



RESEARCH ARTICLE

Ethical challenges in pathogen sequencing: a systematic scoping review [version 1; peer review: 2 approved]

Stephanie Johnson , Michael Parker 

Wellcome Centre for Ethics and Humanities and Ethox Centre, University of Oxford, Oxford, OX3 7LF, UK

V1 First published: 03 Jun 2020, 5:119
<https://doi.org/10.12688/wellcomeopenres.15806.1>
 Latest published: 03 Jun 2020, 5:119
<https://doi.org/10.12688/wellcomeopenres.15806.1>

Abstract

Background: Going forward, the routine implementation of genomic surveillance activities and outbreak investigation is to be expected. We sought to systematically identify the emerging ethical challenges; and to systematically assess the gaps in ethical frameworks or thinking and identify where further work is needed to solve practical challenges.

Methods: We systematically searched indexed academic literature from PubMed, Google Scholar, and Web of Science from 2000 to April 2019 for peer-reviewed articles that substantively engaged in discussion of ethical issues in the use of pathogen genome sequencing technologies for diagnostic, surveillance and outbreak investigation.

Results: 28 articles were identified; nine United States, five United Kingdom, five The Netherlands, three Canada, two Switzerland, one Australia, two South Africa, and one Italy. Eight articles were specifically about the use of sequencing in HIV. Eleven were not specific to a particular disease. Results were organized into four themes: tensions between public and private interests; difficulties with translation from research to clinical and public health practice; the importance of community trust and support; equity and global partnerships; and the importance of context.

Conclusion: While pathogen sequencing has the potential to be transformative for public health, there are a number of key ethical issues that must be addressed, particularly around the conditions of use for pathogen sequence data. Ethical standards should be informed by public values, and further empirical work investigating stakeholders' views are required. Development in the field should also be under-pinned by a strong commitment to values of justice, in particular global health equity.

Keywords

pathogen genomics, ethics, global health, infectious disease, NGS

Open Peer Review

Reviewer Status  

Invited Reviewers

1

2

version 1



03 Jun 2020



report



report

1. **Gregory L. Armstrong** , Centers for Disease Control and Prevention, Atlanta, USA
2. **Yann Joly** , McGill University, Montreal, Canada

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Stephanie Johnson (stephanie.johnson@bdi.ox.ac.uk)

Author roles: **Johnson S:** Conceptualization, Formal Analysis, Writing – Original Draft Preparation; **Parker M:** Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by the Wellcome Trust through a Wellcome Centre Grant [203132] and a Wellcome Strategic Award [096527].

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2020 Johnson S and Parker M. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Johnson S and Parker M. **Ethical challenges in pathogen sequencing: a systematic scoping review [version 1; peer review: 2 approved]** Wellcome Open Research 2020, 5:119 <https://doi.org/10.12688/wellcomeopenres.15806.1>

First published: 03 Jun 2020, 5:119 <https://doi.org/10.12688/wellcomeopenres.15806.1>

Introduction

Genetic information derived from pathogens is an increasingly essential input for infectious disease control, public health and research¹. Although routine sequencing of pathogens was, until recently, unthinkable, the Centers for Disease Control (CDC), Food and Drug Administration (FDA), state, and global public health laboratories now routinely sequence more than 200 foodborne bacterial isolates a day and more than 6,000 influenza virus genomes a year^{2,3}. In the United Kingdom, Public Health England now engages in routine clinical genomic diagnostics and drug sensitivity testing for *Mycobacterium tuberculosis*⁴. In the research setting, phylogenetic analysis (the study of evolutionary relationships among pathogens) is being used to track and understand factors associated with the spread of infections such as HIV⁵ and to monitor the global spread of drug-resistant infections⁶. Mobile genomic sequencing technology is also being applied to disease outbreak investigation, most publicly in the case of the Ebola outbreak in West Africa⁷⁻⁹. Going forward, the routine implementation of genomic surveillance activities and outbreak investigation is to be expected.

While the technical developments of sequencing technology are being implemented at a rapid pace, the non-technical aspects of implementing this technology are still being broadly discussed between the different stakeholders involved¹. The successful implementation of this rapidly developing technology will, for example, require sharing of samples and metadata, interdisciplinary global collaborative partnerships, and will need to offer useful evidence for public health decision-making. Importantly, the successful and appropriate response to these challenges will also require the systematic identification, analysis and addressing of a number of complex ethical, legal and social issues. A number of factors will contribute to the types of ethical issues that arise in different instances. These are likely to include characteristics of the disease, the environmental, political and geographical context, existing laws and policies, public attitudes, and cultural differences¹⁰. In the work reported in this paper, taking these and other ethical issues as our focus, we sought to systematically examine the available literature to: identify the emerging ethical challenges and proposed solutions; and to systematically assess the gaps in ethical frameworks or thinking and identify where further work is needed to solve practical challenges.

Methods

Scoping reviews seek to identify literature relevant to a research objective and may include a variety of research formats and conceptual literature¹¹⁻¹³. This study sought to review published literature on ethical aspects of pathogen sequencing. Inclusion criteria for the study encompassed a broad range of article types, including empirical studies, news articles, opinion pieces, features, editorials, reports of practice, and theoretical articles.

Search strategy

We systematically searched indexed academic literature from PubMed, Google Scholar, and Web of Science from 2000 to April 2019 for peer-reviewed articles that substantively engaged in discussion of ethical issues in the use of pathogen genome

sequencing technologies for diagnostic, surveillance and outbreak investigation. The search was then updated in January 2020. The initial search strategies were developed through an iterative process and used a combination of controlled vocabulary (MeSH terms) and free text words. An example MEDLINE search strategy is provided in [Table 1](#). Reference lists of included articles were searched for relevant articles and further database searches were conducted using the names of researchers commonly publishing in this field. Finally, we also reviewed relevant international research and clinical practice guidance for relevant guidelines e.g. website of the World Health Organization.

Selection criteria

We sought to maximize the literature included in the review by reviewing guidelines, frameworks, commentaries and original research reviews related to pathogen sequencing. We also included studies on molecular typing where enough accuracy could be included to include transmission tracking, as this was thought to provide useful insights into the ethical challenges pathogen sequencing technologies may pose. We excluded studies considering genomics outside of infectious disease or focusing on host response studies as these were not deemed relevant or specific enough to the topic under investigation.

Selection of studies

Duplicates were removed. SJ undertook title and abstract screening to remove obviously irrelevant studies, borderline cases were discussed with MP and a decision reached by consensus. Data was then abstracted by SJ and cross-checked for accuracy by MP. Names of study authors, institutions, journals of publication and results were non-blinded.

Analysis

SJ initially inductively analysed all studies, recording the aims and main findings in Microsoft Excel 2016 (16.0.4699.1000), and developing descriptive codes to chart the broad themes describing the literature base. Similar findings were then grouped according to topic area and a preliminary list of themes was developed in collaboration with MP. Both authors engaged in iterative discussions about organization of findings, after which the final themes were decided. SJ subsequently re-examined each study and extracted data using a standard format (design, results and recommendations).

Results

The search produced 466 articles after duplicates were removed. All articles were initially screened by title and abstract and thirty-nine full text articles were assessed for eligibility. Twenty eight articles were included in the final analysis; three ethical guidelines or frameworks¹⁴⁻¹⁷; seven empirical research studies^{1,18-23}; eight reviews of the ethics^{2,17-21}; and ten publications that contained a section on the ethical aspects of pathogen sequencing²⁴⁻³³. The literature largely originated from the US and other high-income countries (HICs, country determined by lead author institution): nine from the United States (US), five from the United Kingdom (UK), five from The Netherlands, three from Canada, two from Switzerland, one from Australia, and one from Italy. Only two publications, by the same author,

Table 1. Medline search strategy.

Genomics	Infectious disease	Ethics
1. genomics [Mesh] *Genomics/es [Ethics]	11 Bacterial Infections/ge, tm [Genetics, Transmission] bacterial	23. Ethics Committees, Clinical/ or Ethics, Research/ or Ethics/ or Ethics, Medical/ or Ethics Committees, Research/ or ethic*.mp.
2. sequencing [tiab]	12. infectious disease	24. Ethic*
3. Sequence Analysis, DNA [MESH]	13. viral	25. bioethics
4. DNA, Bacterial [MESH]	14. pathogen.mp	26. Genomics/es [Ethics]
5. genom*	15. infectious disease.mp. or Communicable Diseases/	27. policy
6. sequenc*	16. outbreak.mp. or Disease Outbreaks/	28. Guide line
7. Phylogeny [MESH]	17. Population Surveillance/ or Public Health Surveillance/ or surveillance.mp.	29. framework
8. DNA, Viral/ or viral.mp. or Genome, Viral/ DNA, Bacterial/ or Sequence Analysis, DNA/ or sequenc*.mp.	18. epidemic*.mp.	30. Socio-cultural
9. DNA, Bacterial/ or Sequence Analysis, DNA/ or sequenc*.mp. or Phylogeny/	19. epidemic*.mp.	30. 22 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29
10. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9	20. Epidemiology/ or epidemiology. mp. or Molecular Epidemiology/	31. 10 and 22 and 31
	21. public health	
	22. 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21	

originated from the global south (South Africa). Eight articles were specifically about the use of sequencing in HIV^{14,15,17,20,22,28,34,35}. Eleven were not specific to a particular disease. Table 2 presents a summary of studies.

Results were organized into four themes: Tensions between public and private interests; difficulties with translation from research to clinical and public health practice; the importance of community trust and support; equity and global partnerships; and the importance of context.

Theme 1: Tensions between private and public interests

When considering the implications of collecting, using and sharing pathogen genome sequence data, the interests and rights of individuals were universally acknowledged. In particular, the literature pointed to the importance of considering and/or protecting individual rights to autonomy^{19,21,32,36}, and to privacy^{18,21,22,28}. Many pointed out that the potential of sequencing techniques to detect the origin and routes of transmission of an outbreak may result in negative consequences for the individuals involved²⁵. Consequences may include stigmatization, penalties, economic risks, problems with interpersonal relationships (e.g. inadvertent disclosure of infidelity), emotional distress and the capacity for discrimination^{19,21,22,24,25,27,30,32,37–39}. There was also concern that sequencing could lead to serious legal consequences, particularly with regards to the criminalization of HIV transmission^{14,15,17,22,24,28,34,35}. It was also acknowledged that

individuals may have an interest in avoiding the use of information about them for purposes they do not endorse, such as to support anti-gay sentiment or as part of a criminal investigation^{20,24}. Some were concerned about forced testing either of certain groups such as gay men²⁸ or healthcare workers^{19,21,38–40}. There was acknowledgment of individual professional interests of researchers and practitioners to ownership and use of data^{1,16,32}.

Sequencing was also seen to carry risks for communities and groups. Many authors noted that certain groups can be placed at risk through characterization as high risk or likely to transmit virus (HIV), including geographically defined groups, sexual or gender minorities, or those defined by ethnicity, nationality, or migration status¹⁴. Similarly, data regarding transmission patterns of multidrug-resistant tuberculosis (TB) could be used for discrimination based on ethnicity, and possible challenges to immigration²⁰. Institutions could be subject to increasing numbers of legal claims, or companies could suffer reputational or economic damage³⁹. It was also noted that some communities may be particularly at risk of being exploited by research, especially during emergency outbreak situations^{14,32}.

On the other hand, it was acknowledged that widespread availability and use of sequence data contributes important benefits to the clinical and research communities¹⁷. In particular, the rapid sharing of data can help identify etiological factors, predict disease spread, evaluate existing and novel treatments, symptomatic care and preventive measures, and guide the deployment of

Table 2. Summary of studies.

Authors, Year, Country	Year	Country	Disease
Empirical studies			
Davies, A., Scott, S., Badger, S, Török, M. E., & Peacock, S.	2015	United Kingdom	Tuberculosis
Degeling C, Johnson J, Gilbert GL.	2019	Australia	Flu, MRSA and Listeria
Ribeiro Dos, Santos., Martine Y. van Roode , George B. Haringhuizen, Marion P. Koopmans, Eric Claassen, & Burgwal 2018	2018	The Netherlands	Non specific
Mutenherwa F, Wassenaar DR, de Oliveira T. 2018	2018	South Africa	HIV
Rump, C. Woonink, F. Van Steenbergen, J. Verweij, M. Hulscher, M	2017	The Netherlands	Food borne Pathogens
Schairer C, Mehta SR, Vinterbo SA, Hoenigl M, Kalichman M, Little S	2017	United States	HIV
Shean, R, & Greninger, A	2018	Switzerland	Non specific
Ethics frameworks or guidelines			
Coltart CEM, Simwinga M, Eba P, Grabowski MK, Amon JJ, Baggaley R, <i>et al.</i>	2018	United Kingdom	HIV
David Evans, Nannette Benbow	2017	United States	HIV
World Health Organisation	2018	Switzerland	Non specific
Reviews of the ethical aspects			
Greninger AL.	2019	United States	Non specific
Gilbert M, Swenson L, Unger D, Scheim A, Grace D.	2016	Canada	HIV
Johnson SB, & Parker M.	2019	United Kingdom	Non specific
Mehta, S. R., Vinterbo, S. A., Little, S. J.	2014	United States	HIV
Mutenherwa F, Wassenaar DR, de Oliveira T.	2019	South Africa	HIV
Ribeiro, Koopmans & Haringhuizen	2018	The Netherlands	Non specific
Rump, B., Cornelis, C. Woonink, F. & Verweij, M.	2013	The Netherlands	Non specific
Rump, B. & Woonink, F	2012	The Netherlands	Hepatitis A, Group A Streptococcus
Contained a section on ethics			
Bhattacharya S	2014	United Kingdom	Hepatitis C and HIV
Boccia S, Pasquarella C, Colotto M, Barchitta M, Quattrocchi A, Agodi A.	2015	Italy	Hospital Acquired Infections
Gardy JL, Loman NJ.	2018	Canada and United Kingdom	Non specific
German D, Grabowski MK, Beyrer C.	2017	United States	HIV
Grubaugh, Nathan D	2019	United States	Non specific
Gwinn M, MacCannell DR, Khabbaz RF.	2017	United States	Non specific
PHG Foundation	2015	United Kingdom	Non specific
Thaler DS, Head MG, Horsley A.	2019	Switzerland	Non specific
Thorogood A, Zawati MH, Knoppers BM.	2014	Canada	Non specific
Yozwiak NL, Schaffner SF, Sabeti PC.	2015	United States	Ebola
Empirical studies			
Davies, A., Scott, S., Badger, S, Török, M. E. & Peacock, S.	2015	United Kingdom	Tuberculosis
Degeling C, Johnson J, Gilbert GL.	2019	Australia	Flu, MRSA and Listeria
Ribeiro Dos, Santos., Martine Y. van Roode, George B. Haringhuizen, Marion P. Koopmans, Eric Claassen, & Burgwal	2018	The Netherlands	Non-specific
Mutenherwa F, Wassenaar DR, de Oliveira T.	2018	South Africa	HIV
Rump, C. Woonink, F. Van Steenbergen, J. Verweij, M. Hulscher, M	2017	The Netherlands	Food borne Pathogens
Schairer C, Mehta SR, Vinterbo SA, Hoenigl M, Kalichman M, Little S	2017	United States	HIV
Shean, R, & Greninger, A	2018	Switzerland	Non-specific

Authors, Year, Country	Year	Country	Disease
Ethics frameworks or guidelines			
Coltart CEM, Simwinga M, Eba P, Grabowski MK, Amon JJ, Baggaley R, <i>et al.</i>	2018	United Kingdom	HIV
David Evans, Nannette Benbow	2017	United States	HIV
World Health Organisation	2018	Switzerland	Non-specific
Reviews of the ethical aspects			
Greninger AL.	2019	United States	Non-specific
Gilbert M, Swenson L, Unger D, Scheim A, Grace D.	2016	Canada	HIV
Johnson SB, & Parker M.	2019	United Kingdom	Non-specific
Mehta, S. R., Vinterbo, S. A., Little, S. J.	2014	United States	HIV
Mutenherwa F, Wassenaar DR, de Oliveira T.	2019	South Africa	HIV
Ribeiro, Koopmans & Haringhuizen	2018	The Netherlands	Non-specific
Rump, B., Cornelis, C. Woonink, F. & Verweij, M.	2013	The Netherlands	Non-specific
Rump, B. & Woonink, F	2012	The Netherlands	Hepatitis A, Group A Streptococcus
Contained a section on ethics			
Bhattacharya S	2014	United Kingdom	Hepatitis C and HIV
Boccia S, Pasquarella C, Colotto M, Barchitta M, Quattrocchi A, Agodi A.	2015	Italy	Hospital Acquired Infections
Gardy JL, Loman NJ.	2018	Canada and United Kingdom	Non-specific
German D, Grabowski MK, Beyrer C.	2017	United States	HIV
Grubaugh, Nathan D	2019	United States	Non-specific
Gwinn M, MacCannell DR, Khabbaz RF.	2017	United States	Non-specific
PHG Foundation	2015	United Kingdom	Non-specific
Thaler DS, Head MG, Horsley A.	2019	Switzerland	Non-specific
Thorogood A, Zawati MH, Knoppers BM.	2014	Canada	Non-specific
Yozwiak NL, Schaffner SF, Sabeti PC.	2015	United States	Ebola

limited resources¹⁶. Many of the ethical, legal and social issues arising from sequencing studies were seen to reflect this tension between interests that arise at the level of the individual and these other important societal interests (in public health)^{14,16-19,21,22,27,28,32,33,37,39}. Much of the literature either explicitly or implicitly reflected on how to balance these tensions. Usually, by reflecting on the permissibility of conducting sequencing studies, and on the necessary conditions of collection, storage and use of data.

Permissibility of sequencing studies and conditions of use

There was strong support for the permissibility of conducting sequencing studies, as long as potential risks were thoughtfully mitigated^{14,18,19,21,22}. In one empirical study, patients and health-care workers were asked if the benefits of HIV molecular epidemiology outweigh the risks; all said yes²². Three-quarters of respondents answered with an unqualified, yes, and one quarter gave a positive answer with qualifications, such as 'It's very necessary, just as long as parameters are set in place and

they're kept', or 'with proper protections in place, the benefits outweigh the risks'²². In another study, expert Delphi panelists held that the protection of the public was of overriding importance, but that most of the potential harms could be managed¹⁹.

There were differences across the literature in the priority afforded to different conditions of use, and to the types of risk or amount of risk deemed acceptable. It was broadly agreed that any research should have a favourable risk benefit ratio^{14,20}, and that maximizing the utility of data must be weighed against concerns over interests of individuals and that policies on data collection and release should seek to align the interests of different parties³⁶. There was, however, disagreement as to whether privacy concerns or public interest should take precedence^{24,25,31} and some noted that the balance between the public health benefit and personal privacy risk for individuals whose genetic data (personal or pathogen) are included is difficult to delineate, since neither the true benefit nor the actual risk to participants has been adequately defined^{27,35}.

Below we set out the key recommendations from the literature on the conditions of use for pathogen sequence data. **Box 1** summarises recommendations from the literature for future research focus and study design.

Publication

It was clear that the release of information of relevance to public health should not be delayed by publication timelines or concerns over academic ownership of data^{16,26,32}. Recommendations to address such conflicts of interests included: that medical journals should update their policies to support pre-publication sharing of pathogen sequence data related to outbreaks¹⁶; publication disclaimers prohibiting use of sequence data for publication without permission^{16,32}; acknowledgment of data sharing contributions and the inclusion of such criteria to the assessment of academic research credit¹; establishment of governance structures and dispute resolution mechanisms that can mediate where disagreements arise¹⁶.

Anonymization and privacy protection

Much of the literature suggested that traditional methods of de-identification or anonymization of data are insufficient to meet their purpose in the context of pathogen sequencing^{2,14,17,23,28,33,39}. Existing approaches to minimize the risk of privacy loss to participants are based on de-identification of data by removal of a predefined set of identifiers¹⁷. However, this has three key limitations in the context of pathogen genomics. First, sharing of corresponding sample metadata (minimally time and place of collection, ideally with demographic, laboratory, and clinical data) is essential to enhance the interpretation and the value of genomic data¹⁶, therefore removal of key identifiers such as geographic location may severely limit the utility of genomic data³⁹. Second, removing predefined identifiers may be ineffective at protecting privacy and confidentiality^{17,19,22,39}. For example, one study demonstrated how sample collection dates associated with microbiological testing at a large tertiary hospital were highly correlated with patient admission date (protected health information), meaning data is re-identifiable²³. Small study populations may also mean that individuals who are part of a transmission chain may be able to identify others during the course of routine contact tracing (e.g. sexual partners)¹⁷. Third, anonymization of data does little to mitigate potential risk to communities and groups^{14,17,39}. Perhaps a consequence there was clear support for re-visioning of existing privacy standards, and for privacy policies specific to the context of sequencing studies.

Consent

There was debate in the literature around the importance of consent to the use of sequence data and associated meta-data in epidemiological investigation. In the research setting, Coltart *et al.*¹⁴ state that research participants and patients whose samples are being used for phylogenetic analysis should ideally have consented to such use, but suggest that when using data from previous studies, where only broad consent for HIV-related research might have been obtained, waivers of specific consent are allowable when samples are no longer linked to identifiers, or when broad consent for sample collection for research and storage in future studies was given¹⁴. In the public health and clinical setting, an Australian study reported that one of the key differences amongst participants in a modified Delphi

Box 1. Recommendations from the literature for future research focus and study design

Social science and ethics research

1. Community engagement should occur early in the research design process, ensuring that phylogenetic research is relevant to participating communities and that local perspectives are included in the design and overall conduct of research studies¹⁴.
2. Preparation for future outbreaks should include provisions for rapidly building new bridges and establishing community norms³².
3. Future research on the perceptions of stakeholders should include standardised background information³⁵.
4. Social and behavioural research into conceptual and normative aspects should be backed up by empirical research³⁵.
5. New inter-disciplinary collaborations including microbiologists, engineers and bioethicists³⁰.
6. As real-time and other intervention strategies that build on HIV phylogenetic information continue to emerge, it will be critical to address questions of efficacy for cluster growth interventions to ensure that the benefits outweigh potential risks. Implementation science research may also inform best practices for discussing the meaning and limitations of sequence data and cluster membership with community members and help to identify acceptable and evidence-based approaches that impose the least risk to persons within specific contexts. These might involve partnerships with providers for non-intrusive patient follow-up related to clusters, more detailed consent procedures for future follow-up related to HIV test results or partner services referrals, and specific guidelines and education to mitigate criminalization risks²⁸.
7. Communication methods that increase the understanding of phylogenetic studies need to be designed and evaluated. These must emphasise potential harms, thoughtful mitigation of harms to risk groups, processes for monitoring risk, and clear protection procedures to minimise risks¹⁴.
8. Collaboration between stakeholders is necessary, with an active exchange of experiences and best practices. The first step, should be sought in creating awareness and consensus within sectors on the causal factors of barriers to sharing of sequence data¹.

Ethical conduct of studies

9. Need to pre-define exceptional circumstances where un-validated techniques might be used in emergency situations²⁷.
10. To ensure scientific validity, researchers and their associates should be competent to implement the proposed study design. In order to maximize scientific validity, the researchers should ensure that they have all necessary resources, that the community accepts the protocol and that a competent and independent research ethics committee (REC) or institutional review board (IRB) reviews and approves the protocol³⁵.
11. The scientific objectives of research should guide the choice of participants and determine the inclusion criteria and appropriate recruitment strategies. It is unethical to use privilege, convenience and/or vulnerability as criteria for selecting participants. Exclusion of certain population sub-groups or communities in a research study without appropriate scientific justification is also considered unethical³⁵.
12. Risk mitigation strategies must also provide for redress mechanisms in cases of abuse or misuse of phylogenetic data. These strategies might require the establishment of ties with local legal services, organisations working to protect people with HIV, and criminalised or stigmatised populations, to ensure that they have access to the means to protect their rights¹⁴.

study included the necessity for consent before testing and data-linkage. No panelists agreed with the statement “under no conditions should a study be conducted without prior consent”, although only ten of thirty agreed that consent is not required under any conditions²¹. In a Dutch study, outbreak managers thought intervention without seeking explicit consent of all individuals involved is justified when there is at least be a substantial public health threat, realistic expectation that deploying the techniques will help to mitigate the outbreak, and that source and contact tracing would most likely not be successful without the use of molecular typing techniques²¹.

Data access

There was a strong commitment to rapid and open data-sharing, particularly in emergency or outbreak situations^{16,27,32} and in such conditions for incentives and safeguards to encourage rapid and unrestricted access to data release^{16,27,32}. The World Health Organization (WHO) recommended that in emergency outbreak situations “the first set of sequences providing crucial information on the pathogen, genotype, lineage, and strain(s) causing the outbreak should be generated and shared as rapidly as possible. Sharing of corresponding anonymised sample metadata (minimally time and place of collection, ideally with demographic, laboratory, and clinical data) is essential to enhance the interpretation and the value of genomic data”¹⁶. However, access to data gathered as part of clinical care was seen as ethically more contentious as “publicly accessible databases are not an appropriate storage location for the level of metadata required to enable clinical and epidemiological analysis for the purposes of providing patient and population care”³⁹. Suggestions were made for a tiered approach to data release, whereby a separate database governed by appropriate public health authorities would collate and store metadata in a location to which access to data could be limited to users with a legitimate clinical or public health need to use it, and data that cannot be released into public domains but is needed by authorised healthcare and public health professionals for service delivery remains within a suitable secured access database³⁹.

A public survey on TB explored questions related to database access and the potential benefits and risks associated with it. Most felt that medical professionals and the research community should have access to such a database; and a significant proportion thought that other agencies, such as the police (10%) and immigration officials (13%), should also have access to the genomic database. Experts, however, were clear that they felt transmission data should not be used in litigation; this was partially because it was deemed too unreliable^{24,27} and also because of the potential for ‘abuse’ of data^{14,15,24}.

Overall, there was broad support for further work in defining the conditions for collection use and storage of data and samples^{2,27,30,32,37} and for policy and legal clarity to aid the ethical implementation of these technologies. This will require more work to carefully assess and understand risks³⁹; research to decide how much individual privacy might be risk in the name of public health³¹; consideration of alternative strategies required to mitigate this risk, such as suppression of data in the public domain where it may cause serious harm; and adjustments to communication plans¹⁴.

Themes 2: Difficulties of translation from research to clinical and public health practice

Effective phylogenetic work often occurs at the interface between research and public health practice because the same data can be used for both purposes¹⁴. In this regard, pathogen sequencing was described as ‘straddling the boundary between research and clinical use’²⁷. The hybrid nature of sequencing activities imposes important ethical challenges.

Genomics and the clinic

At present, sequencing technology results are produced with a substantial delay from sampling; therefore, any results are unlikely to be timely in informing clinical care^{14,39}. However, with the evolution of real-time phylogenetic and other typing techniques, reporting of results to study participants could result in changes to clinical management¹⁴. This could present a number of new problems for clinical practice. For example, whether HIV acquisition events are linked to the known infected partner might be crucial for interpretation of the efficacy of prevention strategies, but feeding back individual results that indicate source of HIV acquisition in discordant couples would require ethical reflection¹⁴. Likewise, routine surveillance could impact on clinical and public health encounters. In one empirical study interviewees pointed out:

“So, what would it mean? At some point we knock on someone’s door and say, “from the data we have it looks like you are infecting a lot of people and we need you to stop doing that.” This then becomes an issue in the public health setting and it seems you cannot do that in a research study if people can now perceive you can knock at individual doors and say you have a problem here, then people might be very reluctant to participate in further research studies and might be very reluctant to provide information to public health professionals because of the fear that they could be singled out”²⁰.

Clinical implementation of metagenomics sequencing (un-targeted testing) has the potential to detect unexpected or incidental findings that may include infections with hepatitis or HIV³⁹. Incidental findings of a different type may occur if non-germline samples (such as faecal samples) are contaminated with germline cells, which could potentially reveal predictive information about developing inherited disease³⁹. Furthermore, informed consent for phylogenetic studies that are difficult complex and difficult to understand, and in which the benefits and risk may not be fully determined, may also be difficult to achieve³⁵. Mutenherwa suggests that where sequence data are generated for routine clinical management, its subsequent use for research and surveillance may be underestimated by patients²⁰. Others suggested that the right to withdraw from research activities—a key indicator of voluntary participation in research—was overlooked by expert stakeholders²⁰.

Professional boundaries

Understanding and interpreting phylogenetic data requires significant expertise^{25,27} and presents a challenge to established professional boundaries. Expertise in phylogenetic studies creates new obligations for researchers, such as deciding whether or not to participate in forensic investigations and potential

prosecutions of individuals²⁴, and to consider the down-stream uses and misuses of data¹⁴. The routine implementation of pathogen sequencing studies may create new responsibilities for clinical microbiologists (related to public health)², and require major changes in culture such that diagnostic interpretation, therapeutic management decisions and antimicrobial treatment regimes are delegated to physicians instead of microbiologists³¹.

Theme 3: Truth-telling, trust and community engagement

Many noted that there are important reasons to ensure that the public and individuals understand the uses of data collected as part of a sequencing studies, and the potential risks. First, this was seen to have some intrinsic value in that it supports patient autonomy and truth telling is a respected moral virtue^{14,15,19,21}. Second, truth-telling was seen to be important because it may lead to better outcomes in research and public health practice. This is both because this was deemed to promote trust in research and therefore lead to increased participation, and because it promotes disclosure, which is helpful from a public health perspective. Third, promoting understanding of uses and risks of data was also seen as a way of avoiding harm and exploitation of vulnerable individuals and communities, by enhancing understanding of risks that may be specific to that them. In some cases, this was balanced by a number of practical challenges to telling people the truth, such as: risk of fear mongering¹⁹; information needed for legal proceedings in public interest¹⁴; and the fact that it may be difficult to adequately inform the public and/or ensure full understanding²⁰. None-the-less, there was a clear recommendation in the literature to raise public awareness and understanding of these techniques^{19,22,30}, and for early and meaningful community engagement^{15,22,32,34} prior to conducting sequencing studies. This was seen as particularly important when working with vulnerable groups or when the risks of participation are high.

Theme 4: Justice and global partnerships

The notion of justice appeared to be a widely recognized ethical principle in the field¹. Stakeholders pointed to the importance of equitable access to data²⁷, and to benefit-sharing obligations^{16,26}. This included an ethical imperative that outbreak related research and countermeasures, such as diagnostics and vaccines, should be accessible to all affected countries²⁶, and towards reciprocal arrangements such that countries that participate in sequencing activities should derive some corresponding local benefit²⁷.

Collaboration

Collaboration between researchers from Africa and HICs was raised as an important ethical consideration²⁰. In one empirical study, interviewees were concerned that African researchers were not meaningfully engaged in the scientific research process in health research in general and phylogenetic research in particular, and that for equitable and mutually beneficial collaborative research partnerships to be realized, local researchers were encouraged to take leading and active roles throughout the research process²⁰. This type of collaborative research practice was supported elsewhere in the literature^{14,27,37}. For some,

this was to enhance equity as well as to help maximize the utility of data and lead to better public health outcomes. It was argued that local researchers were more likely to understand their health care and research systems and study results were more likely to be easily translated into policy²⁰, and that context specific responses to particular outbreaks were likely to be required. Recommendations were made to conduct studies exploring the nature of existing collaborative partnerships between researchers from low- and middle-income countries (LMICS) and HICs to explore team composition and distribution of roles, including contribution to intellectual property²⁰. Outbreak related research and countermeasures, such as diagnostics and vaccines, must be accessible to all affected countries not only as a legal obligation, but also as an ethical imperative²⁶.

Global cooperation

It was also noted that global and interdisciplinary partnerships are a necessary component of an effective genomic informed response to infectious disease. This was because of the vast range of stakeholders and varying interests involved in control of infectious disease outbreaks, and because issues may resist simple resolution and span multiple jurisdictions²⁷. For example, conflict may result from governments wishing to keep an outbreak quiet and/or from the tension between LMICs with few resources for generating and using data and the researchers or response teams from better-resourced settings^{1,37}.

Ownership of samples and data was seen as an important barrier to global cooperation. The Nagoya Protocol (NP), for example, was developed to facilitate access to genetic resources and the fair and equitable sharing of benefits arising from their utilization¹. Nevertheless, despite the importance of reinforcing sovereignty rights of States over genetic resources in their territory, uncertainties about intellectual property rights and the resulting disputes hamper access to samples^{1,32}. Ribeiro *et al.*¹ explain that:

“The real or perceived possibilities for the commercial valorization of microbial genetic resources (MGR) has enforced their appropriation for further use in research, innovation and product development. The problem for public health surveillance occurs when such appropriation is triggered at initial (upstream) phases of the research and innovation cycle, such as sampling and sequencing of microorganisms, instead of later stages, such as the actual product development (in this case drugs, diagnostics and vaccines). As such, stakeholders are reluctant to share their (intangible) assets even in early phases of the innovation process, decreasing the scope of innovation efforts due to the lack of access to upstream research inputs.”

The same authors suggest that standardized and simplified sharing agreements²⁶, and collaboration between stakeholders with an active exchange of experiences and best practices¹ are required.

In general, recommendations were made for a global approach to ethics, policy and legal frameworks^{16,27,29,32,37}. For example, it was suggested global data sharing arrangements should include “a global data governance or ethical framework,

supplemented by local memoranda of understanding that take into account the local context²⁹; or to investigate how the Global Alliance for Genomics and Health (GA4GH) framework for responsible data-sharing could be adapted for digital pathogen surveillance²⁷.

Themes 5: The importance of context

Lastly, it was clear that the types of ethical issues likely to arise are in part dependent upon the contexts in which studies are conducted, as well as the nature of the pathogen under study. Chiefly, information that may impact on interpersonal relationships was viewed as particularly sensitive and therefore, worthy of additional ethical reflection. Examples included: sexually transmitted infections^{14,20}; consent requirements to use isolates collected from dead neonates for the purposes of epidemiological research¹⁹; and disclosure of family members as the source of infection³⁸. It was suggested that the balance of risks to patients and public health benefits is likely to be affected by the characteristics of the pathogen, in terms of likely morbidity and mortality; infectivity; treatability and drug resistance³⁹. Stakeholders also suggested that the ethical permissibility of sharing data about, particularly with regards to the source of transmission, may be different in professional contexts, where healthcare providers or companies are seen to carry a responsibility to control risk, as opposed to outside of professional contexts where protecting individuals from ‘naming and shaming’ may be of greater concern²¹. It was also noted that the legal and regulatory structures in which studies are conducted may also influence the implementation and ethics of conducting pathogen sequencing studies. In particular, use of phylogenetic analyses in criminal convictions was raised as an ethical risk^{14,17,24}.

Limitations of the review

Although quality assessment of all included materials is desirable in systematic reviews, it was not possible in this case due to the inclusion of a diverse range of research formats and literature, such as commentaries and ethics guidelines. A second limitation of this review is that a large proportion of the literature included related to phylogenetic and HIV specifically (8 out of 28), meaning that the issues relevant to this context may be over-represented.

Conclusion

This review highlights that while pathogen sequencing has the potential to be transformative for public health and clinical practice and to bring about important health benefits, there are a number of key ethical issues that must be addressed. In particular, there was clear support in the literature for innovative and critical thinking around the conditions of use for pathogen sequence data. This includes context specific standards of practice for consent, data collection, use and sharing. These practices should be informed by public values, and further empirical work investigating stakeholders’ views are required. This should include experts in pathogen sequencing, patients and the general public, as well as end users such as public health professionals and clinicians. Lastly, it is both a scientific and an ethical imperative that development in the field is under-pinned by a strong commitment to values of justice, in particular global health equity.

Data availability

All data underlying the results are available as part of the article and no additional source data are required.

References

- Ribeiro CDS, van Roode MY, Haringhuizen GB, *et al.*: **How ownership rights over microorganisms affect infectious disease control and innovation: A root-cause analysis of barriers to data sharing as experienced by key stakeholders.** *PLoS One.* 2018; 13(5): e0195885. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Greninger AL: **Societal Implications of the Internet of Pathogens.** *J Clin Microbiol.* 2019; 57(6): pii: e01914–18. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Allard MW, Strain E, Melka D, *et al.*: **Practical Value of Food Pathogen Traceability through Building a Whole-Genome Sequencing Network and Database.** *J Clin Microbiol.* 2016; 54(8): 1975–83. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- England world leaders in the use of whole genome sequencing to diagnose TB [press release].** 2017. [Reference Source](#)
- Raghwanji J, Redd AD, Longosq AF, *et al.*: **Evolution of HIV-1 within untreated individuals and at the population scale in Uganda.** *PLoS Pathog.* 2018; 14(7): e1007167. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- The Centre for Genomic Pathogen Surveillance (CGPS): **Gonorrhoea surveillance study maps antibiotic resistance across Europe.** 2018. [Reference Source](#)
- Geller G, Dvoskin R, Thio CL, *et al.*: **Genomics and infectious disease: a call to identify the ethical, legal and social implications for public health and clinical practice.** *Genome Med.* 2014; 6(11): 106. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Dudas G, Carvalho LM, Bedford T, *et al.*: **Virus genomes reveal factors that spread and sustained the Ebola epidemic.** *Nature.* 2017; 544(7650): 309–15. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Quick J, Loman NJ, Duraffour S, *et al.*: **Real-time, portable genome sequencing for Ebola surveillance.** *Nature.* 2016; 530(7589): 228–232. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Johns Hopkins Berman Institute of Bioethics: **Center for Bridging Infectious Disease, Genomics, and Society (BRIDGES).** 2019. [Reference Source](#)
- Arksey H, O'Malley L: **Scoping studies: towards a methodological framework.** *Int J Soc Res Methodol.* 2005; 8(1): 19–32. [Publisher Full Text](#)
- Armstrong R, Hall BJ, Doyle J, *et al.*: **Cochrane Update. ‘Scoping the scope’ of a cochrane review.** *J Public Health (Oxf).* 2011; 33(1): 147–50. [PubMed Abstract](#) | [Publisher Full Text](#)
- Bull S, Roberts N, Parker M: **Views of Ethical Best Practices in Sharing Individual-Level Data From Medical and Public Health Research: A Systematic Scoping Review.** *J Empir Res Hum Res Ethics.* 2015; 10(3): 225–38. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Coltart CE, Simwima M, Eba P, *et al.*: **Ethical considerations in global HIV phylogenetic research.** *Lancet HIV.* 2018; 5(11): e656–e66. [PubMed Abstract](#) | [Publisher Full Text](#)
- Evans D, Benbow N: **Ethical Considerations for a public health response using molecular HIV surveillance data: A multi-stakeholder approach.** 2017. [Reference Source](#)
- WHO: **WHO’s code of conduct for open and timely sharing of pathogen genetic sequence data during outbreaks of infectious disease.** 2018. [Reference Source](#)

17. Mehta SR, Vinterbo SA, Little SJ: **Ensuring privacy in the study of pathogen genomics.** *Lancet Infect Dis.* 2014; **14**(8): 773–7.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
18. Davies A, Scott S, Badger S, *et al.*: **Public perceptions of bacterial whole-genome sequencing for tuberculosis.** *Trends Genet.* 2015; **31**(2): 58–60.
[PubMed Abstract](#) | [Publisher Full Text](#)
19. Degeling C, Johnson J, Gilbert GL: **Perspectives of Australian policy-makers on the potential benefits and risks of technologically enhanced communicable disease surveillance - a modified Delphi survey.** *Health Res Policy Syst.* 2019; **17**(1): 35.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
20. Mutenherwa F, Wassenaar DR, de Oliveira T: **Experts' Perspectives on Key Ethical Issues Associated With HIV Phylogenetics as Applied in HIV Transmission Dynamics Research.** *J Empir Res Hum Res Ethics.* 2018; **14**(1): 61–77.
[PubMed Abstract](#) | [Publisher Full Text](#)
21. Rump B, Cornelis C, Woonink F, *et al.*: **Using typing techniques in a specific outbreak: the ethical reflection of public health professionals.** *Epidemiol Infect.* 2017; **145**(7): 1431–6.
[PubMed Abstract](#) | [Publisher Full Text](#)
22. Schairer C, Mehta SR, Vinterbo SA, *et al.*: **Perceptions of molecular epidemiology studies of HIV among stakeholders.** *J Public Health Res.* 2017; **6**(3): 992.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
23. Shean RC, Greninger AL: **Private collection: high correlation of sample collection and patient admission date in clinical microbiological testing complicates sharing of phylodynamic metadata.** *Virus Evol.* 2018; **4**(1): vey005.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
24. Bhattacharya S: **Science in court: Disease detectives.** *Nature.* 2014; **506**(7489): 424–6.
[PubMed Abstract](#) | [Publisher Full Text](#)
25. Boccia S, Pasquarella C, Colotto M, *et al.*: **Molecular epidemiology tools in the management of healthcare-associated infections: towards the definition of recommendations.** *Epidemiol Prev.* 2015; **39**(4 Supplement 1): 21–6.
[PubMed Abstract](#)
26. Dos S Ribeiro C, Koopmans MP, Haringhuizen GB, *et al.*: **Threats to timely sharing of pathogen sequence data.** *Science.* 2018; **362**(6413): 404–6.
[PubMed Abstract](#) | [Publisher Full Text](#)
27. Gardy JL, Loman NJ: **Towards a genomics-informed, real-time, global pathogen surveillance system.** *Nat Rev Genet.* 2018; **19**(1): 9–20.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
28. German D, Grabowski MK, Beyrer C: **Enhanced use of phylogenetic data to inform public health approaches to HIV among men who have sex with men.** *Sex Health.* 2017; **14**(1): 89–96.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
29. Gwinn M, Maccannell D, Armstrong GL: **Next-Generation Sequencing of Infectious Pathogens.** *JAMA.* 2019; **321**(9): 893–4.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
30. Thaler DS, Head MG, Horsley A: **Precision public health to inhibit the contagion of disease and move toward a future in which microbes spread health.** *BMC Infect Dis.* 2019; **19**(1): 120.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
31. Thorogood A, Zawati MH, Knoppers BM: **Point-of-care genetic tests for infectious disease: Legal considerations.** *Current Pharmacogenomics and Personalized Medicine.* 2014; **12**(1): 43–50.
[Publisher Full Text](#)
32. Yozwiak NL, Schaffner SF, Sabeti PC: **Data sharing: Make outbreak research open access.** *Nature.* 2015; **518**(7540): 477–9.
[PubMed Abstract](#) | [Publisher Full Text](#)
33. Grubaugh ND, Ladner JT, Lemey P, *et al.*: **Tracking virus outbreaks in the twenty-first century.** *Nat Microbiol.* 2019; **4**(1): 10–9.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
34. Gilbert M, Swenson L, Unger D, *et al.*: **Need for robust and inclusive public health ethics review of the monitoring of HIV phylogenetic clusters for HIV prevention.** *Lancet HIV.* 2016; **3**(10): e461.
[PubMed Abstract](#) | [Publisher Full Text](#)
35. Mutenherwa F, Wassenaar DR, de Oliveira T: **Ethical issues associated with HIV phylogenetics in HIV transmission dynamics research: A review of the literature using the Emanuel Framework.** *Dev World Bioeth.* 2019; **19**(1): 25–35.
[PubMed Abstract](#) | [Publisher Full Text](#)
36. Garattini C, Raffle J, Aisyah DN, *et al.*: **Big Data Analytics, Infectious Diseases and Associated Ethical Impacts.** *Philos Technol.* 2019; **32**(1): 69–85.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
37. Johnson SB, Parker M: **The ethics of sequencing infectious disease pathogens for clinical and public health.** *Nat Rev Genet.* 2019; **20**(6): 313–315.
[PubMed Abstract](#) | [Publisher Full Text](#)
38. Rump B, Cornelis C, Woonink F, *et al.*: **The need for ethical reflection on the use of molecular microbial characterisation in outbreak management.** *Euro Surveill.* 2013; **18**(4): 20384.
[PubMed Abstract](#) | [Publisher Full Text](#)
39. **Pathogen genomics into practice.** PHG Foundation. 2015.
[Reference Source](#)
40. Rump BO, Woonink F: **Ethical Questions Concerning the Use of Molecular Typing Techniques in the Control of Infectious Diseases.** *Public Health Ethics.* 2012; **5**(3): 311–3.
[Publisher Full Text](#)

Open Peer Review

Current Peer Review Status:  

Version 1

Reviewer Report 24 August 2020

<https://doi.org/10.21956/wellcomeopenres.17333.r39902>

© 2020 Joly Y. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

 **Yann Joly** 

Center for Genomics and Policy (CGP), McGill University, Montreal, Canada

This manuscript presents the results of a scoping review of the literature on the topic of the ethical issues in the use of pathogen genome sequencing technologies. The authors make excellent use of the scoping review methodology which they implemented clearly (good use of tables!) and rigorously.

There are minor formatting issues and typos:

p. 4 under theme 1: **Somewere** concerned about forced testing either of certain groups such as gay men²⁸ or healthcare workers^{19,21,38-40}. There was acknowledgment of individual professional interests of **researchersand** practitioners to ownership and use of data.

p. 9 under theme 3: 'that may be specific to **that** them'

The main substantial issue is the substantial focus of the manuscript on the HIV context (8 articles out of 28) which the authors acknowledge. HIV is a particularly stigmatized, serious condition, with lifelong consequences for patients which is not the case for many other communicable diseases. The problem is that in the past few months, COVID-19 has been a game changer in the field. COVID-19 is both much more contagious but less stigmatizing and, for most people, less dangerous than HIV. Considering that the last update to this scoping review was made in January 2020, none of the COVID-19 emerging literature was considered. The result, which is not the fault of the authors, is that the article will only have a limited relevance to the current global pandemic. Given the importance of COVID to the field and beyond, it could be worth it for the author to take the time to update their research accounting for very recent developments before publishing. Otherwise, the publication maybe perceived as already outdated and not garner much attention from readers.

A second issue of the manuscript is that while it acknowledges the tension and blurry demarcation between the research and the public health context, it doesn't really provide any solution in this regard. For example, in the context of pathogen genome sequencing for outbreak surveillance during a public health emergency, informed consent is often not required. However, the lack of consent can create issue later for data sharing with the research community. Such a scenario is

not really discussed in the manuscript. Similarly, the impact of the public health vs. research situation on the potential requirement for ethics review is not discussed. Perhaps this was not touched upon in the literature, but it is certainly a preoccupation of researchers in the COVID context.

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Ethical, legal and social issues of genomic research, medical law, bioethics.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 06 July 2020

<https://doi.org/10.21956/wellcomeopenres.17333.r39365>

© 2020 Armstrong G. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Gregory L. Armstrong 

Office of Advanced Molecular Detection, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, GA, USA

The article is generally well written, although it would be stronger if it had a more concise, more focused set of recommendations for readers to act on. The focus of the paper is mostly on the use of microbial genomic data for research but also touches a little on the issues around public health use of the data. This reviewer is not an expert in ethics; that said, the ethical concepts brought out

in the paper in the review appear to be appropriate and informative. Some additional concepts that weren't included, possibly because the article presents a review of already published ideas, are the following:

- The potential for civil legal consequences from publication of genomic data. The article discusses at a few different points the potential for criminal prosecution, but publication of data could also open up participants of a study to civil penalties as well.
- The potential for violation of a privacy when cases are rare. For example, early in an outbreak, as is highlighted in the article, there is an urgency to making sequence data from the pathogen in question public. However, it's common early in an outbreak for the media to discover and publish the names of early cases. In that situation (which is relatively common), there's a risk that sequence data made public could be linked to a specific person.
- One concern that is often under-appreciated is that, under certain circumstances, a researcher or public health agency might be legally compelled to release the identity of someone. In general, public health laws are quite strong in shielding public health data, but that may not be the case with research data. If there is uncertainty about whether such data are protected, any participants in research should be notified.
- GH4GE is mentioned in the manuscript, but there's no mention of PHA4GE (<https://pha4ge.github.io/>), which is much more applicable here.
- Issues around the Nagoya Protocol are addressed, but the tension between GISAID and INSDC is not. GISAID provides protections for intellectual property rights that the INSDC members do not, and such protections are key for participation of LMICs. However, where public funds are used to obtain data, there is a strong argument that the data should be made publicly available (i.e., through INSDC) without restriction as long as privacy and confidentiality are not placed at risk. The tension between the two models has been particularly strong with the advent of COVID-19, and the very assertive push by GISAID to prevent researchers from submitting to INSDC or from citing it.

Some (minor) specific issues that need to be addressed:

- I would remove the reference to Microsoft Excel. There's a strong argument for including information about statistical software, but which spreadsheet is used is not particularly relevant.
- "Theme 1: ...": There are typos in the first paragraph.
- Box 1, Item 5: This is not a complete, declarative sentence like the others.
- Reference 29 appears to be the incorrect reference. Also, the quote from it (Pages 9-10) appears to be a quote from a separate document (that should be cited separately, if still available).

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Not applicable

Are all the source data underlying the results available to ensure full reproducibility?

No source data required

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Pathogen genomics to support public health.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.
