

The Pathogen Genomics in Public HeAlth Surveillance Evaluation (PG-PHASE) Framework: An Implementation Science Approach to Evaluating Pathogen Whole Genome Sequencing in Public Health

Angeline S Ferdinand (✉ a.ferdinand@unimelb.edu.au)

University of Melbourne School of Population and Global Health <https://orcid.org/0000-0002-4816-5539>

Margaret Kelaher

University of Melbourne School of Population Health: The University of Melbourne School of Population and Global Health

Courtney R Lane

University of Melbourne

Anders Gonçalves da Silva

University of Melbourne

Norelle Sherry

University of Melbourne

Susan A Ballard

University of Melbourne

Patiyan Andersson

Univrsity of Melbourne

Tuyet Hoang

University of Melbourne

Justin T Denholm

Melbourne Health

Marion Easton

Victorian Department of Health and Human Services

Benjamin P Howden

University of Melbourne

Deborah A Williamson

University of Melbourne

Keywords: Pathogen genomics, evaluation, SARS-CoV-2, Listeria, Salmonella

Posted Date: November 13th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-104489/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background Pathogen whole genome sequencing (WGS) is being incorporated into public health surveillance and disease control systems worldwide and has the potential to make significant contributions to infectious disease surveillance, outbreak investigation and infection prevention and control. However, to date, there are limited data regarding: (i) the optimal models for integration of genomic data into epidemiological investigations, and (ii) how to quantify and evaluate public health impacts resulting from genomic epidemiological investigations.

Methods We developed the Pathogen Genomics in Public HeAlth Surveillance Evaluation (PG-PHASE) Framework to guide examination of the use of WGS in public health surveillance and disease control. We illustrate the use of this framework with three pathogens as case studies: *Listeria monocytogenes*, *Mycobacterium tuberculosis* and SARS-CoV-2.

Results The framework utilises an adaptable whole-of-system approach towards understanding how interconnected elements in the public health application of pathogen genomics contribute to public health processes and outcomes. The three phases of the PG-PHASE Framework are designed to support understanding of WGS laboratory processes, analysis, reporting and data sharing, and how genomic data are utilised in public health practice across all stages, from the decision to send an isolate or sample for sequencing to the use of sequence data in public health surveillance, investigation and decision-making. Importantly, the phases can be used separately or in conjunction, depending on the need of the evaluator. Subsequent to conducting evaluation underpinned by the framework, avenues may be developed for strategic investment or interventions to improve utilisation of whole genome sequencing.

Conclusions Comprehensive evaluation is critical to support health departments, public health laboratories and other stakeholders to successfully incorporate microbial genomics into public health practice. The PG-PHASE Framework aims to assist public health laboratories, health departments and authorities who are either considering transitioning to whole genome sequencing or intending to assess the integration of WGS in public health practice, including the capacity to detect and respond to outbreaks and associated costs, challenges and facilitators in the utilisation of microbial genomics and public health impacts.

Contributions To The Literature

- This is a 'first-in-field' framework for the evaluation of whole genome sequencing (WGS) implementation in public health surveillance and outbreak investigation. Although many studies describe the use of WGS data in infectious diseases, there are currently no frameworks to assess its utility.
- The framework utilises a whole-of-system approach to understanding how interconnected elements in the public health application of pathogen genomics contribute to public health processes and outcomes.

- This framework responds to a global increase in pathogen WGS in surveillance systems and a growing need for rigorous evaluation to support effective and efficient integration of pathogen WGS in public health.

Background

Whole genome sequencing (WGS) is being incorporated into public health surveillance and disease control systems worldwide. The United States and the United Kingdom have both implemented national WGS services [1, 2], and in 2016, 26 European countries reported the use of WGS in routine public health practice [3]. The reduction in cost and advancement of portable sequencing technologies has continued to increase accessibility for low-resource settings and low- and middle-income countries [4–6].

Two major applications of pathogen WGS in disease control are: (i) identifying and investigating outbreaks and (ii) genomic surveillance of pathogens of public health importance. WGS has higher discriminatory power than other genotyping methods and can divide clusters identified through other methods into more detailed groupings [7–9], and aid in mapping patterns of microbial relatedness across different human, animal and environmental samples [10]. A key advantage of WGS is the transferability of sequence data, which allows for rapid and interoperable national, international and cross-sectoral data sharing. This enhances the capacity of surveillance systems to quickly detect related cases where epidemiological links are difficult to identify, such as geographically or temporally dispersed clusters caused by wide food dispersion or long incubation periods [11–14].

For laboratories, the direct costs of transitioning to WGS may include: (i) the purchase of equipment and consumables; (ii) changes in laboratory workflows and methods, safety and training needs, and (iii) workforce remodelling [1]. In addition, the generation, analysis, visualisation and storage of genomic data requires appropriate bioinformatics and data management infrastructure. Different analyses and reports may be produced based on the needs of diverse end users, and the reporting needs of end users may change according to circumstance, such as during an outbreak investigation. At present, there are no standardised guidelines for the reporting of genomic data and end users may find it difficult to interpret the data being provided to them. Moreover, there may be end user variability in understanding the uses of genomics for public health, further impacting the utility of pathogen genomic data. Community expectations, including anticipated benefits and ethical concerns around privacy and stigma also have implications for how pathogen genomic data are used and the level of support and advocacy for its incorporation into public health systems [15, 16].

To date, much of the literature on public health pathogen genomics comes from a research perspective. As such, it is often retrospective, illustrating ‘proof of concept’, and does not account for public health practice or outcomes [17–19]. Where there have been studies exploring the implications of public health utilisation of microbial genomics in relation to specific pathogens, the examination of public health practice has been undertaken in isolation from earlier steps such as sequencing and reporting practices, which have a profound impact on the utility of sequence data [20]. Further, the processes of workflow

management, analysis and reporting have often been examined separately from the incorporation of genomic data into public health practice [1, 21]. To our knowledge, there are no whole-of-system evaluations of WGS in public health, nor any formal evaluation or implementation frameworks in this area. Yet, such an evaluation is critical for continued improvement of current genomic implementations, and to add to the evidence base available for other labs and jurisdictions currently embarking on this undertaking. The potential benefits of evaluating WGS implementation are summarised in Table 1. For these reasons, we developed the Pathogen Genomics in Public Health Surveillance Evaluation (PG-PHASE) Framework, underpinned by the principles of implementation science, to examine the impact of WGS on public health. To illustrate application of the PG-PHASE Framework, we use three pathogens as case studies: *Listeria monocytogenes*, *Mycobacterium tuberculosis* and SARS-CoV-2.

Table 1
Potential benefits of evaluating WGS implementation in public health surveillance

Potential benefits
Inform the development of appropriate guidelines and policies around the use of pathogen whole genome sequencing in public health practice
Contribute to the body of evidence available regarding development of pathogen whole genome sequencing capacity in a considered, informed way
Identify needed investment, infrastructure and training to successfully incorporate microbial genomics into public health practice
Provide clarity around the expected outcomes of incorporating pathogen whole genome sequencing in public health and evidence of whether these outcomes are being realised
Identify unexpected outcomes of pathogen whole genome sequencing in public health (either positive or negative) and take appropriate action as necessary
Ensure the public health benefits of whole genome sequencing technology are equitably distributed across the population
Identify and reduce inefficiencies or redundancies in the system
Application of the PG-PHASE Framework against three case studies

A systems-based approach to evaluating WGS in public health and infectious disease control

Hummelbrunner [22] conceptualises systems approaches as a way to understand interventions that consist of elements that form inter-relationships within a whole. Under this model, the utilisation of pathogen genomics in public health qualifies as a complex system, given the interrelatedness of the individual elements, as well as residing within ‘the domain of the “unknowable,”’ where situations are unique and previous experience is not directly applicable [22]. In evaluating such interventions, failure to account for the interdependence present in a complex system risks an under-appreciation of the level of

uncertainty faced by stakeholders and implementers, an underestimation of the influence contextual factors have in the operation of the system, and misattribution of causal effects [22, 23].

The benefits arising from transition to WGS and its subsequent use in public health surveillance are highly context dependent and rely on each part of the system working in conjunction with the others. Ideally, analysis, reporting and translation into practice work together as interconnected parts of an iterative process. Decisions regarding the selection and identification of isolates for sequencing may be informed by previously provided public health information, and WGS approaches need to be rapidly adaptable in order to respond to new and emerging pathogens and AMR. Analysis and reporting are dependent on the various needs of end users, which in turn are also influenced by the wider context. For this reason, a whole-of-system approach is necessary to understand how each step in the process of data generation, analysis and use are interconnected.

Methods

The Pathogen Genomics in Public HeAlth Surveillance Evaluation (PG-PHASE) Framework draws from several approaches that encompass the major components of WGS implementation. The '*Framework for evaluating public health surveillance systems for early detection of outbreaks: Recommendations from the CDC Working Group*'[24] has been used to conceptualise key elements. The PG-PHASE Framework is concerned not only with the *results* from integrating WGS into public health practice, but also the *means* by which outcomes are achieved. The '*Process evaluation of complex interventions: Medical Research Council guidance*' has been used to support an evaluation framework that enables examination of the relationships between assumptions about how the system will work, implementation, system mechanisms, outcomes, and contextual factors [25]. Finally, the structure of the evaluation framework is directly informed by the three phases of the total testing process in clinical laboratories (the pre-analytical, analytical and post-analytical phases) (Fig. 1). The three phases of the evaluation framework and associated outputs, outcomes and indicators are provided in Supplementary Table 1 with data collection methods outlined in Supplementary Table 2.

Phase 1: Pre-analysis and analysis

Given the relatively limited control the laboratory has regarding the pre-analytical phase (largely encompassing specimen selection, collection and transport), the pre-analytical and analytical phases are combined. This first phase focuses primarily on laboratory workflow in transitioning and undertaking WGS for the relevant pathogens (Fig. 1). Initial 'first-level' data analysis is performed on individual sequences at this stage, including quality control, speciation, extraction of basic typing information and identifying the presence or absence of relevant loci/genes such as AMR and virulence genes. Assessment encompasses (i) changes to workflow processes as 'legacy' laboratory methods are retired and pathogen characterisation is transitioned to WGS; (ii) the number of samples processed and analysed in a defined time period (efficiency), and (iii) costings, including costs for staff, instrumentation (including robotics) and reagents/consumables. Decisions regarding which samples are selected for sequencing and sample

processing and analysis times are also assessed as part of this phase of the evaluation. Elements of 'future-proofing' practices to ensure the future useability of isolates and sequence data, including adequate documentation of sample selection and how particular isolates are isolated, cultured, and maintained may also be explored [26]. Laboratory data, such as numbers of samples analysed, and purchasing data, such as reagent costs, may provide additional information. Interviews with laboratory staff may be useful to understand the impacts of workflow and other changes, such as change management issues, training of staff in new processes and the adequacy of transition plans, while interviews with end users may define which isolates were selected for sequencing and why.

Phase 2: Reporting and communication

Phase two examines reporting and communication processes from the dual perspectives of the laboratory and end-users. During this phase, 'second-level' analysis is undertaken, which relates first-level findings to additional metadata and extends analysis to a group of sequences (as distinct from an individual sequence) within the context of the request for analysis. Assessment in this phase encompasses the timeliness and utility of genomic data presented, and the structures in place to determine the most appropriate reporting formats and mechanisms, according to need. Reporting processes may vary and should be assessed according to (i) context and (ii) end user (e.g. epidemiologist, clinician, etc). Evaluation of reporting processes may include the frequency of reports, level and type of information presented, visualisation of information, and adherence to information design principles [21, 27]. Interviews with end users may include examination of intended use of genomic data and questions posed to inform analysis; satisfaction with communication and reporting processes; information retention; perceived utility of the information presented; and level of understanding of information provided. Interviews with bioinformaticians and genomic epidemiologists can provide insight into the various ways sequence data is being used in addition to routine reporting (e.g. sequence data may be contributed to international databases, utilised in phylodynamic modelling or shared between public health agencies). Interviews may explore processes of data sharing, including data governance structures and legal or logistical barriers and facilitators. Through these interviews areas for two-way learning may be identified, where end users are able to strengthen their understanding of the data provided and bioinformaticians and genomic epidemiologists are better informed about what information is expected by end users and how it is intended to be used.

Phase 3: Implementation in public health practice

This phase consists of two parts. The first is a qualitative examination of how WGS data is integrated into public health practice and used to complement or inform epidemiological investigations. Key informant interviews can assess the perspectives of stakeholders regarding the acceptability, usefulness and sustainability of pathogen genomics in public health. Key informants may include medical professionals such as physicians, microbiologists and epidemiologists, representatives of public health departments and hospital infection control units, and other stakeholders (e.g. regulatory agencies; industry; community members). Interviews may explore possible applications and perceived utility or

risks of pathogen genomics in infectious disease prevention and control. Data may be collected on how the transition to WGS has affected public health decision-making, including confidence in making decisions based on the information provided, time to action and the types of decision-making enabled by genomic data. In addition to interviews, data collected may include documented changes to public health policy or the development of guidelines regarding the use of pathogen genomics in public health practice.

The second part of this evaluation phase is a quantitative examination of public health outcomes following implementation of pathogen genomics, with traditional laboratory processes (e.g. legacy typing methods) as a comparator. Relevant indicators may include epidemiological measures such as the number of identified cases, size, duration and number of identified clusters, the proportion of cases linked to clusters and the proportion of cases/clusters traced to a contamination source. Additional indicators may relate to financial costs such as resource allocation for epidemiological investigations and infection control investigations and actions; direct health care costs; financial losses relating to food-borne disease outbreaks (e.g. food recalls); clean-up costs in relation to water and environmental outbreaks; and costs associated with sick leave, both for employers in the form of reduced productivity and for employees as loss of income. Data collection and analyses may vary significantly depending on the epidemiology of the pathogen under consideration, the type of data available and the context (i.e. routine surveillance or outbreak investigation).

Results

The framework can be broadly applied to many pathogens; however, components of the evaluation may vary according to the specific pathogen and / or disease. Using three case studies, we provide examples of how the evaluation framework can be applied to a range of public health pathogens (Table 2 and Supplementary table 3).

Case study 1: Listeria monocytogenes: Pre/post design

Listeriosis is a notifiable disease in many countries, including Australia and the United States [28–31]. WGS has emerged as a valuable tool for investigation of listeriosis outbreaks and is now routinely used for genomic surveillance in several countries [32, 33]. Previous research has shown that the higher discriminatory power of WGS can identify distinct nested clusters within groups of *L. monocytogenes* isolates that were otherwise indistinguishable using other typing methods [34] and utility in identifying contamination sources [35].

When there is availability of data prior to the introduction of WGS, a pre / post design can be used to examine data in defined time periods before (based on previously used typing methods) and after transitioning to WGS. Relevant variables to assess may include the number, size and geographical spread of identified clusters, the percent of isolates linked to a cluster, numbers of isolates/clusters traced to a common source, number of 'solved' isolates/clusters, and the time taken to identify and resolve outbreaks. Further, the evaluation could encompass the use of and contribution to international genomics databases. Changes in costs relating to epidemiological investigation could be also be analysed, given

the utility of WGS in ruling out transmission links. The evaluation could additionally examine trends in food recalls due to *L. monocytogenes*, including the frequency and magnitude of recalls. This approach allows for a comprehensive understanding of how the use of pathogen genomic data has affected the identification and characterisation of clusters across the surveillance system, as well as resulting effects on public health outcomes and use of public health resources.

Case study 2: **Mycobacterium tuberculosis**: Retrospective analysis

Tuberculosis (TB) is the leading cause of death from a single infectious agent, with drug-resistant tuberculosis identified as a global health crisis [36]. The long incubation period and relatively high rates of asymptomatic and undiagnosed infection mean that it can be difficult to confirm transmission through epidemiological links alone. WGS has been shown to provide superior discrimination compared to other typing methods and may be more cost-effective [37, 38]. Sequencing has been used to understand TB transmission dynamics, including identifying super-spreader events [39]. The use of WGS may therefore support more efficient and effective contact tracing, earlier and more appropriate treatment and the initiation of focused public health interventions.

If historical isolates are available at the time of WGS, retrospective sequencing allows for examination of the question: If WGS had been available during this time, what may have been done differently? Here, a retrospective design may allow evaluators to examine practice and outcomes prior to transition to WGS and compare against what would have been expected to occur if WGS had been available at that time.

Retrospective sequence data in combination with epidemiological data can be used to determine which TB may have been identified earlier, allowing for interventions to disrupt further transmissions. Estimates can then be made regarding the number of possible cases averted along with attendant costs to the health care system, including costs relating to epidemiological investigations that may not have been needed. Given the retrospective nature of this approach, it would be important to incorporate a strong understanding of how public health practice is informed by the use of pathogen genomic data, supported by Phases 2 and 3 of the evaluation framework. Undertaking the evaluation in this way utilises a whole-of-system approach to draw links between how TB genomic data is used and eventual public health outcomes, enabling further refinement of pathogen genomics-informed public health practice. Further, this type of evaluation would provide evidence towards assessing the cost-effectiveness of WGS, which may assist in decision-making regarding resource allocation and investment in pathogen genomics.

Case study 3: SARS-CoV-2: Mathematical modelling

From the first instance of genomic sequencing of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [40], WGS has been integrated into the global public health response to COVID-19. In settings that use WGS, evidence of transmission and identification of clusters using genomic data have been highly influential in public health decision-making [13, 41–43]. If WGS was introduced alongside a significant shift in practice, or surveillance of the pathogen under consideration was recently introduced, there may be a lack of appropriate ‘pre’ data. In the case of COVID-19, a reasonable evaluation question

would therefore be: How has WGS contributed (i) to identification of cases and outbreaks, and (ii) to public health interventions? Public health data from similar contexts where WGS has not been used, in conjunction with existing epidemiological data from the setting under consideration, could form the basis of mathematical models to estimate differences in case numbers and characterised outbreaks. While confounders would need to be accounted for in such a model, the aim would not be to arrive at absolute numbers, but rather relative results, such as the proportion of unidentified cases or the probability of being able to detect outbreaks of a certain size with and without the use of genomics [44]. In the absence of suitable comparative data, available epidemiological data may be examined to determine where identification of transmission events would have been uncertain, or where distinct transmission networks may have been merged without the use of genomic data. Interviews with public health authorities may strengthen understanding of how genomic data has been used in public-health decision-making and the links between genomic data and eventual health outcomes. For laboratories that have used or contributed to international sequence databases, this could also be captured as part of the evaluation.

Discussion

Given the large investments that many laboratories and public health agencies have made in applying WGS to public health pathogen genomics [1, 45], it is essential to have a clear framework for evaluating the clinical, public health and economic impact (positive and negative) of WGS implementation and to derive the best value for money. Here, we present a framework for evaluating the use of WGS in public health surveillance and disease control across all stages, from the decision to send an isolate for sequencing and analysis, report development and delivery, to the use of sequence data in public health surveillance, investigation and responses. The introduction of pathogen genomics into surveillance and disease control systems has had considerable impact on how laboratory and public health systems operate, with attendant uncertainty about the best approaches to: (i) facilitate the integration of genomic data into epidemiological investigations; (ii) define the necessary investment for pathogen genomics, and (iii) evaluate public health impacts. The evaluation of public health programs and interventions, including surveillance systems, is crucial to inform appropriate resource allocation, improve system responsiveness, and ensure that program goals are being met [46, 47]. The three phases of the evaluation are designed to support understanding of WGS laboratory processes, analysis, reporting and data sharing, and how genomic data are utilised in public health practice. Importantly, the phases can be used separately or in conjunction, depending on the need of the evaluator.

Conclusions

WGS is now a major part of public health surveillance and the control of infectious diseases, yet there are no defined ways of measuring the overall utility of this approach. The current evaluation framework is presented as a conceptual model, through a whole-of-system lens, to identify barriers and facilitators to the acceptable utilisation of WGS in public health throughout the process. Rigorous evaluation is critical for continued improvement in public health implementation of pathogen genomics and will increase

clarity among stakeholders regarding expected outcomes and whether the aims of the program are being achieved. Ultimately, it is our hope and expectation that utilisation of the evaluation framework will support more effective and efficient integration of pathogen WGS in public health, leading to improved resource allocation, strengthened and more responsive surveillance systems and improved public health outcomes.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

Not applicable

Funding

This work was supported by a National Health and Medical Research Council (NHMRC), Australia Partnership Grant [APP1149991]. BPH is supported by an NHMRC Practitioner Fellowship [APP1105905]. DAW is supported by an NHMRC Investigator Grant [APP1174555].

Authors' contributions

ASF, MK, DAW and BPH conceived of the presented idea. ASF developed the evaluation framework with support from all authors. ASF wrote most of the manuscript with support from DAW and in consultation with MK and BPH. All authors contributed to the conception of the framework and the final manuscript, and all reviewed and approved the submitted version.

Acknowledgements

The authors would like to acknowledge and thank those individuals from the Communicable Diseases Section of the Victorian Department of Human Health and Services and the Microbiological Diagnostic Unit Public Health Laboratory who have contributed their time to support the understanding that underpinned this article.

Competing interests

All authors: no conflicts.

References

1. Grant K, Jenkins C, Arnold C, Green J, Zambon M. Implementing pathogen genomics: A case study: Public Health England, **2018**.
2. Food and Drug Administration. WGS and GenomeTrakr Q&A: FDA, **2018**.
3. Revez J, Espinosa L, Albiger B, et al. Survey on the use of whole-genome sequencing for infectious diseases surveillance: Rapid expansion of European national capacities, 2015–2016. *Front Public Health* **2017**; 5(347).
4. Hendriksen RS, Lukjancenko O, Munk P, et al. Pathogen surveillance in the informal settlement, Kibera, Kenya, using a metagenomics approach. *PLoS One* **2019**; 14(10): 15.
5. Marquez S, Prado-Vivar B, Guadalupe JJ, et al. Genome sequencing of the first SARS-CoV-2 reported from patients with COVID-19 in Ecuador. *medRxiv : the preprint server for health sciences* **2020**: 2020.06.11.20128330.
6. Gardy JL, Loman NJ. Towards a genomics-informed, real-time, global pathogen surveillance system. *Nature Reviews Genetics* **2018**; 19(1): 9-20.
7. Jajou R, de Neeling A, Rasmussen EM, et al. A predominant variable-number tandem-repeat cluster of *Mycobacterium tuberculosis* isolates among asylum seekers in the Netherlands and Denmark, deciphered by whole-genome sequencing. *J Clin Microbiol* **2018**; 56(2): e01100-17.
8. Ford L, Wang QN, Stafford R, et al. Seven *Salmonella Typhimurium* outbreaks in Australia linked by trace-back and whole genome sequencing. *Foodborne Pathogens and Disease* **2018**; 15(5): 285-92.
9. Rumore J, Tscherter L, Kearney A, et al. Evaluation of whole-genome sequencing for outbreak detection of Verotoxigenic *Escherichia coli* O157:H7 from the Canadian perspective. *BMC Genomics* **2018**; 19(1): 870.
10. Pightling AW, Pettengill JB, Luo Y, Baugher JD, Rand H, Strain E. Interpreting whole-genome sequence analyses of foodborne bacteria for regulatory applications and outbreak investigations. *Front Microbiol* **2018**; 9: 1482-.
11. World Health Organization. Whole genome sequencing for foodborne disease surveillance: Landscape paper. Geneva, Switzerland: WHO, **2018**.
12. Inns T, Lane C, Peters T, et al. A multi-country *Salmonella Enteritidis* phage type 14b outbreak associated with eggs from a German producer: 'near real-time' application of whole genome sequencing and food chain investigations, United Kingdom, May to September 2014. *Euro Surveill* **2015**; 20(16).
13. Seemann T, Lane C, Sherry N, et al. Tracking the COVID-19 pandemic in Australia using genomics. *medRxiv* **2020**: 2020.05.12.20099929.
14. Rockett RJ, Arnott A, Lam C, et al. Revealing COVID-19 transmission in Australia by SARS-CoV-2 genome sequencing and agent-based modeling. *Nature Medicine* **2020**.
15. Johnson SB, Parker M. The ethics of sequencing infectious disease pathogens for clinical and public health. *Nature Reviews Genetics* **2019**; 20(6): 313-5.

16. Degeling C, Carter SM, van Oijen AM, et al. Community perspectives on the benefits and risks of technologically enhanced communicable disease surveillance systems: a report on four community juries. *Bmc Medical Ethics* **2020**; 21(1): 14.
17. McDonnell J, Dallman T, Atkin S, et al. Retrospective analysis of whole genome sequencing compared to prospective typing data in further informing the epidemiological investigation of an outbreak of *Shigella sonnei* in the UK. *Epidemiol Infect* **2013**; 141(12): 2568-75.
18. Whaley MJ, Joseph SJ, Retchless AC, et al. Whole genome sequencing for investigations of meningococcal outbreaks in the United States: A retrospective analysis. *Sci Rep* **2018**; 8: 13.
19. Leekitcharoenphon P, Nielsen EM, Kaas RS, Lund O, Aarestrup FM. Evaluation of Whole Genome Sequencing for Outbreak Detection of *Salmonella enterica*. *PLoS One* **2014**; 9(2): 11.
20. Mook P, Gardiner D, Verlander NQ, et al. Operational burden of implementing *Salmonella Enteritidis* and *Typhimurium* cluster detection using whole genome sequencing surveillance data in England: A retrospective assessment. *Epidemiol Infect* **2018**; 146(11): 1452-60.
21. 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: 25.
22. Hummelbrunner R. Systems thinking and evaluation. *Evaluation* **2011**; 17(4): 395-403.
23. Caffrey L, Munro E. A systems approach to policy evaluation. *Evaluation* **2017**; 23: 463-78.
24. Buehler JW, Hopkins RS, Overhage JM, Sosin DM, Tong V. Framework for Evaluating Public Health Surveillance Systems for Early Detection of Outbreaks: Recommendations from the CDC Working Group: Centers for Disease Control and Prevention, **2004**.
25. Moore GF, Audrey S, Barker M, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ : British Medical Journal* **2015**; 350: h1258.
26. Baltrus DA, Cuomo CA, Dennehy JJ, et al. Future-proofing your Microbiology Resource Announcements genome assembly for reproducibility and clarity. *Microbiol Resour Ann* **2019**; 8(36): e00954-19.
27. Valenstein PN. Formatting pathology reports: Applying four design principles to improve communication and patient safety. *Arch Pathol Lab Med* **2008**; 132(1): 84-94.
28. Australian Department of Health. Listeriosis case definition. Available at: https://www1.health.gov.au/internet/main/publishing.nsf/Content/cda-surveil-nndss-casedefs-cd_listera.htm. Accessed 14 Sept 2020.
29. Centres for Disease Control and Prevention. The Listeria initiative. Available at: <https://www.cdc.gov/listeria/surveillance/listeria-initiative.html>. Accessed 14 September 2020.
30. Public Health England. Notifiable diseases and causative organisms: How to report Available at: <https://www.gov.uk/guidance/notifiable-diseases-and-causative-organisms-how-to-report>. Accessed 14 September 2020.
31. Desai AN, Anyoha A, Madoff LC, Lassmann B. Changing epidemiology of *Listeria monocytogenes* outbreaks, sporadic cases, and recalls globally: A review of ProMED reports from 1996 to 2018.

International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases **2019**; 84: 48-53.

32. Department of Health. National Microbial Genomics Framework 2019 – 2022 Australian Government, **2019**.
33. Brown E, Dessai U, McGarry S, Gerner-Smidt P. Use of whole-genome sequencing for food safety and public health in the United States. *Foodborne Pathogens and Disease* **2019**; 16(7): 441-50.
34. Kwong JC, Mercoulia K, Tomita T, et al. Prospective Whole-Genome Sequencing enhances national surveillance of Listeria monocytogenes. *J Clin Microbiol* **2016**; 54(2): 333-42.
35. Thomas J, Govender N, McCarthy KM, et al. Outbreak of listeriosis in South Africa associated with processed meat. *The New England journal of medicine* **2020**; 382(7): 632-43.
36. World Health Organization. Global tuberculosis report 2018. Geneva, Switzerland: WHO, **2018**.
37. Wyllie DH, Davidson JA, Grace Smith E, et al. A quantitative evaluation of MIRU-VNTR typing against whole-genome sequencing for identifying *Mycobacterium tuberculosis* transmission: A Prospective observational cohort study. *EBioMedicine* **2018**; 34: 122-30.
38. Roetzer A, Diel R, Kohl TA, et al. Whole genome sequencing versus traditional genotyping for investigation of a *Mycobacterium tuberculosis* outbreak: A longitudinal molecular epidemiological study. *PLoS Med* **2013**; 10(2): e1001387-e.
39. Lee RS, Proulx J-F, McIntosh F, Behr MA, Hanage WP. Previously undetected super-spreading of *Mycobacterium tuberculosis* revealed by deep sequencing. *Elife* **2020**; 9: e53245.
40. Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease in China. *Nature* **2020**; 579(7798): 265-9.
41. Watson C. How countries are using genomics to help avoid a second coronavirus wave. Available at: <https://www.nature.com/articles/d41586-020-01573-5#ref-CR1>. Accessed 24 Aug 2020.
42. Meredith LW, Hamilton WL, Warne B, et al. Rapid implementation of SARS-CoV-2 sequencing to investigate cases of health-care associated COVID-19: a prospective genomic surveillance study. *The Lancet Infectious Diseases*.
43. Geoghegan JL, Ren X, Storey M, et al. Genomic epidemiology reveals transmission patterns and dynamics of SARS-CoV-2 in Aotearoa New Zealand. *medRxiv* **2020**: 2020.08.05.20168930.
44. Gerber L, González-Suárez M. Population viability analysis: Origins and contributions. *Nature Education Knowledge* **2010**; 3(10).
45. Gwinn M, MacCannell D, Armstrong GL. Next-generation sequencing of infectious pathogens. *JAMA* **2019**; 321(9): 893-4.
46. Centers for Disease Control and Prevention. Framework for program evaluation in public health. Atlanta GA: US Department of Health and Human Services, **1999**.
47. Calba C, Goutard FL, Hoinville L, et al. Surveillance systems evaluation: A systematic review of the existing approaches. *BMC public health* **2015**; 15: 448.

Figures

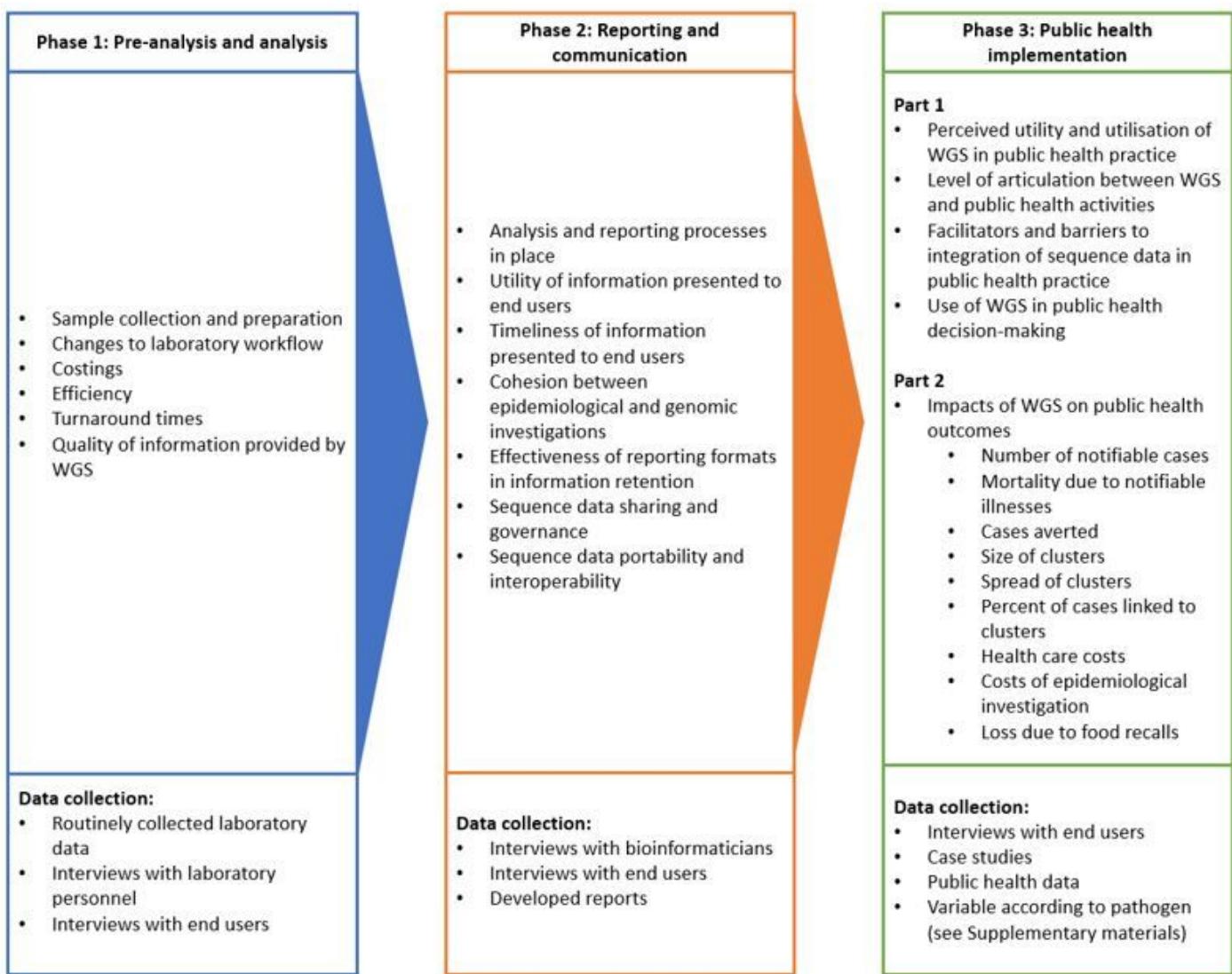


Figure 1

The three phases of the Pathogen Genomics in Public HeAlt Surveillance Evaluation (PG-PHASE) Framework and associated topics and data collection methods.

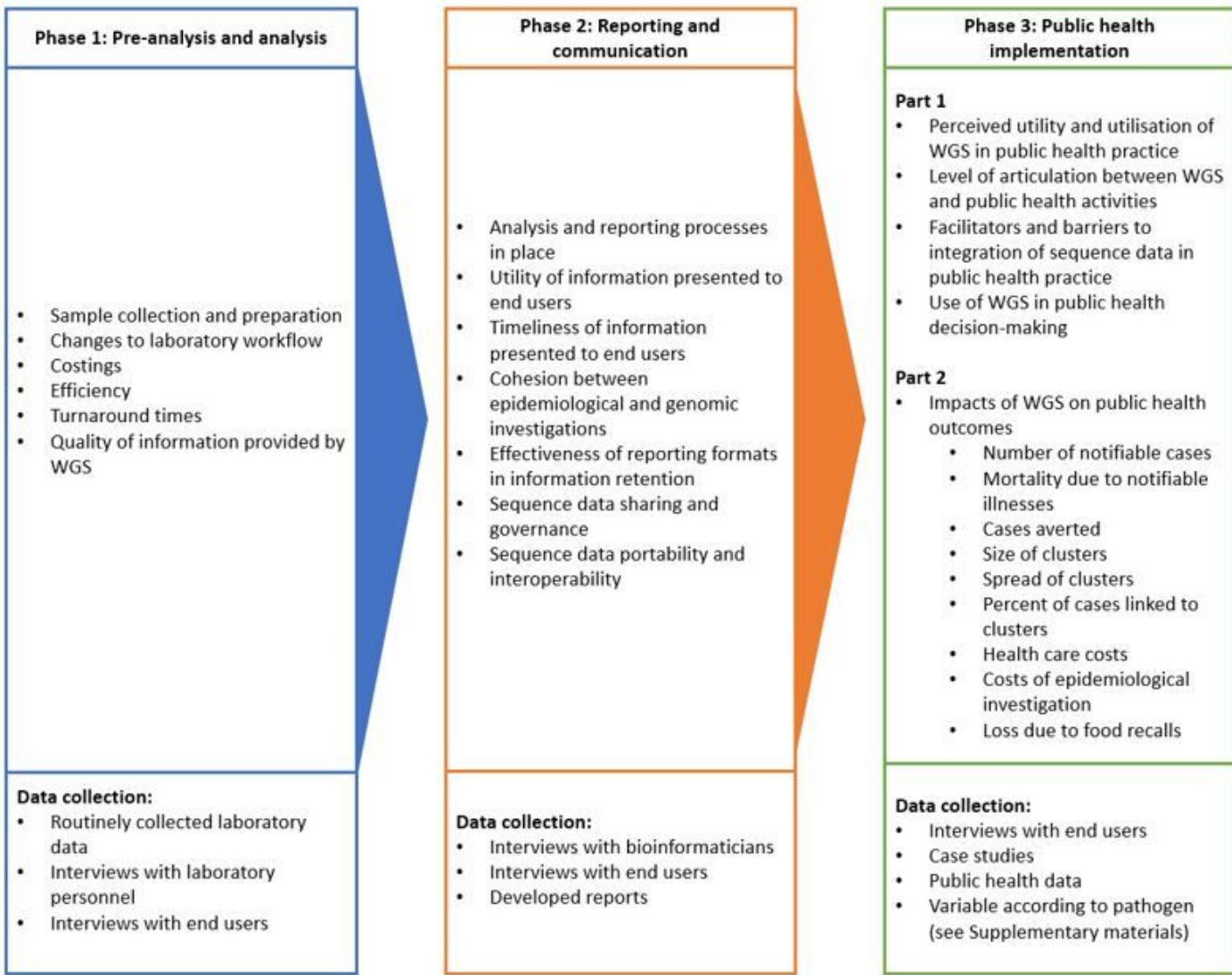


Figure 1

The three phases of the Pathogen Genomics in Public HeAlt Surveillance Evaluation (PG-PHASE) Framework and associated topics and data collection methods.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SuppTables.docx
- SuppTables.docx
- 20201018TIDieRchecklist.docx
- 20201018TIDieRchecklist.docx