Ethics considerations for HIV phylogenetic analyses
Contents

2 Acknowledgments
3 Introduction
5 Ethical questions presented by HIV phylogenetic analyses
  5 Issues relating to consent
  6 Issues relating to participant privacy and data security
  6 Issues relating to results of analyses
  7 Issues relating to human rights violations
8 Minimising risks
  8 Informed consent
  9 Results of analyses
10 HIV Phylogenetics Ethics checklist
12 HIV phylogenetics in public health surveillance
14 Circumstances in which HIV phylogenetic studies should not take place
15 Conclusions
16 References
Acknowledgments

UNAIDS gives special thanks to Lucie Abeler-Dörner, Ana Bulas Cruz, Euzebiusz Jamrozik, Michael Parker, Janet Seeley, Josh Herbeck, David Bonsall, Chris Wymant, Matthew Hall, Christophe Fraser, for their contributions to this publication.
Introduction

Ethics is a discipline or area of study concerned with issues such as benefits, harms, fairness, and rights. As an area of study and practice, ethics covers moral values and questions about what actions should be taken and which policies should be implemented. Ethical answers to these questions depend on what (beneficial or harmful) outcomes are likely to occur, but also the extent to which these outcomes will be distributed fairly through relevant populations, and the rights and interests of the people affected by actions and policies. This is the third brief in a set discussing phylogenetics and HIV in public health will discuss ethics and human rights considerations. Throughout, we keep in mind Article 8 of the Helsinki Declaration: While the primary purpose of medical research is to generate new knowledge, this goal can never take precedence over the rights and interests of individual research subjects.

One of the most familiar representations of this discipline in science is research ethics committees (or institutional review boards) who make judgements about scientific studies, including HIV genomic and phylogenetic research. The main roles of ethics committees are to determine whether, on balance, the expected benefits of a specific scientific study under review outweigh the expected harms (i.e., risks), and whether the research plan will be carried out with adequate respect and safeguards for participants’ dignity, well-being and bodily autonomy, including via seeking their informed consent for participation, as well as in accordance with relevant research ethics guidelines.

The ability to study and compare the genetic make-up of pathogens has revolutionised infectious disease biology. The genetic building plan or genome of each organism can be read by sequencing. Phylogenetic techniques can then be used to compare the genomes of different organisms and draw conclusions about how closely they are related. Phylogenetic analyses based on pathogen sequencing have made it possible to determine if pathogens found in two individuals are closely related to each other or not, thus allowing to learn more about how pathogens travel between human and animal hosts and how they change in the process. The basics of genetics, phylogenetics and sequencing are covered by the first brief in this series.

HIV phylogenetic analyses have been used in research settings since the beginning of the epidemic. HIV phylogenetics can be used in many ways to better understand the HIV epidemic and the risks of transmission in different settings, and to help focus interventions where they are most needed. HIV phylogenetics can reveal how HIV spreads in a population. One of the best-known examples of an HIV phylogenetic analysis is the HPTN 052 trial and the Partners and Opposite Attract studies [1–3]. These studies showed that both same-sex and opposite-sex partners of a person living with HIV very rarely acquired HIV from their partner if the partner was undergoing antiretroviral therapy (ART). The studies established that U=U (undetectable equals untransmissible)—a person on ART with an undetectable viral load is no longer infectious. The impact of HIV phylogenetic studies in public health is discussed in the second brief of this series.
Some HIV phylogenetic analyses can generate sensitive information regarding people living with HIV. This means that, if not well-understood and handled carefully, phylogenetic results have the potential to harm individuals and groups. The use of phylogenetics in HIV requires a strong ethical and rights-based framework that facilitates their use to prevent transmission and inform prevention strategies while respecting the rights and interests of individuals and minimising any risks.

Most phylogenetic analyses have to date been conducted in research settings. In the last few years, however, HIV genomic data has also increasingly been used by public health agencies in high-income countries. Public health agencies should be guided by similar ethical principles as those in public health research, but governance frameworks for public health agencies are different to those for researchers. For example, while research studies always require informed consent and ethics review, most countries permit some public health uses of data without informed consent or ethical review. This brief, the third in a series of three, discusses the ethical issues arising from HIV phylogenetic analyses in both research and public health settings and offers guidance on how to minimise risks for analyses that generate sensitive data.
Ethical questions presented by HIV phylogenetic analyses

Careful attention needs to be paid to a range of important ethical and human rights questions to use phylogenetic analyses to gain useful information while minimising associated risks. Many of the ethical issues presented by the uses of genomics more broadly in healthcare and medical research across different diseases have been discussed in the literature. These include: questions relating to the requirements for valid informed consent; issues relating to the feedback of findings; the need to protect individual privacy; potential stigmatisation and other effects on groups or communities; and practical ethical considerations relating to the collection, storage, sharing and appropriate governance of genomic and other data. There has also been significant discussion of the context-specific ethical issues arising in genomic research in low and middle-income countries (LMICs), including the critical role of community engagement, the need for local capacity building, and the importance of fair research partnerships [4–7].

Several studies have highlighted the ethical issues arising from the use of genomic approaches to understanding HIV [8–11]. These issues mostly relate to analyses that use the rich data generated by deep-sequencing and combine these with epidemiological data to study transmission, including the reconstruction of transmission networks and transmission pairs. Use and mis-interpretation of the strength of results have the potential to cause harm to people.

In jurisdictions where HIV transmission, same-sex sexual activity or sex work are criminalised, identification of key populations has led to their arrest and harassment by police, restriction of access to housing, schooling and employment, and acts of violence, including murder [8]. Phylogenetic analyses which are aiming to identify key populations need to include an assessment of risk for human right violations to take place. Phylogenetic analyses have also been used in several high-profile court cases, e.g., [12–14]. While phylogenetic analyses can show beyond reasonable doubt that one person did not infect another person, it cannot be used to prove that one person did infect another person [8,15,16].

The following paragraphs outline the main ethical issues relating specifically to HIV. The next section discusses possible risk mitigation measures. The focus will be on HIV phylogenetic analyses in research studies, with a section discussing the use of HIV phylogenetics in public health settings.

Issues relating to consent

- The technology and its implications are complex and achieving adequately informed consent can be challenging for both participant and staff carrying out the consent interviews.

- Phylogenetic analyses are relational by nature, i.e., they look at the relationship of samples to each other, which can reveal highly personal information, for example
regarding sexual relationships. This is different from other types of research and can be unintuitive for participants.

- Analyses could lead to, or reinforce stigmatisation, of groups or populations even if they are not part of the study.

- Collecting samples for both clinical use (for example, for viral load and drug resistance testing) and research use potentially complicates the consent process. For example, access to a clinical diagnostic must not be contingent on an individual’s consent to research.

- Research datasets may not be representative of all people living with HIV if certain groups have more concerns about the risks of such analyses and are therefore less likely to participate.

- Phylogenetic techniques are improving with time. If consent has been granted for future research use, new analyses techniques might reveal information that was not possible to obtain at the time consent was sought.

**Issues relating to participant privacy and data security**

- Since HIV mutates in the human host, the population of viruses in a particular host is unique for each person living with HIV, like a fingerprint. This ‘HIV genetic fingerprint’ is stable for a few years which means that two samples from two different studies could be shown to have come from the same person. It may be possible to re-identify individuals from anonymized datasets which makes perfect anonymization very difficult.

- Phylogenetic analyses may reveal highly personal information, for example likely sexual relationships, including extra-marital relationships or whether one person was likely to be taking effective antiviral treatment at a particular time.

- Use of data applying the rules of data minimalism can be complex, that is data should be shared only on a need-to-know basis. For example, clinicians do not need access to transmission networks or social analyses, and researchers in public health do not require participant-identifiable metadata.

**Issues relating to results of analyses**

- Information generated through phylogenetic analysis can illuminate patterns of transmission between populations and can therefore also have implications for relations between and perceptions of groups.

- The raw data are rich and, if stored, future analyses may uncover sensitive information about individuals or groups that was not possible at the time the sample was collected. For example, while current analytical methods infer the directionality and directness of transmission with a degree of uncertainty, the bounds of uncertainty may decrease with improved methods or improved understanding of HIV transmission.

- The results of phylogenetic studies could be misused to increase stigma and discrimination for particular groups or criminalisation of specific activities that may be associated with HIV transmission.
- Phylogenetics can identify key transmission groups, but it cannot uncover the existing structural issues that might be driving HIV transmission.

- Results that suggest that particular groups or types of individuals are more likely to be the source of HIV transmission may increase stigma and/or undermine trust in research.

**Issues relating to human rights violations**

- In countries which criminalise HIV transmission, same-sex sexual activity, sex work or use of intravenous drugs, data from phylogenetic studies could be subpoenaed to legally challenge participants if adequate data protections are not in place.

- The results of phylogenetic studies could be misused to criminalise particular groups for specific activities that may be associated with HIV transmission.
Minimising risks

In some cases, researchers or ethics committees might decide that the best way to avoid a risk associated with a certain phylogenetic analysis is to not perform the analysis at all. For analyses that proceed, efforts should be made to eliminate or mitigate risks. A careful risk–benefit assessment may help to ensure that the benefits outweigh the risks, and that the latter are minimised as much as possible. This assessment should address whether other methods with fewer risks can be achieve similar results, whether benefits accrue to the community equally, and whether risks are too high even if benefits outweigh them.

Informed consent

Information in the consent process needs to be comprehensive and understood by participants. Where consent is sought for phylogenetic studies, the information provided should be as specific as required, and include relevant uncertainties as well as details of data protections and governance mechanisms in place. Where appropriate, tests of comprehension should be used to ensure that the consent process has met these standards.

Some participant concerns may be addressed in discussions during the consent process. Prior community engagement should be used to identify issues raised by relevant populations and prospective participants so that this can shape the research design including the consent.

Where consent includes future use of the data, which might include analyses that are not possible at the time consent it sought, this needs to be pointed out to participants. Where future studies include additional risks, advice from ethics committees should be sought even if the new analysis is covered by existing consent. Researchers and public health authorities should also keep in mind that informed consent alone cannot guarantee the protection of participants.

- Participant privacy and data security. Utmost care should be taken to ensure participant privacy and secure storage of data, with strong protections against accidental release of data. Risk mitigation strategies in this area include:
  - Anonymisation or pseudonymisation of data
  - Storage of data on a secure server with restricted access
  - Storage of data in a jurisdiction that does not criminalize HIV transmission, same-sex sexual activity, illicit injecting drug use and sex work
  - Storage of data by an entity or organisation that is required by law not to release data for purposes other than for which it was collected
  - Storage of data in separate databases and reducing linkage between HIV genomic data and other data (e.g., socioeconomic metadata)
Data sharing agreements that ensure only the minimum required data are shared for a given purpose and with appropriate security standards

Limiting publicly available data to low resolution (e.g., only consensus sequences)

Removing any traces of human genomes or the genomes of non-HIV pathogens before analysis (unless specifically required by the study in question and mentioned explicitly in the consent procedure)

Ethical appraisal by all parties involved, including the study communities can reach consensus on who should and should not have access to specific types of data, and this can inform systems and information governance frameworks that mitigate risks appropriately.

Results of analyses

Risks to individuals or groups arising from the results of population-wide phylogenetic analyses are the hardest to mitigate against, as they are often inseparable from the benefits arising from the analysis. For example, identifying a high level of transmission in a certain age group or in a certain geographical area might increase stigma against the individuals and groups concerned, but also open the way for better prevention strategies for these groups and their contacts. All significant residual risks must be minimised.

The participation of relevant local stakeholders, e.g., via community engagement, is key to discussing concerns with members of the study communities and communicating results in a non-stigmatising and culturally competent way before they are published elsewhere. The scientific presentation of phylogenetic results must involve care to preserve the privacy of individual participants. Population-wide phylogenetic analyses present results averaged over groups rather than for individual participants, and care must be taken to choose groups in a way that does not allow the identification of individual participants. In large international collaborations, local researchers should have leading and active roles throughout the process. This can help minimise harm and maximise benefits as local researchers are more likely to understand the specificities of the local context and health care system and can help translate research into effective policy more easily. Combining phylogenetic analyses with social science research is also crucial in making sure that the results are translated into practical, useful information and that any potential risks are mitigated.
The bioethics literature has identified principles for ethically acceptable HIV phylogenetic research [8–11]. The list below is modified from the list of ethical considerations presented by Coltart et al. [8].

**Risk–benefit analysis.** A careful risk–benefit analysis is needed to identify potential risks to individuals and groups for a given study and weigh them against the benefits. Risks can arise to individuals or to groups. This exercise should be carried out for individuals and on group-level for each of the stakeholder groups involved.

**Development of risk mitigation strategies.** Risks should be eliminated wherever possible. For risks that cannot be eliminated, risk mitigation strategies might include data anonymisation, secure data access policies, community consultations and training researchers to communicate results in a way that minimises risks to individuals and groups.

**Awareness of the social and legal context.** Risk–benefit analyses and risk mitigation strategies need to be study-specific and tailored to the social and legal context of those study communities. Research that is acceptable in one case might, for example, exacerbate existing social tensions in another case.

**Protection of human rights.** The potential for infringements of human rights is significantly increased in countries where HIV transmission, same-sex sexual activity, and/or sex work are criminalized. Many other social and economic factors may make some individuals particularly vulnerable to such infringements. The risk of rights infringements should always be assessed and if the risk is too high for one of the groups, analyses should not take place.

**Public engagement.** Public engagement with the community and wider public is needed to facilitate the involvement of community members and local stakeholders in the design and implementation of models of good practice, and to communicate the risks and benefits of specific studies. The views of members from the study community and the wider public should be sought in advance and should ideally inform the design and conduct of specific studies. Community engagement can be challenging for highly technical research plans, but it is the duty of the researchers to communicate their plans clearly and effectively. When communicating results, community representatives (e.g., community advisory board members) can advise on how to reduce the risk of adverse effects such as stigma. Close work with community representatives is key to conducting successful phylogenetic research which maximises the public health benefit and minimises the risk to participants and all groups represented in the study communities.

**Informed consent and regulatory approvals.** Informed consent is required to ensure that participants have made a voluntary choice to be involved based on a good understanding of how data will be collected, stored, used for research, or shared with others. They should also be aware of the structures in place for the oversight of access to data, including for future research uses. Participants should know which ethics boards have assessed the overall risks to individuals and groups. Informed consent should be as specific to the study as possible and clearly explain which analyses will be carried
out, and what the constraints are on any future unknown uses. It should also inform participants of their right to withdraw. This is a challenge for biobank-type studies where a phylogenetic component might only be a small part of the research program. It is also a challenge when using data from previous studies, where the analysis might only be acceptable if participants gave broad consent for future uses of their data. For these studies, review and authorisation by local ethics committees is particularly important. Since phylogenetics is a relatively new and fast evolving field, focused and continuous training can help ensure that ethics committees can appropriately review studies, protecting participants and assuring accountability without blocking or delaying beneficial phylogenetics studies.

Protection of rights and interests of study participants. The privacy of participants needs to be protected and benefits need to outweigh risks in their perception as well as in the perception of the study organisers. Participants should feel that ‘enough is in there for them’, be it clinically relevant data or a public health aim they can identify with.

Secure data storage. Data must be stored securely to minimise the risk of accidental data leaks and ensure that only ethically approved analyses are being conducted. Secure databases need to be supplemented by a solid framework including training for researchers in the ethical use of data. Researchers should take personal responsibility for conducting only analyses that do not violate the agreed framework. Data should not be released for non-approved purposes.

Equitable data generation and analysis. Currently, sequencing technology, phylogenetic expertise and computing equipment are more concentrated in the high-income countries. For studies carried out in low- and middle-income countries, sequencing and analysis should be performed locally as much as possible and, where needed, these activities should incorporate training to build local capacity.

Return of results. Clinically relevant results should be returned to ensure that participants benefit during their individual care where possible. Previously, phylogenetic studies often had long turn-around times that made information no longer useful for the clinical care of participants. Recent technological developments have enabled a much faster time to results, of weeks rather than months. It has thus become increasingly ethically appropriate to return data relevant to clinical treatment, such as drug resistance profiles, to study participants wherever feasible. Procedures for the return of results should be locally acceptable and mitigate risks—among other things, by excluding clinically irrelevant but potentially sensitive information available to researchers (e.g., any information on sexual networks).

Communication of study results. The broad scientific findings of phylogenetic studies should ideally be presented to local communities and/or available to participants before wider publication. Potentially sensitive results should be discussed in advance with representatives of different groups in the communities. Care should be taken to communicate results through scientific and non-scientific channels in a way that minimises the chance of stigma and averts potential conflict.

Equitable data sharing. Patient privacy must not be compromised, while at the same time as much as possible of the data should become available to the research community to maximise the benefits of the study. Giving researchers, especially locally based researchers, access to the data collected is important to increase the usefulness of the data. However, guidelines on participant privacy, minimising risk to individuals and communities, and appropriate oversight mechanisms must always be followed. This includes obtaining the explicit consent of participants for data sharing as part of the informed consent process. If research cannot take place within the guidelines, it should not be carried out.
HIV phylogenetics in public health surveillance

Public health surveillance is the routine collection, analysis and use of data on health conditions put to use for public health programming. As a government-led activity which often collects sensitive data, working within a strong governance and ethical framework is necessary. A comprehensive overview is provided by the WHO guidelines on ethical issues in public health surveillance [17].

Several countries use phylogenetics or clustering techniques for public health purposes [18–21]. Since 2019, USA public health agencies have used HIV genomic clustering approaches for public health surveillance as part of the programme Ending the HIV Epidemic in the U.S. [22,23]. The aim is to detect fast-growing clusters quickly and help the affected individuals to prevent onward transmission. The US programme has led to a discussion about consent and use of personal data [24,25]. While public health use of personal data does not always require consent, the use of sensitive HIV data in this context initially led to calls for such practices to be halted and further consultation to take place [26]. The current debate centres on topics including consent, data security and patient privacy, community engagement, human rights and racial inequality concerns, programme evaluation, and tailoring the programme to local jurisdictions in order to prevent use of the data by law enforcement, including not rolling out HIV molecular surveillance programme in jurisdictions where this cannot be guaranteed [27,28]. The US programme uses clustering methods as a mitigation strategy because it is unsuitable for determining the direction of transmission. The choice of such methods can help to strike a balance between protecting individual patient privacy and being able to use such data for public health programmes.

There are currently no programmes involving the use of HIV phylogenetic data for public health surveillance in low and middle-income countries, but investment in genomic sequencing capacity during the COVID-19 pandemic may lead to an increase in molecular surveillance programmes for endemic diseases worldwide.

The use of HIV phylogenetic data in public health practice poses additional challenges beyond those in HIV phylogenetic research [29] and requires a strong ethical framework:

- The aims, methods and data security policies of such programmes must be transparent.
- Full informed consent should be sought wherever feasible. Patients should have the opportunity to opt out of phylogenetic linkage without being denied other services.
- The programme must be accompanied by an information campaign regarding HIV molecular surveillance.
- Patient data should be anonymised during phylogenetic analyses. Genomic and other identifiable data should only be used by authorised and trained staff, where required for public health purposes, and only the minimum identifiable information needed should be used.
There must be a clear data governance framework that preserves patient privacy as much as possible.

The programmes are most likely averting infections if participants are safe from prosecution for HIV transmission, same-sex sexual activity, sex work, or drug use.

Great care should be taken not to alienate patients from services that are vital to their well-being.

The benefits and any harms of the programme must be regularly assessed and should be publicly reported.

Public health bodies should see an HIV molecular surveillance programme as an opportunity to engage with people living with HIV and ensure that people living with HIV can actively take part in policy formation. Agencies should be able to demonstrate that the programme can avert infections, find gaps in the health care system, and give more people the chance to live without HIV. This requires the provision of training for health care providers to counsel and answer questions from patients whose data are used for public health purposes.
Circumstances in which HIV phylogenetic studies should not take place

Phylogenetic techniques are tools and, as is the case for all tools, whether they should be employed or not hinges on the balance of expected benefits and risks and whether risks, once minimised, are acceptable. Analyses that are likely to produce significant harm should not be carried out or funded. All other analyses require a careful risk–benefit analysis which considers the positive outcomes of the analysis and devises strategies to minimise any potential risks. Alternative methods to answer the same question, which may eliminate or better mitigate potential risks, must be assessed. If overall risks or risks to a specific group are unacceptably high (as assessed by relevant communities, researchers, or the public health authorities) a study or public health programme should not take place. Risk and benefits depend on the social context. Analyses that are deemed acceptable in one community or location might increase existing social tensions in another context. Risk assessment should therefore always be study specific.

It is very important to have an ongoing discussion with, and genuine involvement of, the members of the communities in which studies are taking place. The ethical acceptability of a decision to carry out a particular study will depend ultimately on its acceptance in the community, the voluntary, informed, and competent consent of community participants, and its conformity with relevant guidelines and regulations.
HIV phylogenetics is a powerful set of tools which, when used correctly and ethically, provides knowledge about HIV epidemics that can inform policies and help reduce the spread of HIV. Some phylogenetic analyses generate highly sensitive participant data, and use of HIV phylogenetics in public health raises additional issues to those in pure research which are more difficult to address, for example regarding questions of consent. Traditionally, public health measures can often be implemented without individual consent. Given the sensitivity of information revealed by some phylogenetic analyses, consent for this kind of molecular surveillance should also be sought in public health contexts if at all feasible. In some circumstances alternative methods, such as cluster analyses, which generate less sensitive data but offer most of the benefits of phylogenetic analyses in a public health context might be a suitable alternative for national molecular surveillance programmes.

Evaluation of currently existing programmes will allow a better quantification of benefits and harms and should inform future implementations of molecular surveillance by public health bodies. A future focus should be the reinforcement and updating of existing governance mechanisms, an ongoing engagement with all stakeholders and relevant communities, and an emphasis on equitable resourcing, analysis and data sharing.

A strong ethical framework is required to ensure that participants can give truly informed consent, participant privacy and data security are guaranteed at all stages of the project, and great care is taken only to undertake analyses in which the benefits to public health and the participants and the members of their community outweigh the risks.
References


