralised analysis at sequencing laboratory as not established at study sites (iii) Genomic and epidemiologic data integrated centrally by study team, presented to IPC teams for further interpretation and action

Bioinformatics analysis	Consider methods for transmission analysis. Examples: core genome SNP analysis, within species or within ST vs cgMLST approach, SNP thresholds, masking recombination Other considerations: frequency of analysis, inclusion of all data vs rolling window (e.g. last 6 or 12 months)	Elected for core genome SNP analysis within ST based on existing accredited bioinformatics pipelines; recombination masked except for MRSA; SNP thresholds selected based on existing data at the time. All samples included in each analysis (approximately monthly), noting computational expense and personnel requirements
Communication of results	Consider audience (IPC/ID/microbiology teams) and understanding of genomics Written reports (frequency) vs discussions in person vs interactive data portal with effective data visualisation	Chose to engage with IPC/ID/microbiology teams using in-person presentations as part of initial implementation, and development of written report format via focus group

MDRO, multidrug-resistant organism; VRE, vancomycin-resistant *Enterococcus*; ESBL-Ec, extended-spectrum beta-lactamase-phenotype *Escherichia coli*; ESBL-Kp, extended-spectrum beta-lactamase-phenotype *Klebsiella pneumoniae*; MRSA, methicillin-resistance *Staphylococcus aureus*; ST, sequence type; cgMLST, core genome multi-locus sequence type; SNP, single nucleotide polymorphisms; IPC, infection prevention and control; ID, infectious diseases.

Table 3. Patient and isolate characteristics

Total no. of isolates included	2275		
Total no. of patients included	1970		
Age (median, IQR)	65 years (47-77yrs)		
Male sex (no., %)	1063 (56.8%)		
Sample collected for clinical purposes (suspected infection)	1428 (62.8%)		
Sample collected for MDRO screening purposes	847 (37.2%)		
Time from admission to isolate collection (median, IQR)	3 days (1-12 days)		
Ward where sample was collected	**All samples (% total samples)**	**Clinical samples (% per ward)**	**Screening samples (% per ward)**
Intensive care unit	373 (16.4%)	117 (31.4%)	256 (68.6%)
Other high-risk ward [a]	434 (19.1%)	121 (27.9%)	313 (72.1%)
General acute ward [b]	1090 (47.9%)	865 (79.4%)	225 (20.6%)
Emergency department	209 (9.2%)	204 (97.6%)	5 (2.4%)
Subacute care ward [c]	169 (7.4%)	121 (71.6%)	48 (28.4%)

MDRO, multidrug-resistant organism; IQR, interquartile range. [a] High-risk wards, includes hematology, oncology, renal ward (including renal transplant), and liver transplant wards.

[b] General acute wards, include all wards not designated as ICU, high-risk or subacute.

[c] Subacute care, includes aged care, rehabilitation, palliative care and spinal wards.

Table 4. Impacts of genomics program for hospital MDRO surveillance

Unexpected MDRO transmission identified, resulting in specific ward interventions	• *vanA* VREfm cluster identified in haematology ward • ESBL-Kp 'super-spreader' patient identified in one ward[a] • Intensive ward cleaning instituted ('super-clean')[b] • Review of patient placement/movements in affected wards • Increased surveillance in affected wards[c]
Outbreak investigation support	• Supported patient-to-patient transmission and implication of shared patient equipment in outbreaks[d] • Identified transmission where epidemiological links were complex • Able to show that new cases in other hospital wards were not linked to an outbreak, demonstrating that the outbreak management had been successful and allowing infection control interventions to be stepped down
Comparison of MDRO transmission rates between networks, resulting in network-wide changes	• Identified higher rate of *vanA* VRE transmission in one network than others, resulting in network-wide change in cleaning practices[e]
Increased awareness of MDRO transmission patterns	• Increased awareness of MRSA and ESBL-Ec transmissions in hospital (previously thought to all be community-acquired) • Confirmed that MDRO identified through screening predicted the MDRO that subsequently caused clinical infection. Useful for advising empiric therapy e.g. in neutropaenic sepsis[38]
Public health impacts	• Large burden of hospital-acquired *vanA* VREfm demonstrated, contributing to new legislative requirement[24] for all *vanA* VREfm cases to be notified to the state health department • Low burden of carbapenem-resistant *Pseudomonas aeruginosa* and *Acinetobacter baumannii* established as baseline before inclusion in state notifiable diseases list

[a] 'Super-spreader' indicating source patient with unusually high number of transmissions to other patients

[b] Ward 'super-cleans' included thorough terminal/discharge-type cleaning for all areas of the ward (including communal areas, corridors, offices, and storerooms) using a bleach-based detergent where possible.

[c] Increased surveillance included additional environmental screening, additional patient screening (added to roster of high-risk wards with weekly patient screening for VREfm and MRGN, and additional focused hand hygiene surveillance

[d] Shared patient equipment includes any equipment that is not restricted to a single patient, e.g. blood glucometers, blood pressure cuffs, weigh chairs

[e] Cleaning practices changed from steam wands and microfibre cloth cleaning to bleach cleaning

Table 5. Areas for further development for implementation of infection control genomics workflow

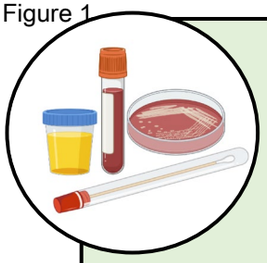
Issue and domain	Description	Potential solutions
Delays in time to results		
a. Hospital laboratory	Delayed time to send isolates to sequencing laboratory • Inadequate staffing • Batch processing of isolates (weekly or longer intervals) • Major disruptions from introduction of new LIS at two laboratories (inability to perform data extractions for study, affecting three sites)	Dedicated staff (supervising microbiologist and technician) More frequent processing of isolates for transit to sequencing lab Ensure robust data extraction capabilities available when implementing new LIS
b. Sequencing laboratory	Minor delays, due to high volume of samples and lower relative priority	Increase in sequencing capacity, redundancy to account for surges in sample volume
c. Bioinformatic analysis	Delays in analysis due to inadequate staffing Increasing complexity of analysis and computational requirements towards end of study due to large isolate numbers (all study samples of same MLST included in each transmission analysis)	Increase dedicated computing resources for complex bioinformatic analyses Work towards partial automation of transmission analysis Develop alternative bioinformatics methods to deal with genomics 'big data' Explore effect of rolling windows (e.g. only including last 6 or 12 months of isolates) on detecting transmission
d. Reporting process	In-person presentations to each network used in this study (to engage and educate staff at study sites) did not allow for results to be conveyed in a timely manner	Issue individual reports for each patient isolate, including MLST result (can be used by Infection Control teams to rapidly exclude or include patients in potential transmissions) New report designed for infection control genomics results (aggregated hospital or network-level data) from focus groups Develop innovative approaches for reporting aggregated results on a hospital/network level e.g. interactive dashboards with advanced data visualisation capabilities
Impaired access to patient epidemiologic data	Each hospital network using different patient information management systems; no readily available way to extract patient data; data had to be collected manually	Develop informatics systems to reliably extract patient movement data, and integrate with results of transmission analyses to inform Infection Control interventions

(including admission, ward and bed data)	(transcription errors, variable formatting requiring manual correction)	
Inability to track patient movements between hospital networks	No readily-available centralised system available for tracking admissions of patients across different hospital networks to identify inter-network MDRO transmission	Engage with health department to advocate for development of centralised admissions database, available in real-time
Perceived cost issues	Sequencing costs perceived to be prohibitive for widespread application currently	Optimise utility of prospective genomics by targeting to local MDROs or locations of concern; demonstrate cost-effectiveness of prospective genomics

LIS, laboratory information system; MLST, multi-locus sequence type (a typing method to subset isolates of the same species into smaller, closely-related groups for targeted transmission analysis); MDRO, multidrug-resistant organism

Figure 1

Click here to access/download;Figure;Fig 1 Infection control genomics workflow.pdf

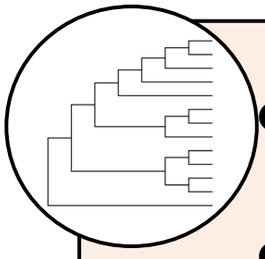


HOSPITAL SITES

1. MDROs isolates from inpatients (clinical or screening samples)
2. Pure subculture sent to central sequencing lab for WGS
3. Demographic and clinical data collected in REDCap

SEQUENCING LABORATORY

1. Single colony picked, subcultured overnight
2. 1-2 colonies selected for WGS
3. DNA extraction, library preparation and QC
4. Illumina short-read sequencing
5. Sequence QC, *de novo* assembly, species identification[a], MLST and AMR gene detection[b]



COMBINED GENOMIC AND EPIDEMIOLOGIC ANALYSIS

6. Phylogenetic analysis performed for all STs with two or more patients[c]
7. Each patient within an ST group compared to each other by pairwise core SNP alignment **(A)**
8. Patient pairs identified as closely related by genomics[d]
9. Patient movement data collected for these patients[e]
10. Overlapping admissions and locations identified **(B)**, and likelihood of transmission inferred[f]

(A)

Pt	A	B	C	D
A		2	7	52
B	2		12	161
C	7	12		83
D	52	161	83	

Closely related by genomics

Not closely related by genomics

(B)

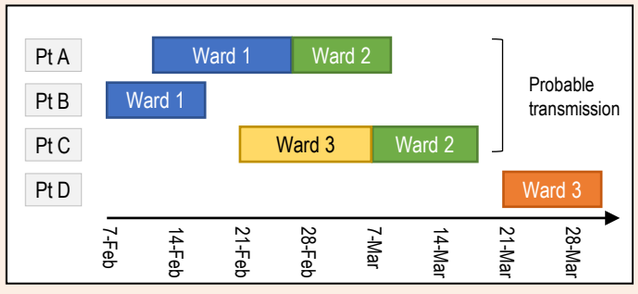


Figure 2

Click here to access/download;Figure;Fig 2 Species by network and screening.pdf

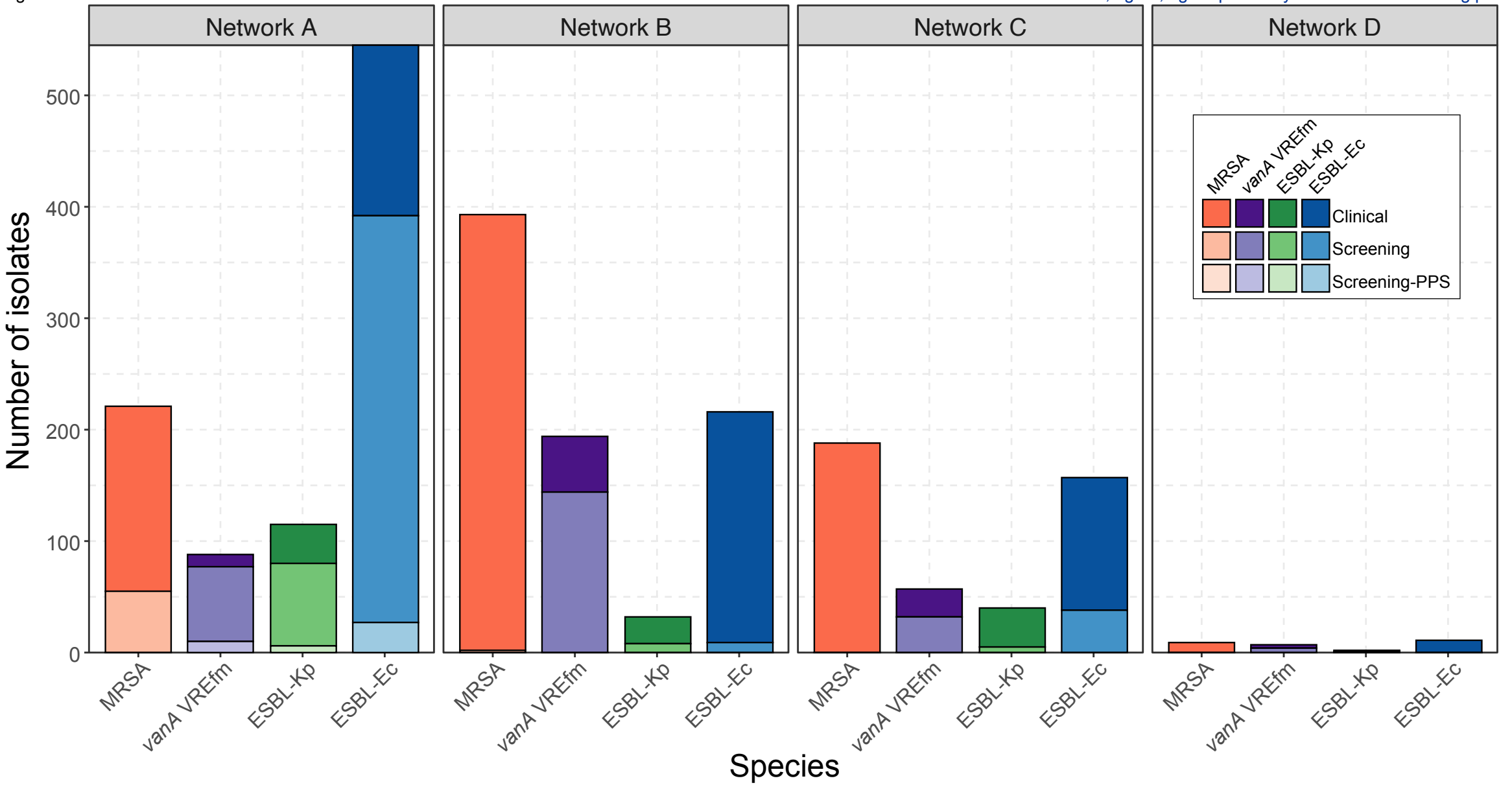
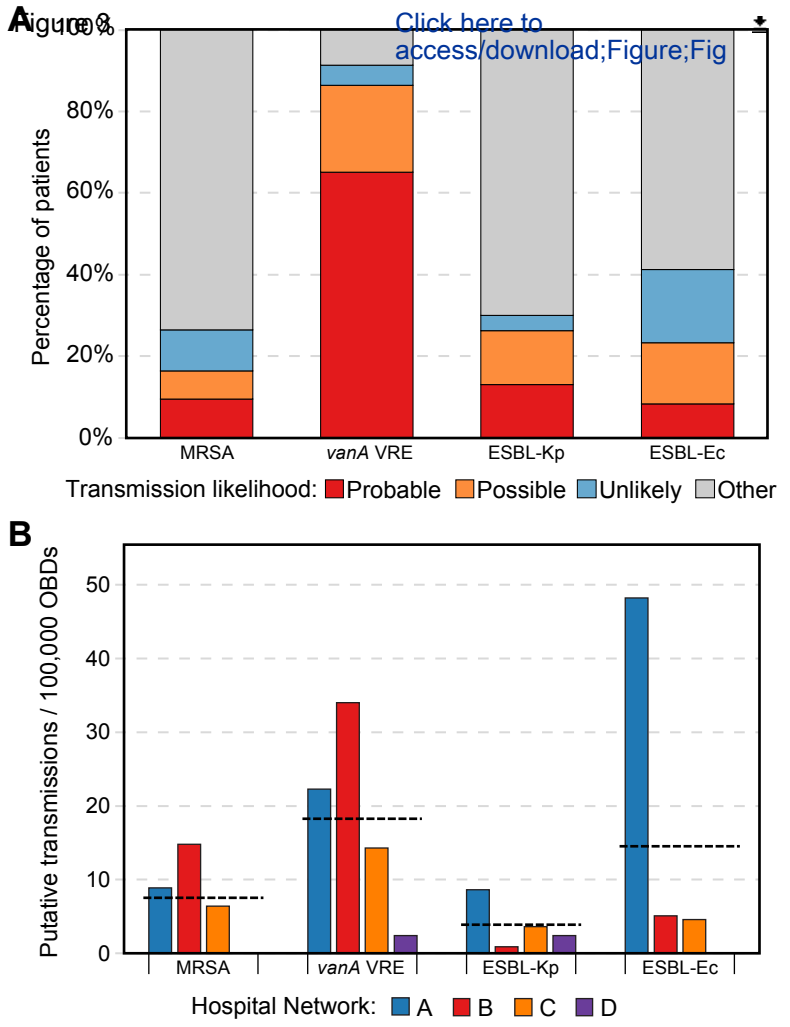
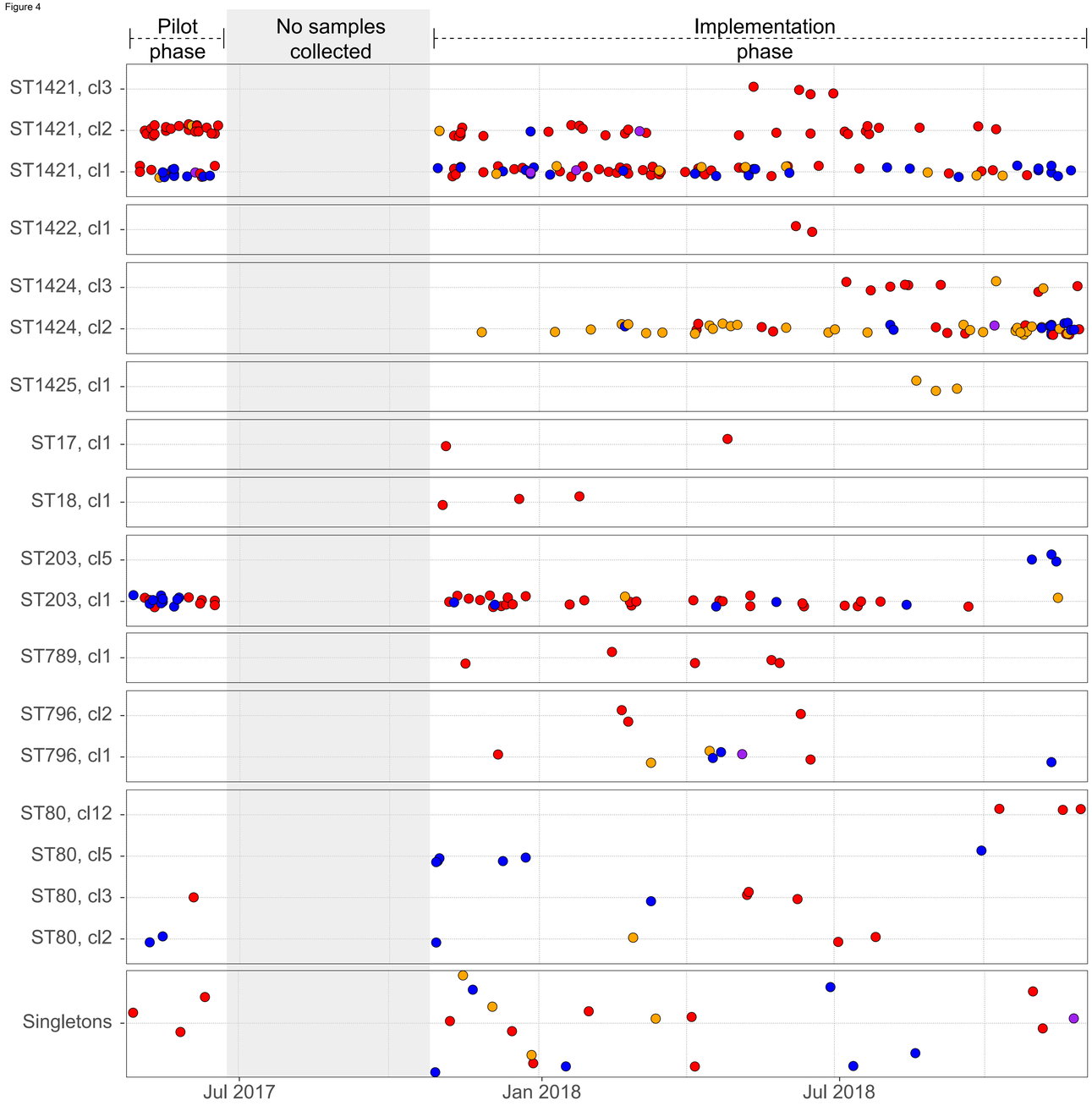


Figure 3





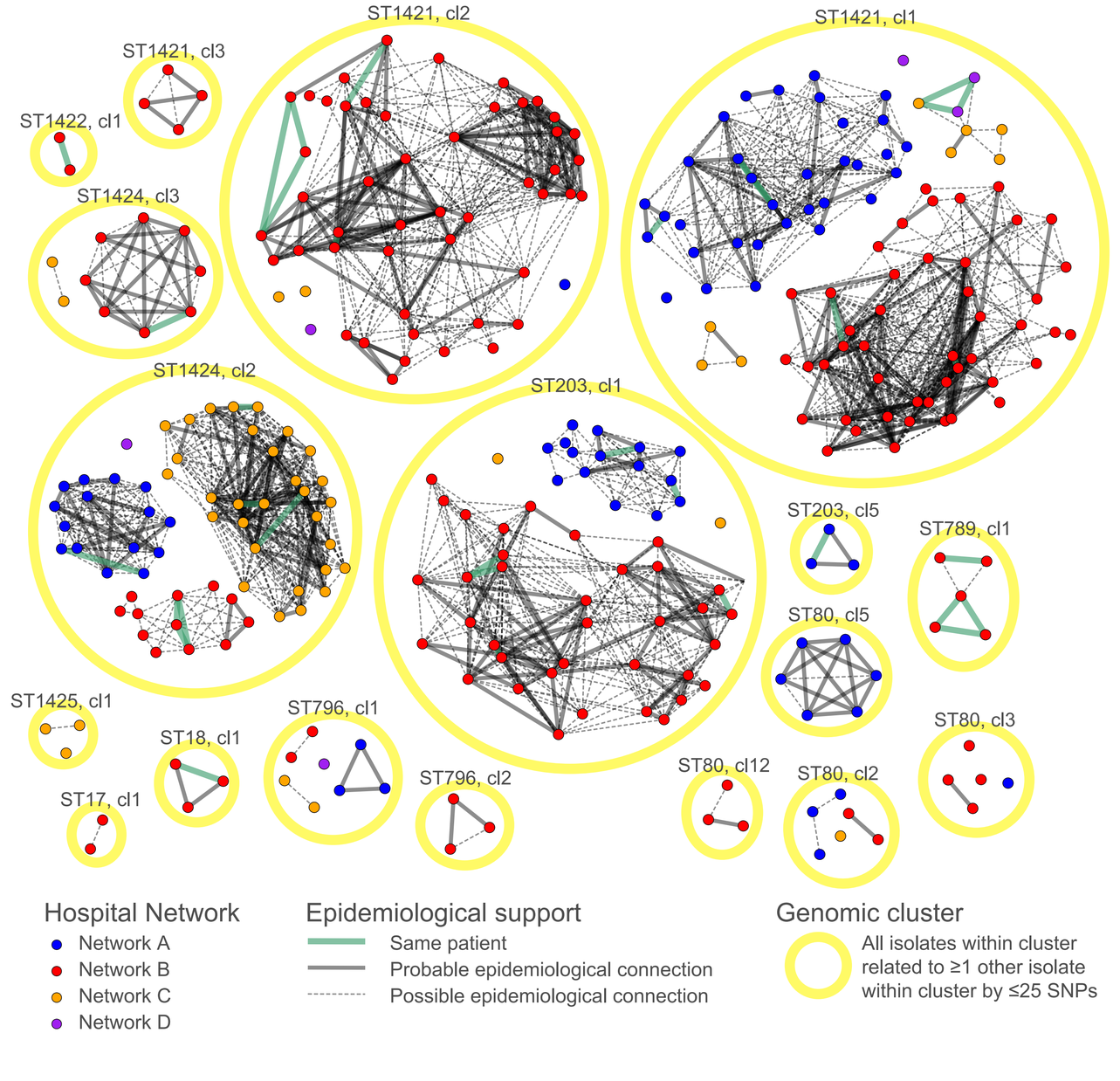


Figure 5

Click here to access/download;Figure;Fig 5 Suggested report format prospective genomics.pdf

Phylogenetic Analysis Report
vanA Enterococcus faecium, Network A, July 2020

ANALYSIS SUMMARY

Six new *vanA* VRE isolates from four patients were submitted for WGS and analysis in this reporting period: four ST80 isolates and two ST203 isolates (full details in Table 4).

- Two new ST203 isolates are **very closely-related** to isolates from three patients in the last four months (Ward W and ICU), and should be investigated for possible in-hospital transmission.
- Multiple isolates from one patient (ST80) are **closely-related** to isolates from three other patients over the last 12 months from three different wards; suggest correlation with epidemiology given isolates detected over a prolonged time period.

Table 1. New isolates in this report and genomic relationships to previous isolates from Network A

	ST	New patient isolates	Collection date	Links to previous patients (year last isolated)	Details in section
	203	PatientB UR64825921 PatientA UR64825851	12/6/20 13/6/20	PatientE UR60020052 (2020) PatientF UR63822546 (2020) PatientG UR64447847 (2020)	**2A**
	80	PatientC UR56321120	9-11/6/20	PatientH UR65932125 (2019) PatientJ UR65832623 (2019) PatientK UR62221253 (2020)	**2B**
	80	PatientD UR61132522	5/4/20	None	**2B**

Key: Genomic relationship between isolates

Very closely related | Closely related | Unlikely related

Reporting checklists (CONSORT, PRISMA, STROBE, ARRIVE, etc)

Title and abstract	Item	STROBE-ID items	STROME-ID items	Responses
Introduction	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found	STROME-ID 1.1: the term molecular epidemiology should be applied to the study in the title or abstract and the keywords when molecular and epidemiological methods contribute substantially to the study	'Genomic epidemiological analysis' included in title
Backgroundrationale	2	Explain the scientific background and rationale for the investigation being reported	STROME-ID 2.1: provide background information about the pathogen populationand the distribution of pathogen strains within the host population at risk	Explained in introduction and methods Population – hospital inpatients colonized or infected with multidrug-resistant organisms (MDROs)
Objectives	3	State specific objectives, including any prespecified hypotheses	STROME-ID 3.1: state the epidemiological objectives of using molecular typing	Stated in introduction – to detect transmission of MDROs in hospitals
Methods				
Study design	4	Present key elements of study design early in the paper	..	
Molecular terminology		..	STROME-ID 4.1: define or cite definitions for key molecular terms used within the study (eg, strain, isolate, and clone)	Definitions for genomic relatedness included in methods (SNP cutoffs)
Molecular markers		..	STROME-ID 4.2: clearly define the molecular markers that were used with a standard nomenclature	Only MLST, methods defined in Supplementary Methods
Infectious diseasecase definition		..	STROME-ID 4.3: clearly state the infectious-disease case definitions	Stated in methods (positive screening or clinical isolate of MDRO, microbiological definitions included in methods and supplementary methods)
Laboratory methodology		..	STROME-ID 4.4: describe sample collection and laboratory methods, including any methods used to minimise and measure cross-contamination, and give the criteria used to interpret strain classification	Detailed in methods Cross-contamination not addressed as clinical samples (i.e. addressed using usual diagnostic microbiology lab methods)
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	STROME-ID 5.1: clearly state the timeframe of the study; consider and appropriately reference the molecular clock of markers if known, and the natural history of the infection	Timeframe clearly stated in methods (detailed dates for 2017-2018)
Participants	6	(a) *Cohort study*—give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—give the eligibility criteria, and the sources and methods of selection of participants (b) *Cohort study*—for matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—for matched studies, give matching criteria and the number of controls per case	STROME-ID 6.1: state the source of participants and clinical specimens, and clearlydescribe sampling frame and strategy	Stated in methods – all MDROs from hospital inpatients at each participating site during the study period
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders,and effect modifiers. Give diagnostic criteria, if applicable		Not directly relevant (observational) Potential confounders for different transmission rates noted in discussion (e.g. different screening practices)
Data sources/ measurement	8*	For each variable of interest give sources of data and details of methodsof assessment (measurement). Describe comparability of assessment methods if there is more than one group		Not relevant
Multiple-strain infections		..	STROME-ID 8.1: describe any methods used to detect multiple-strain infections and measure their effect on the study	Colonisation with multiple MDRO species identified in several cases and reported in Results

			findings	
Bias	9	Describe any efforts to address potential sources of bias	STROME-ID 9.1: describe any efforts made to address discovery or ascertainment bias	Different screening practices and patient populations at each site detailed (Table 1), and discussed with interpretation of results in Discussion
Study size	10	Explain how the study size was arrived at	STROME-ID 10.1: describe any unique restrictions placed on the study sample size	Observational over set time period
Quantitativevariables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	..	SNP cutoffs derived from existing data at the time, discussed in detail in Supplementary methods
Statisticalmethods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) *Cohort study*—if applicable, explain how loss to follow-up was addressed *Case-control study*—if applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—if applicable, describe analytical methods taking account of sampling strategy Describe any sensitivity analyses	STROME-ID 12.1: state how the study took account of the non-independence of sample data, if appropriate STROME-ID 12.2: state how the study dealt with missing data	Median/range/IQR mainly used Chi-square to test differences between groups Incidence differences to compare transmission rates No missing data in this study
Results				
Participants	13*	(a) Report the numbers of individuals at each stage of the study (eg, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage Consider use of a flow diagram	STROME-ID 13.1: Report numbers of participants and samples at each stage of the study, including the number of samples obtained, the number typed, and the number yielding data	Participant numbers (potential and included) described in results with Supplementary Figure S1.
			STROME-ID 13.2: if the study investigates groups of genetically indistinguishable pathogens (molecular clusters), state the sampling fraction, the distribution of cluster sizes, and the study population turnover, if known	Relationships defined primarily by pairwise SNP differences (above or below screening threshold).
Descriptive data	14*	(a) Give characteristics of study participants (eg, demographic, clinical, social) and information on exposures and potential confounders b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study*—summarise follow-up time (eg, average and total amount)	STROME-ID 14.1: give information by strain type if appropriate, with use of standardised nomenclature	Patient (and isolate) characteristics detailed in Table 3. No missing data impacting any results in this study.
Outcome data	15*	*Cohort study*—report numbers of outcome events or summary measures over time *Case-control study*—report numbers in each exposure category, or summary measures of exposure *Cross-sectional study*—report numbers of outcome events or summary measures	..	Outcomes – number of MDRO transmissions, calculated as transmissions per 100,000 occupied bed days, further stratified by correspondence with epidemiologic likelihood of transmission.
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 (b) Report category boundaries when continuous variables were categorised If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	STROME-ID 16.1: consider showing molecular relatedness of strain types by means of a dendrogram or phylogenetic tree	Example of clustering and distribution of clusters between networks over time shown in Figure 4, including correlation with epidemiologic data. Phylogenetic trees have not been included as there are too many to be usefully interpreted (one per species/ST combination).

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	..	Summarised in Discussion (first paragraph)
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	STROME-ID 19.1: consider alternative explanations for findings when transmission chains are being investigated, and report the consistency between molecular and epidemiological evidence	Included in Discussion, including discussion of confounders (different screening practices) and alternative explanations for lack of epidemiologic evidence of transmission for closely-related isolates (e.g. other overlaps in hospital not detected by patient movement data)
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	..	Included in discussion
Generalisability	21	Discuss the generalisability (external validity) of the study results	..	Included in discussion
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	..	Funders had no specific role in the study (provided resources and advice, but the authors had final say over the study design and manuscript)
Ethics	23	..	STROME-ID 23.1: report any ethical considerations with specific implications for infectious-disease molecular epidemiology	No specific ethical implications identified