



Published in final edited form as:

Lancet Infect Dis. 2025 April ; 25(4): e235–e244. doi:10.1016/S1473-3099(24)00729-1.

A Framework towards Implementation of Sequencing for Antimicrobial Resistant and Other Healthcare-associated Pathogens:

Proceedings from the 3rd Geneva Infection Prevention and Control (IPC) Think Tank

Alison Laufer Halpin, PhD^{1,8}, Amy J. Mathers, MD², Prof Timothy R. Walsh, DSc³, Walter Zingg, PD⁴, Prof Iruka N. Okeke, PhD⁵, L. Clifford McDonald, MD¹, Christopher A. Elkins, PhD¹, Prof Stephan Harbarth, MD⁶, Prof Sharon J. Peacock, PhD⁷, Arjun Srinivasan, MD¹, Michael Bell, MD¹, Prof Didier Pittet, MD⁶, Denise Cardo, MD¹ the 3rd Geneva IPC Think Tank

¹Division of Healthcare Quality Promotion, Centers for Disease Control and Prevention, Atlanta, GA 30329, USA

²University of Virginia, Charlottesville, VA 22908-1342, USA

³Ineos Oxford Institute for Antimicrobial Resistance, Department of Zoology, Oxford OX1 3RE, UK

⁴Universitätsspital Zürich, Division of Infectious Diseases and Hospital Hygiene, 8091, Zürich, Switzerland

⁵Department of Pharmaceutical Microbiology, University of Ibadan, Ibadan, 200284, Nigeria

⁶Hôpitaux Universitaires de Genève, Geneva, Switzerland

⁷University of Cambridge School of Clinical Medicine, Cambridge Biomedical Campus, Cambridge CB2 0SP, UK

⁸US Public Health Service, Rockville, MD, USA

Summary

Antimicrobial resistance (AR) continues to be a growing threat globally, specifically in healthcare settings where antimicrobial resistant pathogens cause a substantial proportion of healthcare-associated infections (HAIs). Next generation sequencing (NGS) and the analysis of the data produced therein (i.e., bioinformatics) represent an opportunity to enhance our capacity to address these threats. The 3rd Infection Prevention and Control (IPC) Think Tank brought together experts to identify gaps, propose solutions, and set priorities for the use of NGS for HAI/antimicrobial

Corresponding Author: Alison Laufer Halpin, 1600 Clifton Rd NE, Mail Stop H17-4, Atlanta, GA 30329, USA, ALaufer@cdc.gov. Contributors

Conceptualisation, investigation, methodology, project administration, visualization, writing - original draft, writing - review/editing: ALH. Conceptualisation, investigation, methodology, visualization, writing - original draft, writing - review/editing: AJM. Conceptualisation, investigation, methodology, project administration, writing - original draft, writing - review/editing: WZ. Conceptualisation, investigation, methodology, writing - original draft, writing - review/editing: TRW. Investigation, methodology, visualization, writing - review/editing: SJP. Investigation, writing - original draft, writing - review/editing: INO. Conceptualisation, investigation, methodology, project administration, writing - review/editing: MB, DC, CAE, SH, LCM, DP, AS. Conceptualisation, investigation, methodology, writing - review/editing: MA, JB, ACB, DWE, JJ, TJ, DM, SM, CR, ESS. Investigation, methodology, writing - review/editing: BA, DOA, LB, DRC, VC-CC, AC, NAF, HG, MKH, TJH, HBK, SMK, BP, SP, DP, MHS, JS, RMS, ESS, PS, RS, MCZ.

resistant pathogens. The major deliverable from this meeting was a proposed framework for implementing sequencing of HAI pathogens, specifically those harboring AR mechanisms with the following critical components: wet lab quality, sequence data quality, database and tool selection, bioinformatic analyses, data sharing, and NGS data integration to support public health and IPC actions. Here we detail the framework and discuss in the context for global implementation, specifically in low- and middle- income countries.

I. Introduction

Antimicrobial resistance (AR) remains a critical global health crisis.¹ Healthcare settings serve as amplifiers of both antimicrobial resistant and healthcare-associated infection (HAI) pathogens through transmission within and across healthcare networks, where selective pressure from antimicrobials is high.^{2–4} Over the last decade, next generation sequencing (NGS) and bioinformatic analyses of HAI pathogens and AR have increased, informing surveillance, outbreak responses, and infection prevention and control (IPC) efforts.⁵ Healthcare-associated bacterial pathogens represent a quantitatively staggering and diverse spectrum of genera. Worse, these pathogens can harbor complex genetic elements, including AR, biocide, and other virulence determinants, that can be highly transmissible (e.g., via mobile genetic elements). Although these represent a formidable challenge to the application of NGS for HAI/antimicrobial resistant pathogens, once overcome, NGS will usher in advantages of granular understanding of how those infections spread, their reservoirs, and optimal mitigation measures. These advantages will support global efforts (e.g., national and regional surveillance), local practices (e.g., healthcare facility IPC and healthcare network quality), and clinical diagnostic applications.⁶

As with most healthcare-related challenges, increased complexity requires increased cooperation for interventions to succeed. NGS and bioinformatics innately bring complexity and variability to their applications, and there are many components that need to coalesce for NGS to be useful in IPC and beyond. Given all we have learned to date, it is critical to continue advancing new ideas while applying shared guiding principles and frameworks for IPC and NGS. Thus, after a series of pre-meetings in Summer 2021 with over 60 experts with backgrounds in genomics, bioinformatics, AR, molecular epidemiology, IPC, microbiology, infectious diseases, and healthcare epidemiology, this groups convened at the 3rd Geneva IPC Think Tank, 13-14 September 2021, to address this topic. To ensure we were able to address global implementation of NGS and bioinformatics for HAI pathogens and AR, experts were invited from 16 countries (representing five continents: Asia, Africa, Europe, North America, and South America), including those residing and/or working in low- and middle- income countries (LMICs). The objectives were to explore the role of NGS and bioinformatics in detecting and responding to HAI pathogens and AR, continue to define IPC research and public health priorities related to these topics for the coming years, streamline resources, and advance standards of practice for their use in public health and healthcare. The structure of the 1.5-day meeting comprised multiple sessions where several experts provided a landscape overview or pro/con discussion of a topic to set the stage for moderated discussion. The following serves as a synthesis of perspectives from

the pre-meetings, Think Tank meeting, as well as of literature and advancements in the intervening period since the convening.

II. Applications and Challenges

NGS and bioinformatics have the potential to add new dimensions to our public health capabilities, particularly in healthcare IPC and for local and global IPC and AR detection and tracking. However, in considering the diversity, complexity, and burden of HAI pathogens, the use of genomic science for HAIs/AR pathogens must be fit-for-purpose according to the needs of the users, as opposed to one-size-fits-all.⁷⁻⁹ For the purposes of discussion, applications of NGS for IPC can be categorized based on how sequencing is used: 1) individual diagnostics for patient management, 2) outbreak response^{5,7,9}, and 3) surveillance⁶ and with respect to *when* sequencing is performed: 1) prospective sequencing^{10,11} (i.e., sequence everything or a targeted set of pathogens before other laboratory testing), or 2) retrospective sequencing¹² (figure 1). These are not discrete categories but scalable; different approaches may be selected based on many factors, including the objectives, patient population, resources available, local and regional epidemiology.⁷

Although sequencing in healthcare and for IPC remains at present mainly a research tool, there are increasing examples of its use and benefits in healthcare settings.^{5,13} There might be a day when it can replace the current standard identification and antibiotic susceptibility testing (AST) tools/methods, though AST will require both qualitative and quantitative NGS interpretations (i.e., to predict antibiotic susceptibility phenotype from genotype).¹⁴ Furthermore, for NGS to be useful for IPC and healthcare improvement, consistency and comparability of methods and reporting, broad collaboration across institutions and communities, and appropriate, accessible infrastructure will be required (figure 2).

To date, the application of NGS/bioinformatics has been primarily retrospective or reactive to signals noted through other laboratory testing or astute healthcare staff and for public health rather than diagnostic purposes. Sequence-first strategies in healthcare settings, which rapidly generate and analyze genomes for all HAIs before or without any other laboratory testing towards mitigating (rather than merely identifying) outbreaks, have not been well-studied,^{5,15} in part due to cost, resource-pull, and timeliness. Costs include staff and resources required to stand up and maintain the system. Staffing is particularly challenging in public health laboratories; public health departments are already overextended, often lack bioinformatic-specific job series to aid in recruitment and are challenged to match salaries industry can provide individuals with bioinformatics technical knowledge. Therefore, there are concerns regarding cost-effectiveness and return on investment.^{16,17} Although costs have improved, a blanket sequence-first approach remains unrealistic in most healthcare facilities or country settings, particularly in LMICs.¹⁸ As groups are deciding whether to move to a prospective sequencing approach for outbreak detection^{10,11}, the framework of outbreak detection and response must change as IPC actions become driven by genetic surveillance. (figure 3) However, polyclonal and multi-species HAI outbreaks¹⁹ or promiscuous mobile genetic elements harboring AR genes¹² may be missed if prospective sequencing is not attentive to the entire genetic landscape, not solely strain-based inclusivity/exclusivity

(supplemental page 1). Although currently an outbreak often is detected through traditional epidemiologic or laboratory methods, which are significantly less costly and faster to implement, prospective sequencing may ultimately prevent or better mitigate healthcare outbreaks, detect outbreaks that are not otherwise suspected, and refute suspected outbreaks that are not what they seem.¹⁰ These also generate a wealth of data, which, if able to be meaningfully analyzed in a timely way, can be valuable for continuous surveillance, locally and globally.

Lastly, effective use of NGS requires a workforce of bioinformaticians. Many working in bioinformatics are drawn to this field to innovate - developing novel approaches to interpret complex data and identify emerging or new genetic signatures, AR genes, or other virulence factors. Challenges in bioinformatician recruitment and retention are ubiquitous but even more difficult in LMIC settings, where the pool is smaller and 'brain drain' abounds. Although such innovation is necessary, comparable data, e.g., using standardized tools and curated databases, complying with regulatory requirements, are requisite for many public health and healthcare use cases. Therefore, it remains vital to balance the needs for standardization and innovation, as both are critical for public health and IPC success. A recent example includes the novel implementation of machine learning and artificial intelligence, which will likely be critical as sequencing/bioinformatics is implemented in healthcare settings.¹¹ Supplemental page 2 provides a high-level summary of selected barriers and solutions discussed during the meeting (i.e., not comprehensive)

III. Best Practices for Implementation Framework

Reproducibility/Standards/Comparisons

One of the greatest barriers to leveraging the power of NGS data for HAI/antimicrobial resistant pathogens is the lack of global standards. To maximize epidemiologic impact, we must be able to compare all aspects of the process, from DNA extract quality to instrument performance, raw data generated, genome assembly, and data sharing. Table 1 includes quality metrics for data generation and interrogation, with a focus on metrics and parameters that can be agreed upon to indicate high-quality data and what data and analytic outputs are critical for IPC action.

Wet Lab Processes and Sequencing Data: Given the complexity of sequencing and the diversity of HAI/AR bacterial pathogens, establishing minimum quality metrics is not as straightforward as a single metric. Leveraging multiple metrics to understand sequencing run quality provides a more complete assessment (table 1). Individual laboratories generating sequence data must validate wet lab methods locally and understand the use of controls to monitor the performance of a given sequencing run. Processes to check for sample contamination, quality of reads, cluster density, lack of contamination in the negative control, number of contigs, and expected species based on phenotypic results are all critical to address in such a plan. In addition, tracking run quality longitudinally can help identify subtle declines in instrument performance and wet lab processes. Sequencing data quality checks are critical, particularly for long-read platforms where error rates have been historically higher than short-read platforms, though they are improving.

Finally, understanding the amount of DNA fragmentation that occurs during extraction is also critical for the analysis of mobile genetic elements where genomic context is key to understanding the transmission and/or spread of AR gene-harboring mobile genetic elements. In line with a quality control plan, it follows to utilize multiple components together to assess each run.^{20,21}

Bioinformatic Analyses: Gene identification: Bioinformatic analysis standards may be the most diverse and problematic of the entire sequencing workflow. Even the presence/absence of a given gene, which appears straightforward, remains complex – setting thresholds for percent coverage and identity, comparison database selection, etc., does not lend itself to a binary result (i.e., positive vs negative). Clinical and IPC applications of NGS require robust, validated, version-controlled bioinformatics tools and databases. Further, AR gene databases must be curated, maintained, and should only include clinically significant AR genes, as exemplified by the *mcr-9* gene, subsequently discovered to not actually confer colistin resistance.²² Confirming clinically relevant genes and validating all resistance phenotypes is a slow but necessary process. Re-validation of all tools and/or datasets after any updates, in particular for open-source tools, is not always realistic for real-time implementation (e.g., outbreak response).²³ Despite these challenges, sequence data have an important strength that traditional laboratory tests generally lack. Stored and historical sequence data are readily available for re-analysis as new threats (e.g., AR genes, virulence genes) are identified. These data allow for improved understanding of how and when resistance may have emerged and spread, reconstructing outbreaks, and helping to predict future threats.²⁴

Bioinformatic Analyses: Phylogenetics and Relatedness: For phylogenetic analyses to inform the relatedness of isolates on local, regional, or global levels, method selection is key. Phylogenetics is a challenging area for standardization as the tools and methods are contextual to the questions being asked and the selected dataset and involve differing computational parameters/assumptions. Methods must evolve along with and/or in response to the biology, epidemiology, and computational advances. High quality single nucleotide variant (hqSNV) analyses are appropriate to determine relatedness at the most granular level of a relatively small number of isolates hypothesized to be related (e.g., outbreak/cluster investigations). Larger datasets and/or isolates spanning longer periods (e.g., years) and/or larger geography (e.g., surveillance efforts) may benefit from less granular methods (e.g., k-mer- or Minhash-based approaches, multilocus sequence typing [MLST], core genome or whole genome MLST [cgMLST, wgMLST]) to produce a genomic landscape.²⁵ Although not all HAI bacterial pathogens have established cgMLST or wgMLST schemes, this approach also allows for rapid assessment of large isolate sets to identify potential “hotspots” where subsequent granular (i.e., hqSNV) analyses are needed. This approach may consequently save time and computational needs by bridging the global-local paradigm.

Thresholds (e.g., hqSNV cut-offs) for clonal relatedness are not currently established for most HAI pathogens. Although standards for what defines relatedness at the outbreak-level have been proposed, these will not necessarily apply broadly to all bacterial

species or across methods.^{26–28} Thresholds must consider expected species-level diversity, evolutionary variation in strains (i.e., molecular clock), the analytic approach, reference selection,²⁹ and proportion of the genome analyzed.⁷ Development of thresholds should also consider the possibility of missed links in the transmission pathway (e.g., unidentified reservoirs). Especially in healthcare settings, even prospective sequencing of only clinical isolates will miss environmental, device, and other reservoirs that are often critical steps of the transmission pathway and need to be remediated to halt infections. Finally, some outbreaks occur over days to weeks, while others occur over months to years.^{19,30} There are many phylogenetic standardization challenges and thresholds that need to be sufficiently robust to withstand these pathogen, host, and environmental factors.

Running all sequence data through the same or a centralized workflow and/or bioinformatics pipeline would facilitate comparisons. However, this may not be realistic or simple. For example, individual groups have invested significant time and resources in developing and optimizing pipelines within their group and may be reluctant to adopt a different pipeline, which may not be as optimal for their group, and there may be perceived and/or actual performance differences. Though not a complete solution, collections of high-quality, highly curated specimens support not only the comparison of wet and dry lab approaches across laboratories but also assure accuracy and reproducibility within a single laboratory.^{31,32} (table 1) Designing and distributing a standard set of thoroughly characterized and curated reference isolates for this purpose should be considered. Furthermore, if a centralized workflow is implemented, care should be taken to ensure the data and bioinformatic outputs and interpretations are fed back promptly to the submitters to connect with epidemiologic and clinical data. Building an infrastructure that supports a complete communication loop is critical as a source of data for action and to build and maintain collaborative relationship. This infrastructure and the relationships will enhance our capacity to detect, characterize, and respond to antimicrobial resistance locally and globally.

III. Low- and Middle-Income Countries (LMICs)

While the COVID-19 pandemic accelerated the use of NGS and bioinformatics for public health and IPC, important limitations and challenges remain in LMIC settings, including insufficient/substandard microbiology, resources, supply chains, data handling and sharing, and expertise retention. Despite the potential value of NGS in LMIC settings,³⁴ there are many priorities for clinical diagnosis and IPC, as well as for public health, that may be more pressing requirements, including those with more accessible solutions (e.g., rapid point-of-care diagnostics) to support with limited resources.

There are important examples of LMIC sequence programs domiciled in national public health institutes (e.g., NCID: National Institute for Communicable Diseases in South Africa, RITM: the Research Institute for Tropical Medicine in The Philippines³⁵ and INEI-ANLIS: Argentina's National Institute of Infectious Diseases), or operating with academic or private institutions (often in high- and high-middle income countries (HMICs)) twinned with national public health institutes (e.g., Nigeria,³⁴ ACEGID: African Centre of Excellence for Genomics and infectious Diseases,³⁶ KEMRI: Kenya Medical Research Institute³⁷). However, LMICs with sequencing capacity are too often centered at academic institutions

rather than public health entities. Furthermore, where there are domestic programs, these tend to be at national reference centers and, therefore, somewhat removed from healthcare facilities battling to contain HAIs. Thus, these infrastructures can result in sequencing that is done in isolation from infection prevention control or too focused on a particular phenotype or other characteristic (e.g., convenience sampling, lacking rationale, missing demographic data). NGS implementation efforts should support a public health-oriented design.³⁸ This includes building and maintaining important connections to ensure sequence data and results are promptly shared from the sequencing groups to the submitting partners. This infrastructure will enhance collaboration and strengthen the value of sequence data for public health action.

Studies reporting sequencing data analyses from bacterial strains in LMICs are valuable as surveillance efforts to assess the spread and landscape of major clones and may provide critical data to further our understanding of transmission dynamics, including prevalence in the community of pathogens that are typically healthcare-associated and vice versa. Furthermore, improving the characterization of circulating pathogens in LMICs has a global benefit, considering that pathogens cross borders freely with and without their human or zoonotic hosts.

Although some sequencing platforms (e.g., Oxford Nanopore Technologies) are designed to have a low barrier to entry, establishing the necessary NGS infrastructure in LMICs may require specialized facilities, involve a considerable lag phase, and can be a struggle to sustain. Many public health laboratories in LMICs lack molecular facilities; therefore, there is a need for substantial investment in refurbishing existing laboratory infrastructure to accommodate NGS requirements. Sequencing equipment is highly specialized and expensive and must be installed, maintained, and managed at a high standard. Even when financial resources are available to establish a facility, having ongoing resources for procurement and servicing/maintaining of laboratory equipment is difficult.¹⁸ Because sequencing allows for subtyping of a broad range of bacterial pathogens, shifting from traditional methods to WGS can simplify clinical microbiology supply chains. However, despite what should have been a clear advantage for LMICs in this regard, sequencing, and consumables currently remain challenging to procure in many settings. Sequencing workflows must often be housed in laboratories that are air-conditioned/temperature and humidity-regulated, vibration-free, isolated from other laboratory procedures, with reliable electrical supply (e.g., for freezers and refrigerators) and molecular-grade water source(s), having information technology support, and physical access security (i.e., restricted access). Absent or intermittent internet connection will also constitute a problem in using bioinformatics tools or databases, requiring the development of offline tools and systems administration of local servers. In addition to a lack of continuous internet access and electrical supply lapses, setting up the initial IT infrastructure and data-sharing policies between local hospitals and/or governmental/academic partners remains a hurdle. Outbreaks are often sizeable, spanning lengthy time-periods, and ensuring quality control of sequencing and analysis across long periods may be difficult. However, these scenarios are when sequencing can produce significant impact. NGS training and expertise in LMICs continue to lag; retention of scientists, particularly bioinformaticians, in LMICs is also poor. Accordingly, NGS programs should include train-the-trainer efforts, not only from HMIC

partners but also South-South cooperation and upskilling.^{39,40} Each of these components is a challenging standard to meet individually, let alone in concert, but as there are several exemplars where they have been met, achieving NGS capacity is possible.

Once established, the current cost of NGS remains a significant barrier – most LMICs set a very low ceiling (<10USD) for blood culture, which includes the blood culture bottle, media, and other supplies and reagents for antibiogram profiling. This limits further genomic efforts because a clinical isolate will not be available. Most LMIC hospitals are not fully publicly funded; thus, the costs of diagnostic tests are deferred to the patient (or patient's relatives), who are often uninsured and resource constrained. Thus, any diagnostic test becomes secondary to affording life-saving antibiotics.⁴¹ Discharge against medical advice is common (~15%), with diagnostic/treatment costs being a significant factor. Therefore, finding 40-50USD for isolate sequencing when patients' average wage is at the world poverty level is unrealistic.^{41,42}

Despite these challenges, there can be a valuable return on investment.¹⁷ For example, detection and identification of geographically dispersed outbreaks related to a common source that might otherwise not be linked, such as a contaminated medication, device, or product.^{19,43,44} In all settings, but especially in resource-limited settings, successful application of NGS should be through defining *a priori* use cases where sequencing provides the most utility. For example, identifying multiple clusters with distinct transmission pathways during a general increase of a particular pathogen, like carbapenem-resistant *Acinetobacter baumannii*.⁴⁵ Furthermore, because data can be analyzed remotely, bioinformatics tools and analysis pipelines can be more portable than traditional laboratory infrastructure. Centralized sequencing in a national or regional reference laboratory would also minimize excess costs. Therefore, any investment made in NGS programs will necessarily address specific research questions and identify critical data gaps, furthering our understanding of what IPC interventions are realistically affordable and sustainable. Ultimately, however, in the context of limited resources, priorities should be directed to strengthening IPC measures. Without these capacities, sequencing is not actionable.

IV. Discussion/Conclusion

Next generation sequencing will continue to grow in its utility and accessibility for public health and IPC. Although traditional laboratory approaches remain critical and NGS/bioinformatics does not represent a panacea, we have an important opportunity to intentionally integrate NGS/bioinformatics into routine practice in effective and practical ways. While many healthcare facilities, academic and private laboratories, and public health institutions have established and maintained sequencing capacity, there remain barriers to implementing and increasing the use of HAI/antimicrobial resistant bacterial pathogen NGS for public health and regional IPC. Resource burden, sustainability, and limitations on standardization are often top of the list, underlining the need to ensure sequencing is applied in a flexible, fit-for-purpose manner that matches the objectives to consistently generate comparable high-quality data. Despite this, instances of public health institutions implementing NGS/bioinformatics are growing. For example, the aforementioned RITM in the Philippines^{35,46} and INEI-ANLIS in Argentina. In addition, the Norwegian Institute

of Public Health NOR-WGS program (NOR-WGS)⁴⁷ and the US CDC AR Laboratory Network.⁴⁸ Efforts such as these will facilitate concrete advancement in the implementation of sequencing for public health.

NGS/bioinformatics for IPC must be more than academic; we must create opportunities and systems to use sequence data and bioinformatic analyses to address pragmatic needs for IPC and public health action. Bridging the gap between laboratory, clinical, IPC, and epidemiology domains will enhance implementation. Each of these groups must bring their expertise to the effort, but cross-training and continuing education to ensure all sides can effectively communicate and leverage sequencing in the prevention and containment of HAI/antimicrobial resistant pathogens is critical and remains an ongoing barrier to fully implementing sequencing for IPC.

Other opportunities remain to be fully explored, including how NGS data can drive our fundamental knowledge of healthcare and advance our capability to perform more granular attribution analyses. Unlike other organisms, such as *Mycobacterium tuberculosis*, prediction of phenotype from the genomic data remains a challenge specifically for HAI pathogens but, when used appropriately, may be incredibly informative and support public health as well clinical applications. Finally, in the coming years, discussions around balancing data sharing and privacy, as well as the use of novel tools versus standard methods, will be critical.

These discussions could not provide all the answers, but they aim to foster continued discussion and energy in the effort to implement meaningful NGS/bioinformatics for IPC and combat HAIs and AR. These tools will only continue to advance and evolve. We must find creative approaches to ensure that the need for perfection does not derail current opportunities to leverage NGS/bioinformatics as another tool for preventing infections and saving lives.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

We wish to acknowledge Tcheun-How Borzykowski and Coralie Deleage for their incredible efforts to ensure the success of the 3rd IPC Geneva Think Tank. Further, we wish to thank Dr. Kristy Sands for thoughtful, expert feedback and input on the manuscript.

Disclaimers:

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention/The Agency for Toxic Substances and Disease Registry. Use of trade names is for identification only and does not imply endorsement by the US Centers for Disease Control and Prevention

Declaration of Interests:

ESS reports support from Centers for Disease Control and Prevention (CDC) Cooperative Agreement CK-22-2203 for this manuscript and funding from CDC, ASPR, Massachusetts Institute for Technology, Royalties from: Up to Date (mpox chapter); Consulting fees from BARDA/CARBX/Boston University; Committees: CDC Healthcare and Infection Control and Prevention Advisory Committee (HICPAC), Boston Biosafety Commission, Massachusetts

Infectious Diseases Society, SHEA Public Policy and Government Affairs Committee Chair and High level disinfection and sterilization guidelines committee co-chair, all outside the work reported here. VCCC reports support from the Health and Medical Research Fund (CID-HKU1-16) to his institution and participation on the Society for Healthcare Epidemiologists of America (SHEA) Publication Committee, and Clean Hospitals; serving as Editor or Editorial Board membership for: Journal of Hospital Infection, Infection Prevention in Practice, Infection Control & Hospital Epidemiology, Antimicrobial Stewardship & Healthcare Epidemiology; and Frontiers in Public Health. NAF reports support from UK NIHR (Global Health Professorship). NPG reports support from UK NIHR, CDC, MRC, Gates Foundation to their institution, participation on a Data Safety Monitoring or Advisory Board for ACACIA trial, DREAMM project, 5FC Crypto project; The Federation of Infectious Diseases Societies of Southern Africa (FIDSSA) participation. SH reported consultant fees from Destiny Pharma outside the scope of this manuscript. MKH reports support from CDC; service as President of SHEA Board of Trustees. INO reports no support for this manuscript and support for efforts outside the scope of this manuscript from The Bill and Melinda Gates Foundation, Wellcome Trust, Fleming Fund/Mott McDonald, Grand Challenges Africa Award GCA/DD/rnd3/021, NIHR project #NIHR133307, International Vaccine Institute Award, Royalties for: Genetics: Genes, Genomes and Evolution; Divining without Seeds; and Antimicrobial Resistance in Developing Countries; Thomas Bassir Biomedical Foundation Nigeria Board; International Centre for Antimicrobial Resistance Solutions (ICARS) Technical Advisory Forum. SJP reports service on the Scientific Advisory Board for Next Gen Diagnostics. SP reports no support for this manuscript and reports support outside the scope of this manuscript from the German Center for Infectious Research. MHS reports no support for this manuscript and reports support outside the scope of this manuscript from Department of Veterans Affairs; Centers for Disease Control; Agency for Healthcare Research and Quality. The other authors declared no conflicts of interest.

***3rd Geneva IPC Think Tank members (Alphabetical by surname):**

Mohamed Abbas, MD; Benedetta Allegranzi, MD; Diego O. Andrey, PhD; Prof Judith Breuer, MD; Allison C. Brown, PhD; Liliana Brown, PhD; Prof Douglas R. Call, PhD; Vincent Chi-Chung Cheng, MD; Alejandra Corso, Prof David W Eyre, DPhil; Prof Nicholas A Feasey, PhD; Prof Nelesh P Govender, MMed; Prof Hajo Grundmann, MD; Prof Mary K. Hayden, MD; Thomas-Joerg Hennig, PhD; John Jernigan, MD; Thibaut Jombart, PhD; Prof Hong Bin Kim, MD; Duncan MacCannell, PhD; Prof Surbhi Malhotra-Kumar, PhD; Kalisvar Marimuthu, MBBS; Stefanie McBride, PhD; Benjamin Park, MD; Silke Peter, MD; Diamantis Plachouras, PhD; Chanu Rhee, MD; Prof Matthew H. Samore, MD; Prof Jacques Schrenzel, MD; Erica S. Shenoy, MD; Rachel M. Smith, MD; Evan S. Snitkin, PhD; Padmini Srikantiah, MD; Richard Stanton, PhD; & Marie-Céline Zanella, MD

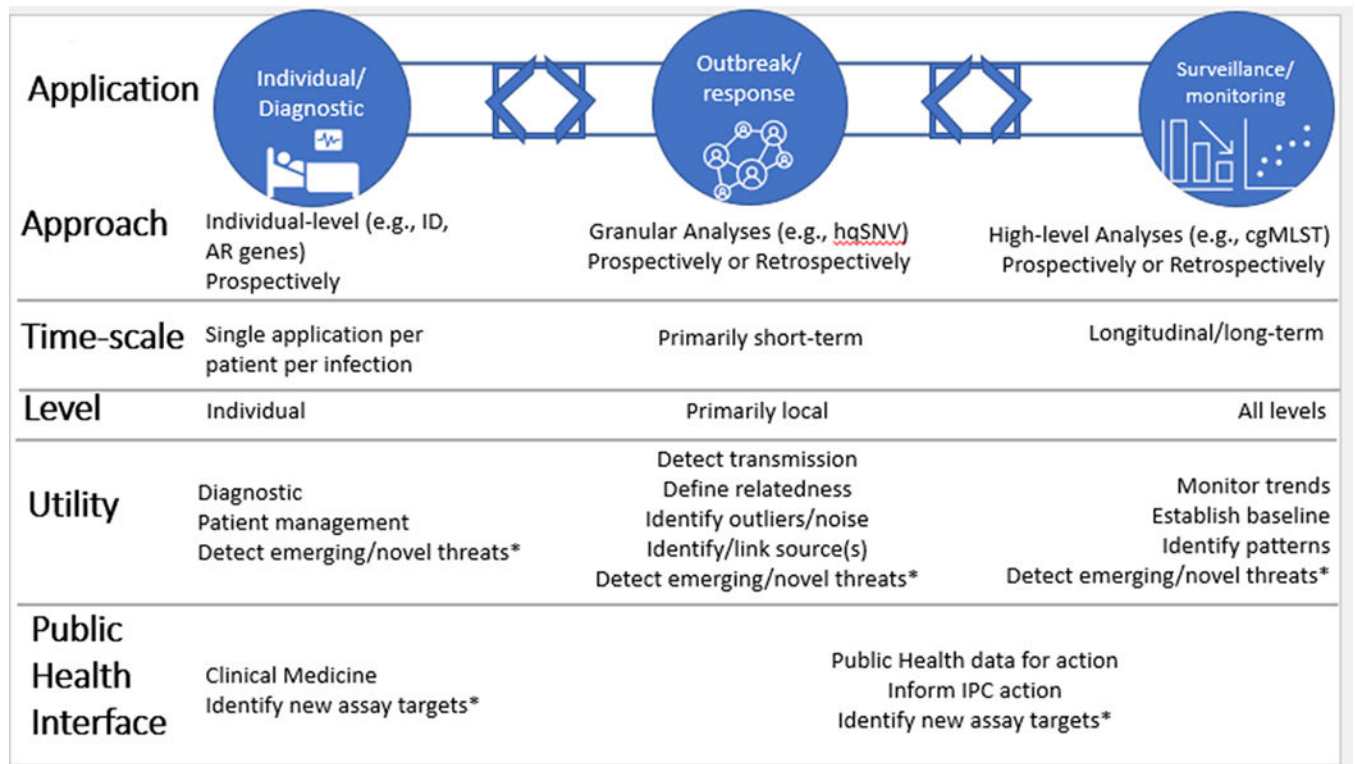
References

1. Murray CJ, Ikuta KS, Sharara F, Swetschinski L, Robles Aguilar G, Gray A, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet*. 2022; 399(10325): 629–55. [PubMed: 35065702]
2. Won SY, Munoz-Price LS, Lolans K, Hota B, Weinstein RA, Hayden MK, et al. Emergence and Rapid Regional Spread of *Klebsiella pneumoniae* Carbapenemase-Producing Enterobacteriaceae. *Clin Infect Dis*. 2011; 53(6):532–40. [PubMed: 21865189]
3. Woodworth KR, Walters MS, Weiner LM, Edwards J, Brown AC, Huang JY, et al. Vital Signs: Containment of Novel Multidrug-Resistant Organisms and Resistance Mechanisms — United States, 2006–2017. *MMWR Morb Mortal Wkly Rep*. 2018; 67(13):396–401. [PubMed: 29621209]
4. U.S. Centers for Disease Control and Prevention. Antibiotic resistance threats in the United States, 2019. Internet. 2019; accessed 2022 Dec 23. Available from: <https://stacks.cdc.gov/view/cdc/82532>.
5. Forde BM, Bergh H, Cuddihy T, Hajkiewicz K, Hurst T, Playford EG, et al. Clinical Implementation of Routine Whole-genome Sequencing for Hospital Infection Control of Multi-drug Resistant Pathogens. *Clin Infect Dis*. 2023;76(3):e1277–84. [PubMed: 36056896]
6. Kekre M, Arevalo SA, Valencia MF, Lagrada ML, Macaranas PKV, Nagaraj G, et al. Integrating Scalable Genome Sequencing Into Microbiology Laboratories for Routine Antimicrobial Resistance Surveillance. *Clin Infect Dis*. 2021;73(Suppl_4):S258–66. [PubMed: 34850836]
7. Halpin AL, McDonald LC, Elkins CA. Framing Bacterial Genomics for Public Health(care). *J Clin Microbiol*. 2021; 59(12):10–1128.

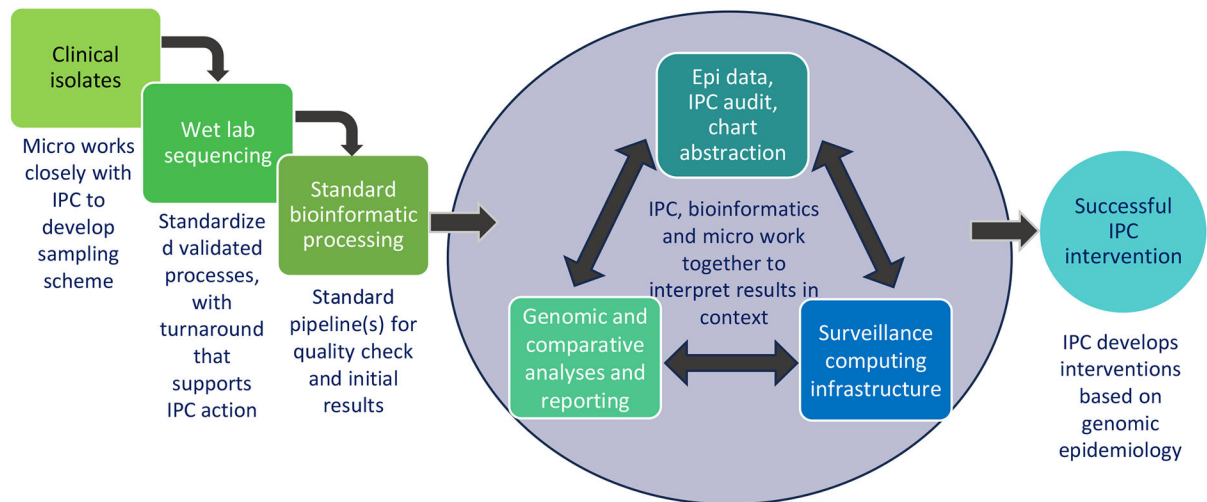
8. Mirande C, Bizine I, Giannetti A, Picot N, van Belkum A. Epidemiological aspects of healthcare-associated infections and microbial genomics. *Eur J Clin Microbiol Infect Dis*. 2018;37(5):823–31. [PubMed: 29340898]
9. Quainoo S, Coolen JPM, van Hijum SAFT, Huynen MA, Melchers WJG, van Schaik W, et al. Whole-Genome Sequencing of Bacterial Pathogens: the Future of Nosocomial Outbreak Analysis. *Clin Microbiol Rev*. 2017;30(4):1015–63. [PubMed: 28855266]
10. 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. [PubMed: 30052172]
11. Sundermann AJ, Chen J, Kumar P, Ayres AM, Cho ST, Ezeonwuka C, et al. Whole-Genome Sequencing Surveillance and Machine Learning of the Electronic Health Record for Enhanced Healthcare Outbreak Detection. *Clin Infect Dis*. 2021;75(3): 476–82
12. De Man TJB, Yaffee AQ, Zhu W, Batra D, Alyanak E, Rowe LA, et al. Multispecies Outbreak of Verona Integron-Encoded Metallo- β -Lactamase-Producing Multidrug-Resistant Bacteria Driven by a Promiscuous Incompatibility Group A/C2 Plasmid. *Clin Infect Dis*. 2021;72(3):414–20. [PubMed: 32255490]
13. Ward DV, Hoss AG, Kolde R, van Aggelen HC, Loving J, Smith SA, et al. Integration of genomic and clinical data augments surveillance of healthcare-acquired infections. *Infect Control Hosp Epidemiol*. 2019;40(6):649–55. [PubMed: 31012399]
14. Mitchell SL, Simner PJ. Next-Generation Sequencing in Clinical Microbiology. *Clin Lab Med*. 2019;39(3):405–18. [PubMed: 31383265]
15. Sherry NL, Gorrie CL, Kwong JC, Higgs C, Stuart RL, Marshall C, et al. Multi-site implementation of whole genome sequencing for hospital infection control: A prospective genomic epidemiological analysis. *Lancet Reg Health West Pac*. 2022;23:100446. [PubMed: 35465046]
16. Kizny Gordon A, Phan HTT, Lipworth SI, Cheong E, Gottlieb T, George S, et al. Genomic dynamics of species and mobile genetic elements in a prolonged blaIMP-4-associated carbapenemase outbreak in an Australian hospital. *J Antimicrob Chemother*. 2020;75(4):873–82. [PubMed: 31960024]
17. Price V, Ngwira LG, Lewis JM, Baker KS, Peacock SJ, Jauneikaite E, et al. A systematic review of economic evaluations of whole-genome sequencing for the surveillance of bacterial pathogens. *Microb Genomics*. 2023;9(2): 000947.
18. Boostrom I, Portal EAR, Spiller OB, Walsh TR, Sands K. Comparing Long-Read Assemblers to Explore the Potential of a Sustainable Low-Cost, Low-Infrastructure Approach to Sequence Antimicrobial Resistant Bacteria With Oxford Nanopore Sequencing. *Front Microbiol*. 2022;13:796465. [PubMed: 35308384]
19. Kracalik I, Kent AG, Villa CH, Gable P, Annambhotla P, McAllister G, et al. Posttransfusion Sepsis Attributable to Bacterial Contamination in Platelet Collection Set Manufacturing Facility, United States. *Emerg Infect Dis*. 2023;29(10):1979. [PubMed: 37561399]
20. Kastanis GJ, Santana-Quintero LV, Sanchez-Leon M, Lomonaco S, Brown EW, Allard MW. In-depth comparative analysis of Illumina® MiSeq run metrics: Development of a wet-lab quality assessment tool. *Mol Ecol Resour*. 2019;19(2):377–87. [PubMed: 30506954]
21. Individualized Quality Control Plan (IQCP) | US Centers for Disease Control and Prevention [Internet]. 2019 [accessed 2024 Jul 11]. Available from: <https://www.cdc.gov/labquality/iqcp.html>
22. Tyson GH, Li C, Hsu CH, Ayers S, Borenstein S, Mukherjee S, et al. The mcr-9 Gene of Salmonella and Escherichia coli Is Not Associated with Colistin Resistance in the United States. *Antimicrob Agents Chemother*. 2020;64(8):10–128.
23. Florensa AF, Kaas RS, Clausen PTL, Aytan-Aktug D, Aarestrup FM. ResFinder - an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. *Microb Genomics*. 2022;8(1):000748.
24. Bennett RJ, Baker KS. Looking Backward to Move Forward: the Utility of Sequencing Historical Bacterial Genomes. *J of Clin Microbiol*. 2019;57(8): e00100–19. [PubMed: 31092597]

25. Masters TL, Toney NC, Ewing TO, McAllister G, Mathis MS, Grigg C, et al. Genomic Epidemiology of Extrapulmonary Nontuberculous Mycobacteria Isolates at Emerging Infections Program Sites – United States, 2019–2020. *J Infect Dis.* 2024;:jiae488.
26. Eyre DW, Cule ML, Wilson DJ, Griffiths D, Vaughan A, O'Connor L, et al. Diverse sources of *C. difficile* infection identified on whole-genome sequencing. *N Engl J Med.* 2013;369(13):1195–205. [PubMed: 24066741]
27. Schürch 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. [PubMed: 29309930]
28. Bush SJ, Foster D, Eyre DW, Clark EL, De Maio N, Shaw LP, et al. Genomic diversity affects the accuracy of bacterial single-nucleotide polymorphism-calling pipelines. *GigaScience.* 2020;9(2):giaa007. [PubMed: 32025702]
29. Valiente-Mullor C, Beamud B, Ansari I, Francés-Cuesta C, García-González N, Mejía L, et al. One is not enough: On the effects of reference genome for the mapping and subsequent analyses of short-reads. *PLoS Comput Biol.* 2021;17(1):e1008678. [PubMed: 33503026]
30. Johnson RC, Deming C, Conlan S, Zellmer CJ, Michelin AV, Lee-Lin S, et al. Investigation of a Cluster of *Sphingomonas koreensis* Infections. *N Engl J Med.* 2018;379(26):2529–39. [PubMed: 30586509]
31. Sichtig H, Minogue T, Yan Y, Stefan C, Hall A, Tallon L, et al. FDA-ARGOS is a database with public quality-controlled reference genomes for diagnostic use and regulatory science. *Nat Commun.* 2019;10(1):3313. [PubMed: 31346170]
32. Lutgring JD, Machado MJ, Benahmed FH, Conville P, Shawar RM, Patel J, et al. FDA-CDC Antimicrobial Resistance Isolate Bank: a Publicly Available Resource to Support Research, Development, and Regulatory Requirements. *J Clin Microbiol.* 2018;56(2):10–128.
33. Petkau A, Mabon P, Sieffert C, Knox NC, Cabral J, Iskander M, et al. SNVPhyl: a single nucleotide variant phylogenomics pipeline for microbial genomic epidemiology. *Microb Genomics.* 2017;3(6): e000116.
34. Okeke IN, Aboderin AO, Egwuenu A, Underwood A, Afolayan AO, Kekre M, et al. Establishing a national reference laboratory for antimicrobial resistance using a whole-genome sequencing framework: Nigeria's experience. *Microbiol Read Engl.* 2022;168(8): 001208.
35. Argimón S, Masim MAL, Gayeta JM, Lagrada ML, Macaranas PKV, Cohen V, et al. Integrating whole-genome sequencing within the National Antimicrobial Resistance Surveillance Program in the Philippines. *Nat Commun.* 2020;11(1):2719. [PubMed: 32483195]
36. Folarin OA, Happi AN, Happi CT. Empowering African genomics for infectious disease control. *Genome Biol.* 2014;15(11):515. [PubMed: 25417758]
37. Kigen C, Muraya A, Kyanya C, Kingwara L, Mmboyi O, Hamm T, et al. Enhancing capacity for national genomics surveillance of antimicrobial resistance in public health laboratories in Kenya. *Microb Genomics.* 2023;9(8):001098.
38. Farzana R, Jones LS, Rahman MA, Andrey DO, Sands K, Portal E, et al. Outbreak of Hypervirulent Multidrug-resistant *Klebsiella variicola* Causing High Mortality in Neonates in Bangladesh. *Clin Infect Dis.* 2019;68(7):1225–7. [PubMed: 30204843]
39. Onywera H, Ondoa P, Nfii F, Ogwel A, Kebede Y, Christoffels A, et al. Boosting pathogen genomics and bioinformatics workforce in Africa. *Lancet Infect Dis.* 2024;24(2):e106–12. [PubMed: 37778362]
40. Kanteh A, Manneh J, Sanyang B, Kujabi MA, Jallow HS, Ndure SL, et al. Simple and structured model to build sequencing capacity in west Africa. *Lancet Glob Health.* 2022;10(9):e1240–1. [PubMed: 35961343]
41. Thomson KM, Dyer C, Liu F, Sands K, Portal E, Carvalho MJ, et al. Effects of antibiotic resistance, drug target attainment, bacterial pathogenicity and virulence, and antibiotic access and affordability on outcomes in neonatal sepsis: an international microbiology and drug evaluation prospective substudy (BARNARDS). *Lancet Infect Dis.* 2021;21(12):1677–88. [PubMed: 34384533]

42. World Bank Blogs [Internet]. [access 2024 Jul 12]. September 2020 global poverty update from the World Bank: New annual poverty estimates using the revised 2011 PPPs. Available from: <https://blogs.worldbank.org/en/opendata/september-2020-global-poverty-update-world-bank-new-annual-poverty-estimates-using-revised>
43. 40.Hasan NA, Epperson LE, Lawsin A, Rodger RR, Perkins KM, Halpin AL, et al. Genomic Analysis of Cardiac Surgery–Associated Mycobacterium chimaera Infections, United States. *Emerg Infect Dis.* 2019;25(3):559–63. [PubMed: 30789131]
44. Grossman MK, Rankin DA, Maloney M, Stanton RA, Gable P, Stevens VA, et al. Extensively Drug-Resistant *Pseudomonas aeruginosa* Outbreak Associated With Artificial Tears. *Clin Infect Dis.* 2024;ciae052.
45. Makke G, Bitar I, Salloum T, Panossian B, Alousi S, Arabaghian H, et al. Whole-Genome-Sequence-Based Characterization of Extensively Drug-Resistant *Acinetobacter baumannii* Hospital Outbreak. *mSphere.* 2020;5(1):e00934–19. [PubMed: 31941816]
46. Enciso AR. RITM talks genomic sequencing with Doherty Institute | Research Institute for Tropical Medicine [Internet]. 2023. [accessed 2024 Jul 12] Available from: <https://ritm.gov.ph/ritm-genomic-sequencing-doherty/>
47. Norwegian Institute of Public Health [Internet]. 2023 [accessed 2024 Jul 12]. NOR-WGS: Improved infrastructure for whole genome sequencing. Available from: <https://www.fhi.no/en/hd/laboratorie-analyser/nor-wgs-improved-infrastructure-for-whole-genome-sequencing/>
48. CDC. Antimicrobial Resistance Laboratory Networks. 2024 [accessed 2024 Jul 12]. Antimicrobial Resistance Laboratory Network. Available from: <https://www.cdc.gov/antimicrobial-resistance-laboratory-networks/php/about/domestic.html>

**Figure 1.**

Sliding scale of NGS Applications. *Indicates public health interface and utility is applicable for more than one purpose. Acronyms: Antimicrobial resistance (AR); Core genome multilocus sequence typing (cgMLST); High quality single nucleotide variant (hqSNV); Identification (ID); Infection prevention and control (IPC)



Considerations	Details	<i>Clostridioides difficile</i> as example
IPC infrastructure	Data integration and management which alerts IPC to unusual patterns	<i>Clostridioides difficile</i> is not unusual, but a cluster of cases would be
Microbiology	Microbiology laboratory prospectively sequencing isolates or, if not possible, saves specific, high priority isolates	<i>C. difficile</i> is not typically cultured in most labs making this an especially challenging organism from a logistics standpoint (i.e., labs would have to identify and purify/grow the isolate in culture)
Strain diversity	If the same strains are found in both community and healthcare, it may be challenging to use WGS data to differentiate between importation and transmission	For some pathogens, the same strains are predominant in the community and healthcare, but this is not often true for <i>C. difficile</i>
Turnaround time	Need results and analysis in an actionable timeframe	Results must be available in real-time to be relevant for IPC decision-making and action
Reliable sequencing	Both wet lab and dry lab processes are standardized to be useful to inform interventions	Lab and computing infrastructure can process isolates reliably, as too many “inconclusive” results can erode trust in the utility/application
IPC Intervention	Defining how genomic results and genomic epidemiology findings would alter/inform IPC action	Often, results are not necessary to initiate intervention(s), such as cleaning or improved hand hygiene

Figure 2. Collaboration and infrastructure are critical for implementing next generation sequencing (NGS). This figure provides a detailed example of how NGS for infection prevention and control (IPC) could be implemented to inform actions and a complementary table of considerations for NGS/IPC applications. Acronyms: Infection prevention and control (IPC)

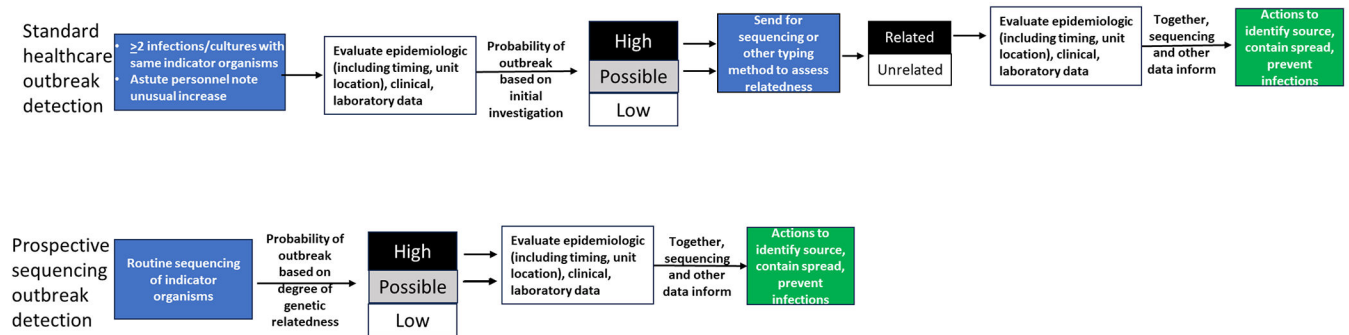


Figure 3.

Standard versus Prospective Sequencing approaches for healthcare outbreak response. Demonstration of how sequencing prospectively may change the framework and timing of outbreak detection and response. Sequencing run times vary based on the instrument selected, throughput, read length, and coverage: Illumina MiSeq instruments run ~17-56 hours, Illumina NextSeq instruments run ~11-30 hours, and Oxford Nanopore Technology instruments run ~6-24 hours. Infection prevention investigations (i.e., evaluating epidemiologic, clinical, and laboratory data) timing will vary based on the facility's resources.

Table 1: Framework for implementing Next Generation Sequencing (NGS) for Public Health and Infection Prevention and Control (IPC).
Best practices for data quality assurance and standardization of reporting healthcare-associated infection (HAI) and antimicrobial resistant bacterial pathogen NGS data and analyses (non-exhaustive list)

	Details and Considerations	Tools, methods, or requirements with example(s)
Component		
Quality: Sequencing instrument performance	- Multiple metrics are important to capture full performance landscape and trends - Metrics to assess sequencing run will vary by platform (e.g., short-read vs long-read)	- Number of reads - Illumina instrument clustering density - Illumina PhiX control - Oxford Nanopore Technology read length distribution - Oxford Nanopore Technology Lambda control
Quality: Sequence data	- Metrics needed to assess both raw reads and assemblies - Metrics to assess initial data generated may vary by platform (e.g., Illumina short-read vs Oxford Nanopore Technology long-read) - Thresholds may differ depending on the objectives (e.g., identifying potential transmission connections vs longitudinal surveillance)	- Quality scores (Q30) - Quality of assembly (e.g., number of contigs using QUAST) - Contamination checks (KRACKEN, GOTTHA) - Identity (genus, species) matches traditional microbiologic methods (if performed) (Average nucleotide identity [ANI]) - Genome completeness (genome length vs expected for species; BUSCO to check for universal single-copy genes) - Depth of coverage (e.g., minimum 30X coverage, 40X for confirming AR gene allele and/or phylogenetic relatedness)
Databases and Tools	- Need reliable databases and tools with standards/thresholds for performance - Methods selection: nucleotide- vs amino acid/protein-level identification	Examples of tools Typing - pubMLST - PlasmidFinder AR genes - AMRFinderPlus - ARG-ANNOT - CARD - ResFinder - Vinulence - Vifdb - VirulenceFinder
Phylogenetic analyses, hqSNV calling, and tracking	- Sequencing platform error rates - Relatedness of a set of isolates on local and/or global levels - Method selection depends on objectives - Reference-based vs <i>de novo</i>/reference-free, each with benefits and trade-offs (Olson paper) - MLST vs cg/wgMLST vs hqSNV - Inappropriate reference genome selection can skew conclusions - Relatedness thresholds are not established for most HAI pathogens - Analyses should account for variables that impact HAI bacterial pathogen diversity and evolution	- Many tools available for phylogenetic analyses with different underlying assumptions, models, parameters²⁴ - pubMLST (web tool) - NCBI Pathogen Detection (web tool) - SNVPhyl³³ - For hqSNV analyses, always include - hqSNV counts “Percent genome included” which indicates proportion of genome scanned for hqSNVs (i.e., high core genome with low hqSNVs supports relatedness)
Sequence Data Sharing	- Timely posting of sequence data passing quality is critical to inform public health actions (e.g., outbreak response, novel and/or emerging threats, contaminated products/devices) to public repositories (e.g., one of the International Nucleotide Sequence Database Collaboration partners - DDBJ, EMBL-EBI, or NCBI) - Posted data provide posterity and carry undefined future utility - Regulatory issues differ by country/region	- Minimum isolate attribute (metadata) sharing proposed for public health utility Organism: Genus and species MLST: (if available), including which scheme(s) (e.g., Oxford, Pasteur) Host: Specimen source (<i>Homo sapiens</i>, animal, environmental, other) Isolation Source: Specimen type (blood, urine, device, surface, etc.) Collection Date: Specimen collection year only Geographic Origin of Sample: Specimen collection country : https://www.insdc.org/

	Details and Considerations	Tools, methods, or requirements with example(s)
	• Always follow regulatory requirements and protect personal identifiable information (PII) • Sufficient anonymization to ensure re-linking to identifiers cannot be accomplished through other pieces of publicly available data	• https://www.insdc-spatiotemporal-metadata-missing-values-update-03-04-2023/news/insdc-spatiotemporal-metadata-missing-values-update-03-04-2023/ • Example: CDC DHQP Seq SuperUmbrella BioProject on NCBI (https://www.ncbi.nlm.nih.gov/bioproject/965738) is a comprised of Umbrella BioProjects and BioProjects that serve as repositories for public health sequence data, standard isolate attributes, etc.
Publication (Repeatability and Reproducibility)	• Sufficiently detailed methods for outside groups to reproduce - consider technical repeatability vs practical repeatability	• Detail tools, version, databases used and which parameters/thresholds/standards • Share scripts and code publicly (e.g., GitHub) • Documented version control of protocols or routine update schedule
Data integration (Data for Action)	• Incorporating additional data is critical (e.g., clinical, epidemiologic, demographic, other laboratory data)	• Center for Genomic Epidemiology (http://www.genomicepidemiology.org/) • Microreact (https://microreact.org/) • iTOL (https://itol.embl.de/) • Pathogenwatch (https://pathogen.watch/)