The Journal of Infectious Diseases

MAJOR ARTICLE

# Addressing Ethical Challenges in US-Based HIV Phylogenetic Research

**Liza Dawson,[1] Nanette Benbow,[2] Faith E. Fletcher,[3] Seble Kassaye,[4] Amy Killelea,[5] Stephen R. Latham,[6] Lisa M. Lee,[7] Thomas Leitner,[8] Susan J. Little,[9] Sanjay R. Mehta,[9] Omar Martinez,[10] Brian Minalga,[11] Art Poon,[12] Stuart Rennie,[13] Jeremy Sugarman,[14] Patricia Sweeney,[15] Lucia V. Torian,[16] and Joel O. Wertheim[9]**

[1]Walter Reed Army Institute of Research, Silver Spring, Maryland, USA, [2]Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA, [3]School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama, USA, [4]Georgetown University Medical Center, Washington, District of Columbia, USA, [5]National Alliance of State and Territorial AIDS Directors, Washington, District of Columbia, USA, [6]Yale Interdisciplinary Center for Bioethics, New Haven, Connecticut, USA, [7]Virginia Tech, Blacksburg, Virginia, USA, [8]Los Alamos National Laboratory, Los Alamos, New Mexico, USA, [9]University of California at San Diego, San Diego, California, USA, [10]Temple University, Philadelphia, Pennsylvania, USA, [11]Fred Hutchinson Cancer Research Center, Seattle, Washington, USA, [12]University of Western Ontario, London, Ontario, Canada, [13]University of North Carolina, Chapel Hill, North Carolina, USA, [14]Johns Hopkins University, Baltimore, Maryland, USA, [15]Centers for Disease Control and Prevention, Atlanta, Georgia, USA, and [16]New York City Department of Health, New York, New York, USA

**(See the Editorial Commentary by Lapointe and Harrigan, on pages 1939–40.)**

In recent years, phylogenetic analysis of HIV sequence data has been used in research studies to investigate transmission patterns between individuals and groups, including analysis of data from HIV prevention clinical trials, in molecular epidemiology, and in public health surveillance programs. Phylogenetic analysis can provide valuable information to inform HIV prevention efforts, but it also has risks, including stigma and marginalization of groups, or potential identification of HIV transmission between individuals. In response to these concerns, an interdisciplinary working group was assembled to address ethical challenges in US-based HIV phylogenetic research. The working group developed recommendations regarding (1) study design; (2) data security, access, and sharing; (3) legal issues; (4) community engagement; and (5) communication and dissemination. The working group also identified areas for future research and scholarship to promote ethical conduct of HIV phylogenetic research.

**Keywords.** HIV/AIDS; ethics; phylogenetics; public health.

In recent years, phylogenetic analysis of human immunodeficiency virus (HIV) sequence data from individuals living with HIV has been used to investigate transmission patterns between individuals and groups. The term HIV phylogenetic analysis refers to a specific set of techniques that compares HIV genetic sequences from different sources to analyze the evolutionary relationships between the HIV sequences from those sources. There are other types of analyses of HIV sequence data, other than phylogenetic techniques, that also raise similar ethical issues. However, this report is focused on phylogenetic techniques due to heightened interest in using these particular techniques to understand HIV transmission patterns, which has particular implications for individuals and communities. These methods have been used in HIV prevention trials to determine relatedness of viral sequences between sexual [1] and needle sharing partners [2], in molecular epidemiology [3, 4], and in public health surveillance programs [5], to identify rapidly growing transmission clusters [6]. This technology provides valuable information for HIV prevention interventions, but it also has risks. Identification of transmission patterns might stigmatize groups vulnerable to HIV, and phylogenetic analysis could contribute evidence to inferences about specific transmission events between individuals. A 2017 consultation on ethical considerations in molecular HIV surveillance (MHS) [7] explored the use of phylogenetics in public health programs. A recent commentary [8] highlighted concerns about MHS used by public health departments [9], and concerns have been expressed about the lack of community engagement about MHS programs or about phylogenetic research studies [7, 10]. Two recent studies [11, 12] demonstrated that the technical ability to predict directionality of transmission is increasing, and an associated commentary [13] highlighted ethical concerns about the risks to individuals and communities from these developments. Further exploration of the burdens and risks of phylogenetic research is warranted to develop appropriate ethical safeguards.

The National Institutes of Health (NIH) organized a working group in fall 2017 to identify and address ethical issues in HIV phylogenetic research. The group's recommendations are reported here.

## METHODS

The working group included members with expertise in virology, molecular epidemiology, public health, bioethics, community engagement, social work, community-based HIV research, and law.

Received 28 October 2019; editorial decision 3 February 2020; accepted 6 March 2020; published online June 11, 2020.

Correspondence: Liza Dawson, PhD, Walter Reed Army Institute of Research, 503 Robert Grant Ave., Silver Spring, MD 20910 (liza.dawson.civ@mail.mil).

**The Journal of Infectious Diseases® 2020;222:1997–2006**



DOI: 10.1093/infdis/jiaa107

A full list of working group members is in Supplementary Material. A background paper was prepared based on bioethics literature, deliberations in HIV research ethics, and working group phone calls. The group met in April 2018 and discussed 6 priority areas: (1) study design; (2) data security, access, and sharing; (3) community engagement; (4) the interface between research, public health, and clinical care; (5) legal issues; and (6) communication. Preliminary recommendations were discussed. Subsequently, 6 subgroups, each constituted with members of the original working group, were created for each of the 6 priority topics above. Each subgroup convened conference calls and exchanged emails; no formal consensus methodology was used but the entire set of recommendations was reviewed and approved by the full working group.

## SCOPE OF THE REPORT

This report was tasked to address ethical issues arising in the conduct of HIV phylogenetic research in the United States. The working group was not constituted to assess conditions in other countries, due to the extensive variation in economic, cultural, legal, and political contexts outside the United States. However, the recommendation in this report are expected to be broadly applicable, especially in similar contexts where recommendations on study design, data-sharing, community engagement, and communication are generally applicable.

One of the main concerns raised about HIV phylogenetic research is the possibility of estimating the probability of specific HIV transmission events between 2 individuals. An individual transmission event currently cannot be proven on the basis of sequence data alone; additional epidemiological or demographic information would be required. However, phylogenetic analysis implicating possible transmission clusters might lead to stigmatization or harm to individuals or groups. This report is designed to address these risks, as well as risks to privacy and the potential for fueling community mistrust of public health or research programs. The working group's recommendations aim to promote ethical conduct of research to benefit public health and scientific inquiry.

Some research aims to better elucidate viral evolution *within* an individual patient using phylogenetics; these analyses are less likely to generate concerns regarding stigma, legal consequences, or mistrust, so the main focus was on studies using phylogenetics to understand transmission patterns among groups of individuals.

The focus of this report is on the use of phylogenetic analysis in research, rather than public health. Public health activities have distinct legal and ethical mandates and guidance specific to their activities. The working group discussed legal and ethical issues relevant to the use of public health surveillance data in research.

## SUMMARY OF WORKING GROUP RECOMMENDATIONS

Recommendations for best practices and ethical conduct were developed in 5 categories: (1) study design; (2) data security, access, and sharing; (3) legal issues; (4) community engagement; and (5) communication and dissemination. In this report, all the recommendations from the subgroup on public health/research interface were included in other sections.

### Study Design

Phylogenetic research projects should be designed to produce valuable knowledge, ensure that risks are reasonable relative to the value of knowledge gained, and risks should be minimized to the extent possible. Projects that draw on data from public health surveillance programs must be ethically justified on the basis of potential benefits to public health.

#### *Recommendations*

Research projects using HIV phylogenetic analysis should use data that are anonymized or deidentified to the greatest extent possible, consistent with the project's scientific aims. Deidentification refers of the process of removing personally identifiable information to minimize risk of disclosure of identity. Anonymization refers to procedures that remove all information that could be linked back to the individual. Using deidentification procedures, sometimes links are maintained that allow reidentification of data, whereas anonymization is irreversible (for further description see NIH Data Sharing Policy and Implementation Guidance [14]). Sometimes identifiers are needed to link sequence data to other information; such use must be justified by scientific necessity. However, identifying information should never be released with sequence data; it should be used only in secure systems with restricted access and identifiers should be removed once linking activities are complete. Robust data security measures should be used in all projects.

Researchers should consider the provenance of previously collected data used for phylogenetic analyses. Researchers and oversight boards should determine whether the data are considered identifiable under the US federal human subjects regulations (45 CFR 46) and under the Health Insurance Portability and Accountability Act, if applicable. Data that are not identifiable do not require consent under these regulations [15], although there might be ethical reasons for seeking informed consent even when not legally required. Researchers should determine whether new informed consent is needed, or whether prior consent or waiver of consent from relevant institutional review boards (IRBs) is sufficient. Any waiver of consent for use of identifiable data must be consistent with federal regulations; 45 CFR 46.116 outlines 4 criteria for waiver: (1) the study must be minimal risk; (2) rights and welfare of participants not affected by waiver; (3) informed consent not practicable; (4) when appropriate, additional information provided to study participants.

In general, the default approach should be to seek informed consent for sharing research data for research purposes. Use of public health surveillance data is an exception to this default, as

it is collected under specific legal authority without individual informed consent. Other exceptions might be appropriate in specific circumstances, when justified by the project's scientific value. In all cases, individuals should be protected from risks of reidentification or social harm to the greatest extent possible, whether or not specific consent was obtained for research.

#### Data Security, Data Access, and Data Sharing

Data security is important in all research, and especially in research on HIV. In phylogenetic analysis, disclosure of HIV diagnosis, identification of transmission clusters, transmission categories, or groups at increased risk of HIV transmission/acquisition are all potentially stigmatizing [16] and responsible data management is essential.

Public health programs have specific legal and ethical mandates to collect and use data to protect the health of populations [17]. Using public health HIV surveillance data in research studies is complex because the data are collected under public health authority, which does not require informed consent [18]; data collection for research is subject to different rules. (All research funded by the Department of Health and Human Services [HHS] is subject to human subjects regulations at 45 CFR 46 (the Common Rule) [19]; 16 other federal agencies are also signatories to the Common Rule and are subject to these same standards [20]). While regulatory standards are different, public health and research activities involve the same broad ethical commitments to deliver societal benefits and minimize harms [21].

There are also different expectations between public health and research agencies regarding data sharing. Within the United States, sharing of public health surveillance data might be prohibited unless specifically authorized by public health law or policy; in contrast, research funders generally expect data sharing among researchers to improve scientific output, maximize return on investment, and promote replication of scientific findings. (For examples see NIH policy on data sharing [22] and the Gates Foundation open access policy [23]).

Some researchers have pioneered the use of secure databases for combining sources of public health data across jurisdictions in a "black box" method [24], to combine and match identifiable data across different geographic jurisdictions [25]. This method protects the confidentiality of data while enabling data sources to be effectively combined for public health purposes.

##### *Recommendations*

HIV phylogenetic research projects with identifiable or sensitive data should use appropriate security measures to prevent unauthorized use, including but not limited to use of secure encrypted servers and secure methods for electronic data transmission of data; restricting access to essential staff; requiring multifactor authentication steps for entry into data systems; restricting the copying of datasets onto hard drives or servers; prohibiting remote access; providing data security training for research staff; requiring renewable confidentiality pledges and training; and maintaining limits on data transfers via data use agreements among collaborators.

Although data sharing is required by many research sponsors and scientific journals, researchers working with public health surveillance data are prohibited by laws and policies from depositing those data in publicly accessible databases such as GenBank [26]. These exceptions should be accepted by sponsors and journals in certain cases involving public health surveillance data. Also, sometimes research data sets should not be shared in their entirety due to concerns about group stigma and/or potential reidentification of individuals. Alternative methods of sharing, such as sharing randomly selected subsets of data, or stripping the sequence data of all associated metadata (the clinical and demographic data associated with the sequence data), can be appropriate when permissible. The sensitivities around sharing research data must be reviewed on a case-by-case basis. Decisions on data sharing must be ethically justified based on considerations of the value of the data to the scientific community and benefits to the community at large; the potential risks to groups or individuals; and risk mitigation mechanisms. Relevant stakeholders for these discussions include research sponsors, researchers, representatives of affected communities, and journals.

Use of public health surveillance data for research purposes must be consistent with laws and policies in the jurisdictions where data were collected and where research takes place, if different. State laws and policies on public health surveillance do not always directly address the issue of research use of surveillance data (Box 1), and it is recommended that requests for data sharing be reviewed by an appropriate research oversight committee, IRB, or privacy board [28], or another review process for studies not considered human subjects research.

Data use agreements should be implemented whenever HIV sequence data are shared for research. The agreements should describe the specific research objectives; plans for publication; permissions for and monitoring of access to the data; data storage, security, and confidentiality; allowances for copying or remote use, if any; deidentification plans; data destruction protocols; and identification of parties responsible for data analysis and data security.

#### Legal Issues

Clarification of legal risks and relevant legal protections is essential in HIV phylogenetic research and applicable legal, policy, and regulatory standards were identified. Because some research is conducted with public health surveillance data, legal standards for both research and public health were discussed.

Thirty-four states currently have HIV criminalization statutes [29], and many states have prosecuted HIV exposure using common-law crimes such as reckless endangerment or assault

**BOX 1. STATUTES, POLICIES AND PROCEDURES REGARDING SHARING OF PUBLIC HEALTH DATA FOR RESEARCH**

**STATUTORY REQUIREMENTS**

Explicit authority: statutes specifically allow health department HIV surveillance data to be shared for research purposes (statute may require that only deidentified data is shared or if identifiable data is shared, that the published results be deidentified).

Implicit authority: statutes are silent as to health department data sharing for research purposes, and the broader statutory scheme is interpreted to allow HIV surveillance data to be shared.

The specific circumstances under which release is allowed, including whether any identifiable data can be shared, are included in health department policies.

**IRB/DATA GOVERNANCE BOARD APPROVAL**

Many health department policies require that research requests for HIV surveillance data (including deidentified data) go through an IRB, data governance approval, or specific legal counsel review given the sensitive nature of the data.

**DATA-USE AGREEMENT PARAMETERS**

Data-use agreements with health departments for use of HIV surveillance data for research purposes will include specific data security and confidentiality requirements. For instance, some states require that researchers access all HIV surveillance data solely through health department servers and will not transfer data outside of the health department.

For further detail see National Alliance of State and Territorial AIDS Directors, 2018 [27].

Abbreviations: HIV, human immunodeficiency virus; IRB, institutional review board.

with a deadly weapon [30]. Some convictions have resulted in decades-long sentences [31]. Most of the HIV-specific laws target the alleged "exposure" of a person to HIV with alleged nondisclosure of HIV status, whether or not HIV transmission occurred [32]. Most prosecutions have not sought to use phylogenetic data to provide evidence of an exposure or transmission event [33], with a few exceptions. In 2 court cases, phylogenetic evidence from blinded samples was presented to establish direction of transmission, and a defendant was alleged to be the source of infection in both cases. In another case, an individual was accused of deliberately infecting a person with blood from an HIV-infected individual; phylogenetic analysis demonstrated that viral sequences of victim and alleged source were more closely related than any of the controls used and supported the victim sequences as embedded within a group of sequences from the alleged source. Conviction of the perpetrator, however, depended also on circumstantial evidence regarding intentional infection. The phylogenetic analysis could not provide conclusive evidence of direct transmission, nor prove the direction of transmission, from sequence data alone. Nevertheless, there is concern that the existence of phylogenetic data could provide support for prosecutions or civil suits.

The existence of HIV criminalization [34] necessitates tight control over the use of data to minimize risks of misuse, and heightens the need for transparency and good communication, to allay fears about misuse and protect community relationships with public health agencies, health care providers, and health researchers.

Public health authorities have an ethical and legal mandate to collect communicable disease data, conduct disease prevention, and implement other public health programs. Public health surveillance has its own set of robust protections; data are closely held, kept confidential, and used only for public health purposes [28]. Even when public health authorities allow public health data to be used in research studies, the overall purpose of these studies must be to advance public health goals.

Sharing of public health data for research falls under state-specific laws that stipulate conditions for disclosure. Specific state laws, regulations, and policies address HIV surveillance data, with significant variability among states [27]. Many states give latitude to health authorities regarding sharing of data (see Box 1). Health departments often have strict policies regarding data sharing, may involve their legal counsel in responding to data requests, and may include protective measures. Current guidance on research use of HIV surveillance data indicates that such use should only be allowed for legitimate public health purposes supporting HIV prevention and treatment efforts within the jurisdiction. When health departments share data with the Centers for Disease Control and Prevention (CDC), these data are protected by the Assurance of Confidentiality under the Public Health Service Act [35].

In research, collection and use of data are governed by federal human subjects regulations, which mandate that, when appropriate, researchers keep identifiable information confidential [15]. In addition, all research data collected or used with research funding from the Department of HHS are covered under the Public Health Service Act, as recently amended by the 21st Century Cures Act (the Act), which mandates that Certificates of Confidentiality (CoC) apply to all research data collected with federal support [36] and prohibits researchers from disclosing data for nonresearch purposes such as civil or criminal procedures [37].

Under the Act, which modifies the Public Health Service Act §301(d) [38], all human subjects' research data in federally funded studies that involve "identifiable, sensitive information"

are protected by CoC; including data that are identifiable or *potentially identifiable*, even if lacking direct identifiers. Section 2012 of the 21st Century Cures Act defines *identifiable, sensitive information* as information about an individual collected or used in research *and* through which the individual is identified; *or* there is "at least a very small risk, as determined by current scientific practices or statistical methods, that some combination of the information, a request for the information, and other available data sources could be used to deduce the identity of an individual."

The Act states that the Secretary of HHS shall issue Certificates to persons engaged in biomedical, behavioral, clinical, or other research activities in which identifiable, sensitive information is collected. CoCs prohibit researchers from releasing research data to outside entities such as law enforcement, and protect the data from subpoena or disclosure, "in any Federal, State, or local civil, criminal, administrative, legislative or other proceeding" [39].

Furthermore, the Act applies to data collected previously and used in subsequent research projects, including use of public health surveillance data in research. To our knowledge, the current version of the Act has not been challenged in court; however, reviews of the earlier version of CoC, which was in effect from 1988 to 2015 [40], found that CoCs were rarely challenged in court. Other mechanisms to protect data from disclosure were often used. In some cases institutional legal counsel simply pointed out the existence of CoC and requests for data were dropped [41]. A recent commentary [42] advocated for better education of researchers and reviewers about CoC so that researchers are fully aware of their requirements.

Phylogenetic analysis has been used in studies involving serodifferent couples [43–45] to determine the genetic relatedness of HIV between these couples in cases of putative HIV transmission. Three partner studies reported that there were no phylogenetically linked transmissions from virally suppressed individuals to their HIV-negative partners. However, each study reported several unlinked transmissions, meaning HIV was acquired outside the participants' main partnership. Reporting phylogenetic results of linked/unlinked transmissions can create legal risks for study participants and their partners. Researchers conducting the Opposites Attract study developed a harm reduction approach to minimize risks of legal harm (Box 2). While not eliminating all risk, this approach helped reduce risk and reassure study participants that research data would not be used against them in legal proceedings.

***Recommendations***

Researchers, research sponsors, and oversight bodies should familiarize themselves with the CoC and associated policies. Researchers whose phylogenetic studies are not federally funded should seek coverage through the voluntary CoC procedure.

**BOX 2. RISK MITIGATION PROCEDURES IN THE OPPOSITES ATTRACT STUDY**

Opposites Attract was a study of 343 cisgender male couples in which 1 partner had human immunodeficiency virus (HIV) and the other did not. Prior to study implementation, Opposites Attract researchers recognized the possibility that study evidence, including phylogenetic data, could be used against study participants in criminal proceedings. The study team therefore sought legal guidance and subsequently developed 4 procedures to mitigate the risk of prosecution for study participants with HIV at baseline.

1. Participants who were HIV-negative at baseline were required to declare in writing that they had full knowledge of their primary partners' HIV-positive status. This requirement protects participants from accusations of HIV status nondisclosure, upon which many HIV criminalization laws are based.
2. All participants were required to successfully complete a knowledge examination on HIV transmission and prevention methods. This requirement provides evidence that HIV-negative partners shared the responsibility of informed sexual decision-making and that HIV prevention is not the sole responsibility of partners with HIV.
3. Only HIV-negative participants reported their sexual behavior on study questionnaires; participants with HIV never reported their sexual behavior. This data collection method prevents self-reported sexual behavior data from being used against participants with HIV. It also documents the sexual decision-making of HIV-negative participants, reinforcing the notion of shared responsibility in HIV prevention.
4. All participants were required to agree that under no circumstances would they or their clinicians have access to phylogenetic results in cases of seroconversion.

Source Bavinton et al [45].

Researchers, community advisory boards, and oversight bodies should consider the legal risks entailed in phylogenetic studies. Researchers should seek guidance from legal experts with regard to minimizing legal risks, and research teams should include someone knowledgeable in this area. When reporting results of phylogenetic analysis that could lead to inferences about individual transmission events, researchers should take steps to minimize risks of social harm or legal repercussions. At minimum, researchers should inform study participants about what information will be reported from the study, and legal and data security protections in place.

### Community Engagement

Community engagement, a cornerstone of HIV research, is used for several reasons, including assessing community knowledge and beliefs, providing information and education about research, identifying and mitigating risks, prioritizing research questions valued by the community, adapting study procedures to reflect community norms and practices, and developing and implementing plans for communication about research findings. Engagement with communities is also ethically important in light of past research abuses that fueled distrust in health research and medical care. These harms are compounded by systematic social injustice linked to racism, homophobia, violations of civil rights, lack of access to care, and stigma. Robust community engagement processes can help improve relationships and trust between researchers and communities participating in or affected by the research [46]. Given the sensitivity of HIV phylogenetic data, substantial community input in the research process is warranted.

Questions remain regarding the best way to conduct community engagement. Methods include community advisory boards, including community advocates on the research team, town hall meetings, stakeholder meetings involving community representatives and advocates, online dissemination of information [47–50] and collection of feedback, community participatory research, crowdsourcing [51], and other methods [52]. Previous scholarship has identified challenges in defining communities and their interests, identifying representative processes and structures, and determining when and how engagement is successful at accomplishing its ethical goals. This work is ongoing in many areas of research [53].

#### *Recommendations*

Community input should be elicited in all HIV phylogenetic research projects that involve potential impact on people with or vulnerable to HIV. Researchers should consider risks to groups as well as individuals. The extent and type of community input will vary depending on the project. In some cases, researchers will need to consult with stakeholders when planning a project to determine what groups could be affected by the research findings and, if so, how to engage them. Researchers should consider partnering with community-based organizations in this process. Risks related to stigma and social harm should be mitigated through planning with stakeholders who may include people with or vulnerable to HIV, community advocates, medical providers, community-based organizations, public health agencies, media outlets, educational organizations, or other groups.

Researchers and other stakeholders should engage relevant communities during planning and study design and not wait until the research is underway or completed. Early engagement provides an opportunity to work with communities on research question development, study design, and mitigating potential social harm. Community input is helpful when presenting and interpreting research findings to relevant stakeholders. In some cases, medical providers and community-based organizations use information from phylogenetic research studies to increase their outreach and education efforts in the community.

Community engagement methods should be designed based upon existing scholarship and models of engagement with consideration of the appropriateness of different methods of engagement for the proposed research [54].

### Communication

Communication and dissemination of results are activities distinct from community engagement. Dissemination of results of phylogenetics studies to stakeholders may be particularly challenging due the complex scientific concepts involved and lack of familiarity among the public. A recent study of stakeholder perceptions demonstrated substantial misunderstanding of phylogenetic research and its implications [10]. Communication strategies need to be tailored to different populations based on sociocultural, economic, and political contexts [55]. Effective communication and education are also important to mitigate potential social harms from misinterpretation of study results. Community engagement strategies can provide critical information for communication plans.

#### *Recommendations*

Researchers should consider implications of reporting findings, including what level of granularity to report with regard to transmission clusters, whether masking of identifiable characteristics is needed, as well as framing effects. (Framing effects is a term used in psychology that refers to the fact that people may respond to statements using a different choice of words, even when the basic information conveyed is the same. For example, a statement about a treatment with 50% mortality rate could be presented as a treatment that saves 50 lives out of every 100 patients treated, or a treatment in which 50 people will die for every 100 treated. Research in cognitive psychology has shown that the positive and negative frames affect how the information is perceived). For example, specific geographic areas or subpopulations might be stigmatized if high rates of HIV transmission are identified. Researchers need to understand the social and political implications of their findings and seek relevant input from stakeholders such as community representatives, policy makers, health communication experts, and journalists, regarding the best way to communicate findings. For example, presentation of HIV transmission network diagrams from HIV sequence data might be misinterpreted as representations of social or sexual network diagrams—an erroneous interpretation that could stigmatize specific groups. In some cases, it is appropriate to reduce the granularity of the observations to limit inferences in order to mitigate risks to

individuals, groups, and communities by, for example, aggregating count data, or reducing granularity of network visualizations by excluding small clusters or removing network links en masse, or other measures.

Researchers should disseminate research findings in accessible formats, such as social media, town hall meetings, or dissemination forums, in addition to publication in scientific journals and conferences. Social media can be used to provide information about new research findings for community groups or for the general public. Researchers should consult with community-based organizations about presentation of findings to lay audiences. Communicating research findings to community members and the public should be prioritized, particularly when findings can directly inform those who are working at the community level to prevent and treat HIV.

Research organizations and public health agencies should generate broad-based educational initiatives to communicate with the general public and communities affected by HIV about HIV phylogenetic analysis in public health and research. Members of the public are often unfamiliar with phylogenetic research and there is currently no lay educational program on the topic.

#### Areas for Further Empirical Research

The working group identified questions and gaps in the literature on ethical, social, behavioral, and legal aspects of HIV phylogenetic research. Further research is needed to address these gaps. The following research topics were identified.

1. Development and assessment of data security measures to determine the best methods for securing data; assessing the actual risks of breaches (in practice), the actual risks of reidentification (in practice); and evaluating methods to effectively and permanently deidentify or anonymize data so that reidentification is not possible.
2. Evaluating possible unintended consequences of phylogenetic research, including whether individuals are deterred from HIV testing or care due to fears about their test results being used in phylogenetic research or criminalization. Eliciting attitudes and beliefs among communities most affected by HIV about existing privacy protections for phylogenetic data.
3. Promote community engagement research to understand community knowledge, concerns, expectations, and attitudes regarding phylogenetic research, including informing previously unaware groups, similar to HIV preexposure prophylaxis (PrEP) [56] and HIV cure research [57]. Evaluate engagement strategies, assessing methods to provide education; addressing perceptions and misconceptions about the research; optimizing engagement methods to promote trust and trustworthiness among stakeholders, developing appropriate research protections, and promoting transparency in research [58]. Stakeholders include IRB members, public health researchers, and other groups as well as communities affected by HIV.
4. Develop and test best methods of dissemination and communication to communities and the public at large about results of HIV phylogenetic research findings.
5. Investigate the use of different sources of data in research (public health data, commercial data, public sequence data, etc.) and the utility and ethical implications of these sources.
6. Assess the use of phylogenetic methods to inform public health practice to determine if and to what extent their use has a beneficial effect, either at the individual or population level, to increase engagement in care and reduce HIV incidence, compared with more traditional epidemiologic measures and public health activities.
7. Research and scholarship on legal epidemiology initiatives, including exploration of legal and policy approaches that might help increase the recommended practices for phylogenetic research.
8. Evaluate the use of data use agreements to determine if they are effective in ensuring that data are used appropriately and not transferred to unauthorized users or used for unapproved purposes.

## CONCLUSION

These recommendations, developed by a multidisciplinary working group, are designed to promote ethical conduct of HIV phylogenetic research projects. The working group recommendations are not an official policy or regulatory approach, but rather suggested approaches aiming to be useful for researchers, oversight bodies, sponsors, community representatives, community-based organizations, and other stakeholder groups. As this field of research evolves and community-based responses change over time, these ethical considerations should be revisited. HIV phylogenetic research in jurisdictions outside the United States will also need to be explored in detail. The working group advocates for further empirical and scholarly research on these topics, to better inform research while serving the interests of communities and individuals affected by HIV.

### Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

## Notes

***Acknowledgment***. On behalf of the NIH Working Group on Ethical Issues in HIV Phylogenetic Research. The full list of members is given in the Supplementary Material.

***Disclaimer***. The views expressed in this paper are those of the authors, representing the work of the NIH Working Group on Ethical Issues in HIV Phylogenetic Research. The views expressed here do not represent any policy or position of the National Institutes of Health, the Department of Health and Human Services, the Department of Defense, or any of the component agencies of those departments.

***Financial support.*** Funding for the working group in-person meeting in April 2018 was provided by National Institutes of Health/National Institute of Allergy and Infectious Diseases Division of AIDS.

***Potential conflicts of interest.*** J. S., S. R., and N. B. report grant support from NIH. S. K. reports grant support from NIH and CDC. S. J. L. reports gr7ant support from NIH and Gilead. J. O. W. reports grant support from NIH, CDC, and Gilead. P. S. is an employee of the CDC. J. S. is a member of Merck KGaA's Bioethics Advisory Panel and Stem Cell Research Oversight Committee, a member of IQVIA's Ethics Advisory Panel, and consulting for Portola Pharmaceuticals, Inc., none of which are related to the material discussed in this paper. All other authors report no potential conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

## References

1. Ping LH, Jabara CB, Rodrigo AG, et al. HIV-1 transmission during early antiretroviral therapy: evaluation of two HIV-1 transmission events in the HPTN 052 prevention study. PLoS One **2013**; 8:e71557.
2. Peters PJ, Pontones P, Hoover KW, et al; Indiana HIV Outbreak Investigation Team. HIV infection linked to injection use of oxymorphone in Indiana, 2014-2015. N Engl J Med **2016**; 375:229–39.
3. France AM, Oster AM. The promise and complexities of detecting and monitoring HIV transmission clusters. J Infect Dis **2020**; 221:1223–5.
4. 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**; 14:89–96.
5. Oster AM, Wertheim JO, Hernandez AL, Ocfemia MC, Saduvala N, Hall HI. Using molecular HIV surveillance data to understand transmission between subpopulations in the United States. J Acquir Immune Defic Syndr **2015**; 70:444–51.
6. Wertheim JO, Chato C, Poon AFY. Comparative analysis of HIV sequences in real time for public health. Curr Opin HIV AIDS **2019**; 14:213–20.
7. Evans D, Benbow N. Ethical considerations for a public health response using molecular HIV surveillance data: a multi-stakeholder approach. Project Inform and Northwestern University, **2018**. https://digitalhub.northwestern.edu/downloads/1bff6ec0-af92-48dd-a8e2-d17e4ac4f1ec. Accessed 9 May 2020.
8. McClelland A, Guta A, Gagnon M. The rise of molecular HIV surveillance: implications on consent and criminalization [published online ahead of print 20 February 2019]. Critical Public Health doi: 10.1080/09581596.2019.1582755.
9. Gilbert M, Swenson L, Unger D, Scheim A, Grace D. Need for robust and inclusive public health ethics review of the monitoring of HIV phylogenetic clusters for HIV prevention. Lancet HIV **2016**; 3:e461.
10. Schairer C, Mehta SR, Vinterbo SA, Hoenigl M, Kalichman M, Little S. Perceptions of molecular epidemiology studies of HIV among stakeholders. J Public Health Res **2017**; 6:992.
11. Ratmann O, Kagaayi J, Hall M, et al. Quantifying HIV transmission flow between high-prevalence hotspots and surrounding communities: a population-based study in Rakai, Uganda. Lancet HIV **2020**; 7:e173–83.
12. Zhang Y, Wymant C, Laeyendecker O, et al. Evaluation of phylogenetic methods for inferring the direction of HIV transmission: HPTN 052 [published online ahead of print 10 January 2020]. Clin Infect Dis doi: 10.1093/cid/ciz1247.
13. Taylor B, Sapién H. Determining the direction of HIV transmission: benefits and potential harms of taking phylogenetic analysis one step further [published online ahead of print 10 January 2020]. Clin Infect Dis doi: 10.1093/cid/ciz1248.
14. National Institutes of Health. NIH data sharing policy and implementation guidance, **2003**. https://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm#hs. Accessed 5 June 2020.
15. Protection of human subjects. 45 CFR 46. Fed Regist **2017**; 82:7149–274.
16. Mehta SR, Vinterbo SA, Little SJ. Ensuring privacy in the study of pathogen genetics. Lancet Infect Dis **2014**; 14:773–7.
17. Sweeney P, Gardner LI, Buchacz K, et al. Shifting the paradigm: using HIV surveillance data as a foundation for improving HIV care and preventing HIV infection. Milbank Q **2013**; 91:558–603.
18. Lee LM, Heilig CM, White A. Ethical justification for conducting public health surveillance without patient consent. Am J Public Health **2012**; 102:38–44.

19. Department of Health and Human Services, Office for Human Research Protections. 45 CFR 46. https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html. Accessed 5 June 2020.
20. Federal policy for the protection of human subjects. 82 FR 7149, **2017**. https://www.federalregister.gov/documents/2017/01/19/2017-01058/federal-policy-for-the-protection-of-human-subjects. Accessed 5 June 2020.
21. World Health Organization. WHO guidelines on ethical issues in public health surveillance. Geneva, Switzerland: WHO, **2017**.
22. National Institutes of Health. NIH sharing policies and related guidance on NIH-funded research resources. https://grants.nih.gov/policy/sharing.htm. Accessed 5 June 2020.
23. Bill and Melinda Gates Foundation. Bill and Melinda Gates Foundation open access policy. https://www.gatesfoundation.org/How-We-Work/General-Information/Open-Access-Policy. Accessed 5 June 2020.
24. Hamp AD, Doshi RK, Lum GR, Allston A. Cross-jurisdictional data exchange impact on the estimation of the HIV population living in the District of Columbia: evaluation study. JMIR Public Health Surveill **2018**; 4:e62.
25. Ocampo JMF, Smart JC, Allston A, et al. Improving HIV surveillance data for public health action in Washington, DC: a novel multiorganizational data-sharing method. JMIR Public Health Surveill **2016**; 2:e3.
26. National Center for Biotechnology Information. GenBank. https://www.ncbi.nlm.nih.gov/genbank/. Accessed 8 May 2020.
27. National Alliance of State and Territorial AIDS Directors. HIV data privacy and confidentiality: legal and ethical considerations for health department data sharing, **2018**. https://www.nastad.org/resource/hiv-data-privacy-and-confidentiality. Accessed 9 May 2020.
28. Centers for Disease Control and Prevention. Data security and confidentiality guidelines. https://www.cdc.gov/nchhstp/programintegration/docs/PCSIDataSecurityGuidelines.pdf. Accessed 9 May 2020.
29. Center for HIV Law and Policy. HIV criminalization in the United States: a sourcebook on state and federal HIV criminal law and practice, **2020**. https://www.hivlawandpolicy.org/sourcebook. Accessed 8 May 2020.
30. Latham SR, Polaris J. HIV criminalization by state. Law Atlas Policy Surveillance Program, **2013**. http://lawatlas.org/datasets/hiv-criminalization-statutes. Accessed 8 May 2020.
31. Lazzarini Z, Galletly CL, Mykhalovskiy E, et al. Criminalization of HIV transmission and exposure: research and policy agenda. Am J Public Health **2013**; 103:1350–3.
32. Galletly CL, Pinkerton SD. Conflicting messages: how criminal HIV disclosure laws undermine public health efforts to control the spread of HIV. AIDS Behav **2006**; 10:451–61.
33. Harsono D, Galletly CL, O'Keefe E, Lazzarini Z. Criminalization of HIV exposure: a review of empirical studies in the United States. AIDS Behav **2017**; 21:27–50.
34. Sweeney P, Gray SC, Purcell DW, et al. Association of HIV diagnosis rates and laws criminalizing HIV exposure in the United States. AIDS **2017**; 31:1483–8.
35. Centers for Disease Control and Prevention. Assurances of confidentiality. https://www.cdc.gov/od/science/integrity/confidentiality/. Accessed 9 May 2020.
36. National Institutes of Health. Notice of changes to NIH policy for issuing certificates of confidentiality. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-17-109.html. Accessed 9 May 2020.
37. Wolf LE, Patel MJ, Williams Tarver BA, Austin JL, Dame LA, Beskow LM. Certificates of confidentiality: protecting human subject research data in law and practice. J Law Med Ethics **2015**; 43:594–609.
38. Cornell Law School, Legal Information Institute. 42 U.S. Code § 241. https://www.law.cornell.edu/uscode/text/42/241. Accessed 20 May 2020.
39. 21st Century Cures Act, Pub L No. 114–255, §2012 (**2016**).
40. Wolf LE, Beskow LM. New and improved? 21st century cures act revisions to certificates of confidentiality. Am J Law Med **2018**; 44:343–58.
41. Wolf LE, Dame LA, Patel MJ, Williams BA, Austin JA, Beskow LM. Certificates of confidentiality: legal counsels' experiences with and perspectives on legal demands for research data. J Empir Res Hum Res Ethics **2012**; 7:1–9.
42. Wolf LE, Beskow LM. Genomic databases, subpoenas, and certificates of confidentiality. Genet Med **2019**; 21:2681–2.
43. Cohen MS, Chen YQ, McCauley M, et al; HPTN 052 Study Team. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med **2011**; 365:493–505.
44. Rodger AJ, Cambiano V, Bruun T, et al; PARTNER Study Group. Sexual activity without condoms and risk of HIV transmission in serodifferent couples when the HIV-positive partner is using suppressive antiretroviral therapy. JAMA **2016**; 316:171–81.
45. Bavinton BR, Jin F, Prestage G, et al. The opposites attract study of viral load, HIV treatment and HIV transmission in serodiscordant homosexual male couples: design and methods. BMC Public Health **2014**; 14:917.
46. MacQueen KM, Bhan A, Frohlich J, Holzer J, Sugarman J; Ethics Working Group of the HIV Prevention Trials Network. Evaluating community engagement in global health research: the need for metrics. BMC Med Ethics **2015**; 16:44.

47. The Legacy Project. MHS phylogenetics series. Part 1. Without our consent: centering people living with HIV in HIV genetic sequencing. https://fredhutch.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=cf3bfa85-c901-40ee-a87a-aa1e00fb8d3a. Accessed 20 May 2020.
48. The Legacy Project. MHS phylogenetics series. Part 2. These trees cast no shade: understanding HIV phylogenetics research. https://fredhutch.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=f96d8e7d-90d9-4f0b-8b98-aa33013572c2. Accessed 20 May 2020.
49. The Legacy Project. MHS phylogenetics series. Part 3. Understanding molecular HIV surveillance: public health and human rights implications. https://fredhutch.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=d793f67f-4fc2-43b9-91b3-aa3a0185d57e. Accessed 20 May 2020.
50. The Legacy Project. MHS phylogenetics series. Part 4. cluster tracking is people: community town hall on molecular HIV surveillance. https://fredhutch.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=f01c3966-0587-4662-9182-aa4e016a6967. Accessed 20 May 2020.
51. Mathews A, Farley S, Blumberg M, et al. HIV cure research community engagement in North Carolina: a mixed-methods evaluation of a crowdsourcing contest. J Virus Erad **2017**; 3:223–8.
52. Community Partners. Recommendations for community engagement in HIV/AIDS Research. A guide for communities and researchers, version 2.0, **2014**. https://www.hanc.info/cp/resources/Documents/Recommendations%202014%20FINAL%206-5-14%20rc.pdf. Accessed 9 May 2020.
53. MacQueen KM, Eley NT, Frick M, Hamilton C. Using theory of change frameworks to develop evaluation strategies for research engagement: results of a pre-pilot study. J Int AIDS Soc **2018**; 21:e25181.
54. Corbie-Smith G, Wynn M, Richmond A, et al. Stakeholder-driven, consensus development methods to design an ethical framework and guidelines for engaged research. PLoS One **2018**; 13:e0199451.
55. Rimer BK, Glanz K. Theory at a glance: a guide for health promotion practice. Bethesda, MD: Department of Health and Human Services, National Institutes of Health, National Cancer Institute, **2005**.
56. Auerbach JD, Kinsky S, Brown G, Charles V. Knowledge, attitudes, and likelihood of pre-exposure prophylaxis (PrEP) use among US women at risk of acquiring HIV. AIDS Patient Care STDS **2015**; 29:102–10.
57. Sylla L, Evans D, Taylor J, et al. If we build it, will they come? Perceptions of HIV cure-related research by people living with HIV in four U.S. cities: a Qualitative Focus Group Study. AIDS Res Hum Retroviruses **2018**; 34:56–66.
58. Fisher CB. HIV prevention research ethics: an introduction to the special issue. J Empir Res Hum Res Ethics **2014**; 9:1–5.